

AAPS Advances in the Pharmaceutical Sciences Series 47

Seetharama D. Jois *Editor*

Peptide Therapeutics

 aaps®

 Springer

AAPS Advances in the Pharmaceutical Sciences Series

Volume 47

Series Editor

Yvonne Perrie, Strathclyde Institute of Pharmacy and Biomedical Sciences
University of Strathclyde, Glasgow, UK

The AAPS Advances in the Pharmaceutical Sciences Series, published in partnership with the American Association of Pharmaceutical Scientists, is designed to deliver volumes authored by opinion leaders and authorities from around the globe, addressing innovations in drug research and development, and best practice for scientists and industry professionals in the pharma and biotech industries. Indexed in

Reaxys

SCOPUS Chemical Abstracts Service (CAS) SCImago

EMBASE

Seetharama D. Jois

Editor

Peptide Therapeutics

Fundamentals of Design, Development,
and Delivery

 Springer

 **aaps**[®]

Editor

Seetharama D. Jois
School of Basic Pharmaceutical and Toxicological Sciences
University of Louisiana at Monroe, College of Pharmacy
Monroe, LA, USA

Editorial contact

Carolyn Spence

ISSN 2210-7371

ISSN 2210-738X (electronic)

AAPS Advances in the Pharmaceutical Sciences Series

ISBN 978-3-031-04543-1

ISBN 978-3-031-04544-8 (eBook)

<https://doi.org/10.1007/978-3-031-04544-8>

© American Association of Pharmaceutical Scientists 2022

Dual Imprint with AAPS Springer

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publishers, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publishers nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publishers remain neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

It is estimated that by 2024, the projected market for peptide therapeutics will be around \$50 billion. Thus, a book that explains with fundamental principles how the peptide-based drugs are designed and what are the steps needed to develop a peptide-based therapeutic is essential. Peptides are natural ligands for many receptors in the human body, and many, such as hormones, neurotransmitters, and growth factors, have physiological roles. Historically, peptide drugs were originated from naturally occurring ligand peptides that were modified for therapeutic purposes. However, advances in synthetic chemistry of peptides, more available raw material for synthesis, and efficiency of peptide synthesis on an industrial scale have made the synthetic peptide market successful. At present, peptide-based therapeutics are in demand in two main areas: metabolic diseases and oncology. More than 80 peptide-based drugs approved by the FDA are now on the market, and more than 150 types of therapeutic peptides are in the pre-clinical stage.

The book covers the design concept of peptide therapeutics from fundamental principles using structural biology and computational approaches. The chapters are arranged in a linear fashion. A fresh science graduate or a scientist who works on small molecules can follow the design and development of peptide therapeutic to understand the basic concepts. Each chapter is written by experts from academia as well as industry. Rather than covering extensive literature, first few chapters of the book provide fundamental concepts of design, synthesis, and computational aspects of peptide therapeutics. Chapters on peptide delivery and stability cover basic as well as recent literature applicable to peptide drugs. In terms of regulations to approve the peptide drugs for therapeutic purposes, there were no special FDA regulations a decade ago. Larger peptides were treated as proteins, and smaller peptides were treated as small molecules. In 2013, the United States Pharmacopeial Convention (USP) formed a therapeutic peptide expert panel for guidance on peptide drugs, and FDA's Center for Drug Evaluation and Research (CDER) released a draft guideline that specifically addresses peptide drugs. With this new development, there are guidelines for peptide therapeutics to bring it to the market. The last chapter covers regulatory issues related to peptide therapeutics and some new developments in FDA regarding peptide drugs. The examples provided in the chapters

can be used to perform peptide formulation considerations for the designed peptides. Some chapters are written as a combination of basic principles and protocol so that scientists can adopt these methods to their research work. The book has nine chapters, and each chapter can be read as an independent chapter on a particular concept. The book can be used as a reference for a pharmaceutical or biomedical scientists or graduate students who want to pursue their career in peptide therapeutics. Most of the academic labs work with the design, synthesis, and structure-activity relationship of peptides. These scientists collaborate with drug delivery scientists to propose peptide delivery. However, this book comprehensively provides all the concepts for a reader. The editor would like to thank the contributors as well as reviewers for their time and effort. The editor would also like to thank Mr. Prajesh Shrestha, a graduate student in the School of Basic Pharmaceutical and Toxicological Sciences, College of Pharmacy, University of Louisiana at Monroe, for helping with final editing and formatting.

Monroe, LA, USA

Seetharama D. Jois

Contents

1	Basic Concepts of Design of Peptide-Based Therapeutics.	1
	Seetharama D. Jois	
2	Peptide Synthesis: Methods and Protocols	51
	Ted Gauthier and Dong Liu	
3	Computational Methods for Peptide Macrocyclic Drug Design	79
	Vikram Khipple Mulligan	
4	Strategies to Optimize Peptide Stability and Prolong Half-Life.	163
	Sophia M. Shi and Li Di	
5	Therapeutic Peptide Delivery: Fundamentals, Formulations, and Recent Advances	183
	Deepal Vora, Amruta A. Dandekar, and Ajay K. Banga	
6	Liposome Nanocarriers for Peptide Drug Delivery	203
	Jafrin Jobayer Sonju, Achyut Dahal, and Seetharama D. Jois	
7	Peptides and Their Delivery to the Brain	237
	Waleed Elballa, Kelly Schwinghamer, Eric Ebert, and Teruna J. Siahaan	
8	Emerging Peptide Drug Modalities for Intracellular Target Space.	267
	Tomi K. Sawyer	
9	Regulatory Issues for Peptide Drugs.	287
	Seetharama D. Jois	
	Index.	307

About the Author

Seetharama D. Jois is a Professor of Medicinal Chemistry in the School of Basic Pharmaceutical and Toxicological Sciences, College of Pharmacy, University of Louisiana at Monroe, USA. Before joining the University of Louisiana at Monroe in 2006, he worked as an assistant professor in the Department of Pharmacy at the National University of Singapore for 6 years. Dr. Jois obtained his PhD degree in 1994 from the Molecular Biophysics Unit, Indian Institute of Science (IISc), Bangalore, India, and a master's degree in physics from the University of Mysore, India. After his PhD degree, Dr. Jois obtained extensive training in pharmaceutical chemistry research at the University of Kansas in Lawrence, KS. During the past 20 years, he has established research in a unique area of drug design discovery, namely protein-protein interaction and drug design. He has secured funding from NIH and state agencies for his research. Dr. Jois has published more than 100 publications and has 3 US patent applications. Dr. Jois has edited a book *Drug Design and Discovery: Methods and Protocols* (Humana Press 2011) and special issues of journal series including *Current Pharmaceutical Design*, "Adhesion Molecules and Drug Targeting" (2006, Bentham Science Publisher LTD); and *Current Pharmaceutical Design*, "Biomaterial" (2008 Bentham Science Publisher LTD). He has 20 years of teaching experience in pharmacy schools and has published his innovative method of teaching in the *American Journal of Pharmacy Education* (2010, 74(8): 147).

Chapter 1

Basic Concepts of Design of Peptide-Based Therapeutics



Seetharama D. Jois

Contents

1.1	Introduction.....	2
1.2	Three-Dimensional Structure of Peptides.....	3
1.3	Targeting a Biochemical Pathway.....	5
1.4	Designing Drug When the Target Receptor Structure Is Known.....	6
1.4.1	The Binding Site and 3D Structure of Receptor Cavity/Surface.....	6
1.4.2	Agonist or Antagonist?.....	6
1.4.3	Pharmacophore Identification.....	7
1.5	Peptidomimetic.....	12
1.5.1	Backbone or Side-Chain Modification.....	12
1.5.2	Secondary Structure Mimics.....	14
1.6	Protein-Protein Interaction and Peptide Drug Design.....	14
1.7	Peptides from Natural Sources and Modification for Drug Design.....	16
1.8	Screening Assays.....	17
1.9	Binding Assays.....	17
1.10	Examples of Peptide-Based Drug Design.....	18
1.10.1	Enzyme-Based Drug Design.....	18
1.10.2	Receptor Ligand-Based Drug Design.....	22
1.10.3	Example of a Peptide from Natural Resources: Polymyxin B.....	28
1.10.4	Protein-Protein Interaction and Drug Design.....	32
1.11	Summary.....	41
	References.....	41

Abstract Peptides are natural ligands for numerous receptors in the human body, and many of these, such as hormones, neurotransmitters, and growth factors, have physiological roles. Peptides are highly selective and relatively well tolerated in the human body. More than 80 peptide-based drugs approved by the FDA are now on the market. A major challenge in peptide-based therapeutics is in vivo stability of peptides for oral administration. There have been several attempts to increase the

S. D. Jois (✉)

School of Basic Pharmaceutical and Toxicological Sciences, College of Pharmacy, University of Louisiana at Monroe, Monroe, LA, USA

e-mail: jois@ulm.edu

enzymatic stability of peptides using backbone or side-chain modification. These methods include cyclization, stapled peptide approach, and N-methylation, in addition to grafting peptides to cyclic peptide or cyclotide frameworks. Peptide therapeutics gained momentum during the last decade as the three-dimensional structures of many proteins were elucidated. The functions of these proteins and their relevance in the biochemical pathways of diseases were revealed. In this chapter, the fundamentals of peptide-based drug design are covered, with an emphasis on the structural biology aspects of peptide drug design. Four examples, namely, the design of an angiotensin-converting enzyme inhibitor; an RGD peptide cilengitide; the use of a naturally occurring peptide, polymyxin; and, finally, a relatively new area in peptide-based drug design, protein-protein interaction inhibitors, are described.

Keywords Peptide design · Peptidomimetics · Cyclization · Protein-protein interaction · Enzyme-based drug design · Natural peptides

1.1 Introduction

The use of peptides as drugs is not new; hormones were used as early as 1950 for therapeutic purposes (Hruby and Cai 2013). At present, there are more than 80 peptide-based drugs on the pharmaceutical market, and more than 150 peptide drugs are in clinical trials (Fosgerau and Hoffmann 2015; Lau and Dunn 2018; Muttenthaler et al. 2021). Peptide therapeutics gained momentum during the last decade as the three-dimensional structures of many proteins were elucidated, and the functions of these proteins and their relevance in the biochemical pathways of diseases were revealed. Peptides are natural ligands for numerous receptors in the human body, and many of these, such as hormones, neurotransmitters, and growth factors, have physiological roles. Naturally occurring ligand peptides can be modified for therapeutic purposes. At present, peptide-based therapeutics are in demand in two main areas—metabolic diseases and oncology (Fosgerau and Hoffmann 2015). Some peptide-based therapeutics, for example, Lupron™ for prostate cancer therapy and glucagon-like peptide-1 (GLP-1) for treatment of diabetes mellitus, have reached sales of more than a billion dollars (Kaspar and Reichert 2013). In terms of therapeutic agents, there are two classes of peptides, those that are extracted from natural resources and those that are designed and synthesized. Naturally occurring peptides can be extracted and modified for therapeutic purposes. Designed synthetic peptides may be from natural sources, based on peptides that occur in the human body, or from de novo design based on the protein receptor structure. More than 7000 naturally occurring peptides that have pharmacological activity have been reported (Padhi et al. 2014). At present, there are attempts to make the peptide orally available using different strategies (Qiu et al. 2017; Rader et al. 2018).

Peptides can be considered small pieces of proteins. Peptides and proteins are formed from amino acids by joining the carboxyl group of one amino acid with an amino group of another amino acid and releasing a water molecule in a chemical

reaction. Twenty different amino acids can be added in any combination, resulting in proteins of any length/size. According to the FDA guidelines (Feb 2019 proposal <https://www.biosimilarsip.com/2019/01/07/fda-proposes-to-amend-the-definition-of-biological-product/>), a polymer composed of 40 or fewer amino acids is a peptide. A protein is any alpha-amino acid polymer with a specifically defined sequence that is greater in size than 40 amino acids. A polymer of any alpha-amino acid that is made entirely by chemical synthesis and is greater than 40 but less than 100 amino acids in size is called a polypeptide. For example, leuprolide (Wilson et al. 2007), a hormonal peptide used for the treatment of prostate and breast cancer, consists of 10 amino acids, whereas teriparatide (Lindsay et al. 2016) is a form of parathyroid hormone consisting of 34 amino acids. Epoetin alfa, a 165-amino acid protein used for the treatment of patients with anemia associated with various clinical conditions, is a protein (Aapro et al. 2018). In this chapter, we will restrict our description of the design of peptides to those of up to 40 amino acids.

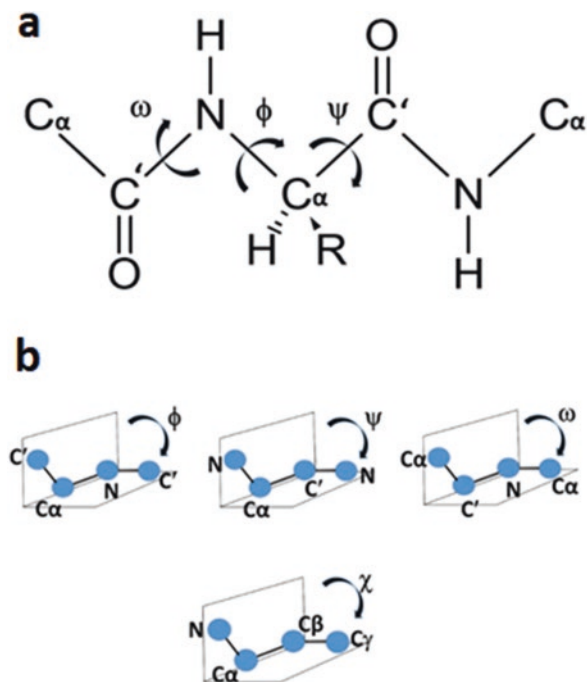
There are different approaches for peptide-based drug design, namely, (a) screening of naturally occurring peptides for biological or pharmacological activity; (b) enzyme and receptor-based drug design, also called rational drug design; and (c) design of drugs where the receptor is unknown (Mavromoustakos et al. 2011; Sable et al. 2017). A new class of design, protein-protein interaction and drug design, can be considered under receptor-based drug design.

1.2 Three-Dimensional Structure of Peptides

Peptides are short pieces of proteins, and peptides with less than 15 amino acids may or may not have a stable secondary structure in solution (Dyson et al. 1988). Many peptides acquire transient secondary structure in solution, and when they bind to their receptor, they may acquire a suitable secondary structure. The three-dimensional structure of peptides is defined by the backbone and side-chain torsion angles (ϕ , ψ , and χ) of the peptide chain in solution (or solid state). In fact, the structure of proteins was proposed based on the study of backbone angles of short peptide sequences by Ramachandran and co-workers (Ramachandran et al. 1963). The definition of torsion angles that dictate the three-dimensional structure of peptides is provided in the schematic diagram (Fig. 1.1a & b). The two amino acids are joined by a peptide bond, which has a partial, double bond character and is not freely rotatable at room temperature and, hence, is rigid. This torsion angle/dihedral angle is defined by four atoms, $C\alpha-C'-N-C\alpha$, and the angle is represented by ω , $\omega = 180$ for trans peptide bonds (in most cases, peptide bonds are in “trans” configuration) and $\omega = 0$ for “cis” peptide bonds. Phi and psi torsion angles are defined by the rotation around the single bonds around $C\alpha$ atom (Fig. 1.1). These torsion angles are also defined by the angles between the planes formed by three backbone atoms as shown in Fig. 1.1b.

Side-chain conformation of amino acids in the peptide is defined by the torsion angle χ that is defined by four atoms, $N-C\alpha-C\beta-C\gamma$. Although all possible torsion

Fig. 1.1 Backbone dihedral angles define the three-dimensional structure of peptides and proteins. (a) Phi, psi, and omega dihedral angles along the peptide main chain, (b) definition of phi, psi, omega, and chi formed by planes of atoms in the peptide backbone or side chains



angles with values from 0 to 360° are possible for ϕ , ψ , and χ , because of steric hindrance of atoms that come close together, only certain values are possible for these torsion angles that define the 3D structure of a peptide. For details on the analysis of side-chain torsion angles, readers can refer to Chakrabarti et al. and McGregor et al. (Chakrabarti and Pal 2001; McGregor et al. 1987). The most possible secondary structures or open structures that a peptide can acquire are shown in Fig. 1.2. Why is the 3D structure of the peptide important in drug design? The backbone structure and the side-chain conformation of the peptides are crucial in receptor-ligand and enzyme-substrate interactions, and hence, understanding the 3D structure of peptide ligands and interaction of these peptides with receptors is important in peptide-based drug design.

In peptide-based drug design, there are two main categories (Strømgaard et al. 2017; Craik 1996; Hruby and Cai 2013): (1) design of peptides for known target receptors and (2) design of peptides for unknown target receptors. If the target receptor is known then, in the majority of cases, the 3D structure of the target is evaluated based on experimental methods such as X-ray crystallography and NMR or on homology modeling. In the case of an unknown target, the design is based on the screening of peptide structures by different molecular biology methods or from the knowledge of natural product therapy of peptides from the literature. In such cases, optimization of peptide-based design is heavily dependent on the known peptide ligand structure.

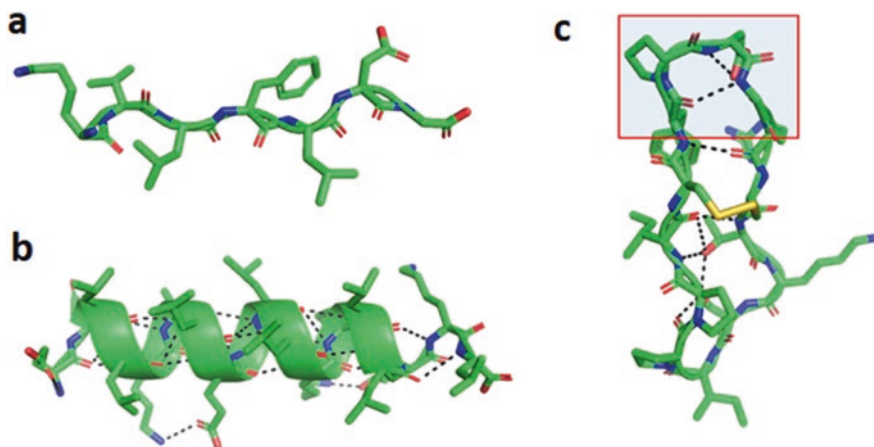


Fig. 1.2 Possible three-dimensional structures of peptides in solution. (a) Linear or open structure, (b) helical structure, (c) beta-turn or beta-hairpin structure (PDB ID 4xoj). Structures generated by PyMol Software (Schrodinger LLC)

1.3 Targeting a Biochemical Pathway

Biochemical pathways control many physiological processes in the human body. In the case of target-based rational design, the identification of biochemical pathways that lead to disease states is a crucial step. Choosing the target protein receptor or enzyme for drug design is the first step. For example, in the case of the renin-angiotensin system, the biochemical pathways are controlled by at least three enzymes and two receptors (renin, angiotensin-converting enzyme, angiotensin-converting enzyme-2, angiotensin receptors type I and II) (Acharya et al. 2003). To control blood pressure, any of the biochemical pathways related to the enzymes and receptors is possible as long as enhancing or blocking this pathway does not lead to significant changes in the normal physiological process and produce severe side effects. Another example is the case of cholesterol control in the body. Apart from its consumption in food, cholesterol is synthesized in the body; hence, to reduce the cholesterol in the body, its synthesis must be controlled. There are nearly 20 steps of synthesis for the cholesterol produced in the body, and many enzymes are involved in this biochemical pathway (Lemke and Williams 2012). However, the enzyme that controls the rate-limiting step and does not cause side effects is the best one to choose. Thus, a major enzyme that controls cholesterol synthesis, HMG-CoA reductase, is targeted for drug design. There are some receptors with multiple sites on the surface that are related to the physiological process, and it is very difficult to control one pathway without affecting the other. One such example is G-protein-coupled receptors (GPCR) (Weis and Kobilka 2018; Hauser et al. 2017), and many membrane receptors have a GPCR type of structure. Most of these receptors have

similar general protein structures, and their extracellular domains have subtle changes for binding to specific peptide substrates to control the physiological process. Designing a peptide to target one site is extremely hard as binding at one site may affect other sites. In general, based on the literature, one has to choose a pathway that has the fewest possible side effects.

1.4 Designing Drug When the Target Receptor Structure Is Known

1.4.1 The Binding Site and 3D Structure of Receptor Cavity/Surface

The target protein receptor could be an enzyme or a protein cell surface receptor/intracellular receptor or, in protein-protein interactions, one of the protein partners. Depending on whether it is an enzyme or a receptor, or a protein-protein interaction site, a binding site could be a compact pocket or a shallow pocket or a surface with possible binding sites. In HIV protease (Martin et al. 1999; Arts and Hazuda 2012) or angiotensin-converting enzyme receptor (Acharya et al. 2003), the binding cavity is deep inside the protein. The natural substrate binds to the enzyme with multiple interactions (hydrogen bonding, hydrophobic interaction, or electrostatic interaction (metal-ion or charge-charge interaction from side chains of amino acids)). In the HIV protease enzyme, the substrate peptide interacts with an enzyme that has a C2 symmetry (Fig. 1.3a). The substrate forms hydrogen bonding and hydrophobic interaction with the receptor. In the case of the ACE enzyme, the substrate is buried deep in the pocket with hydrogen bonding and hydrophobic interactions (Fig. 1.3b). On the other hand, an angiotensin receptor is a cell surface protein receptor that has a concave surface, and the peptide is much more exposed to the surface on one side (Fig. 1.3c). In the case of protein-protein interaction between CD2 and CD58, the two domains form a perpendicular interaction in a hand-shaking pattern. The PPI surface interaction has 1200 Å² surface area with 10 salt bridges, 5 hydrogen bonds, and 2 hydrophobic interactions. The surface area is slightly concave, and there is no cavity (Wang et al. 1999) (Fig. 1.3d). In the design of peptides in which the target is known, familiarity with the ligand-receptor surface and shape is important in the design.

1.4.2 Agonist or Antagonist?

In the design of peptides for modulating the biochemical pathway for a possible cure for the disease, the desired effect of modulation is important. For hormones and neurotransmitters, naturally occurring peptides are ligands for receptors, and they

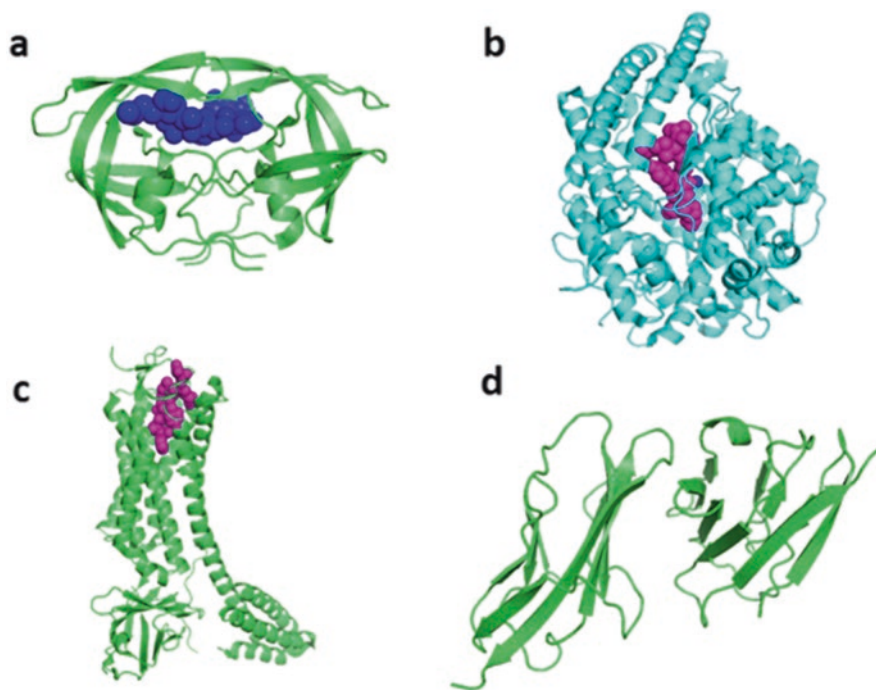


Fig. 1.3 Enzymes and receptors with binding and protein-protein interaction sites. Crystal structure of (a) HIV protease with substrate bound (blue spheres) (PDB ID: 1f7a), (b) ACE with substrate peptide buried deep inside the cavity (PDB ID:4aph), and (c) ACE type I receptor with ang II peptide (PDB ID:6os0). One part of the peptide is buried in the receptor-binding site, and the other side is exposed to solvent and (d) CD2-CD58 complex with large PPI surface area of binding (PDB ID:1qa9)

have agonist activity. The binding of the peptide leads to enhancement of the activity. In such a case, the designed peptide should enhance the receptor activity (agonists) (Hruby 2002). On the other hand, if the receptor function must be modulated to decrease the activity, antagonists are needed. In both cases, the starting point is the natural ligand. Details of how the natural ligand binds to receptors and causes enhancement of activity need to be elucidated.

1.4.3 Pharmacophore Identification

The establishment of detailed structural aspects of natural ligands and receptors is a time-consuming task and involves molecular biology, bioanalytical assays to evaluate binding, and NMR or X-ray crystallography to elucidate the detailed structural aspects of interaction. Once the structural elements of the receptor-ligand interactions are established, the important structural elements that are required for binding

Fragments of long peptide with overlapping sequences

NH₂-DLSYSLRRKPSTSPNLQWEQTKRYLLSTSFPHYHILLDYLSK-COOH

Overlapping sequences of the peptide

NH₂-DLSYSLRRKPSTSPNLQWEQTKRYL-COOH

NH₂-QWEQTKRYLLSTSFPHYHILLDYLSK-COOH

NH₂-RKPSTSPNLQWEQTKRYLLSTSFPHYHILL-COOH



Truncation of peptide, alanine scanning

Truncation of the peptide

NH₂-ALKYTRSFPGQNDLSLSRIK-COOH
NH₂-LKYTRSFPGQNDLSLSRIK-COOH

NH₂-ALKYTRSFPGQNDLSLSRI-COOH

Alanine scanning

NH₂-ALKYTRSFPGQNDLSLSRIK-COOH

NH₂-AAKYTRSFPGQNDLSLSRIK-COOH

NH₂-ALAYTRSFPGQNDLSLSRIK-COOH

NH₂-ALKATRSFPGQNDLSLSRIK-COOH

NH₂-ALKYARSFPGQNDLSLSRIK-COOH

NH₂-ALKYTASFPGQNDLSLSRIK-COOH

NH₂-ALKYTRAFPGQNDLSLSRIK-COOH

NH₂-ALKYTRSAPGQNDLSLSRIK-COOH

-----A-----

-----A-----

Fig. 1.4 Approaches to the design of peptides as agonists or antagonists. Determination of pharmacophores in the peptide using peptide mapping, truncation, or ala scanning

will be identified (pharmacophore). Natural peptide ligands may have more amino acids than are actually needed for binding. Some amino acids may not be involved in binding; however, they might be important for maintaining the 3D structure of the peptide ligand/substrate. Modification of natural ligand to improve agonists or antagonist activity involves a systematic approach that is established in the peptide-based drug design literature. Truncation of peptides and replacement of similar or unnatural amino acids or D-amino acids are viable options. Furthermore, these peptide drugs must be stable in vivo as they are administered via intravenous (IV) in most cases. Thus, pharmacophore elements that provide enzymatic stability in vivo for the peptide need to be considered. Here, we describe some general approaches.

Truncation of Peptide Truncation of the peptide is carried out to evaluate whether the entire peptide chain is needed for activity at the receptor or if a short version of the peptide can be as effective as a full-length peptide. Typically, the designed peptide is truncated from N- or C-termini with one amino acid deletion at a time, generating a series of peptides of different lengths (Vlieghe et al. 2010) (Fig. 1.4). All of these peptides are evaluated for pharmacological activity, and the peptide with the best activity among them is chosen for further modification. In the case of long-chain peptides with more than 20 amino acids, a peptide is fragmented with at least 10–12 amino acids with overlapping regions, and its activity is evaluated.

Alanine Scanning Pharmacophore functional groups in peptides are limited to 20 amino acids that occur in nature. The physicochemical properties of side-chain functional groups are considered to design a pharmacophore for peptide-based drug design. The importance of functional groups in peptide sequences can be determined by alanine scanning (Morrison and Weiss 2001). In this method, each amino

acid in the designed peptide is replaced with alanine, and these designed peptides are evaluated for biological activity. An increase or decrease in the activity of the peptide after replacement of alanine determines the importance of functional groups in the peptide (Fig. 1.4).

Determination of the Structure of Biologically Active Conformation As mentioned above, peptides with less than 15 amino acids may or may not have a stable conformation in solution. If the bound substrate peptide or interacting protein partners have a particular secondary structure that helps to bind to the receptor/enzyme, then the secondary structure of the bound substrate/ligand is important. In peptide-based drug design, once the minimum number of amino acids and the key amino acids needed for biological/pharmacological activity are determined, an attempt to determine the active conformation should be made. Simple methods such as circular dichroism spectroscopy provide the overall conformation of the peptide, indicating whether the peptide acquires an α -helical, beta-turn/beta-hairpin, or completely open and unordered structure. NMR and X-ray crystallographic methods can be used to determine the detailed conformation of the peptide.

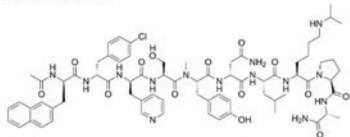
Replacement with D-Amino Acids Changing the chirality of amino acids changes the orientation of the side chain that is exposed to the receptor. Thus, in short-chain peptides, the replacement of L-amino acids with D-amino acids is useful to determine the structure-activity relationship. Furthermore, D-amino acids are known to make the peptides relatively stable in solution. However, one should consider some precautions here. If all the L-amino acids are replaced by D-amino acids, the sequence of the peptide may need to be reversed as side-chain orientation completely changes with a change in chirality.

Conformational Constraints When the peptides are designed as agonists or antagonists to bind to a receptor, in many cases, the design is based on the naturally occurring substrate peptide or protein ligand that binds to the receptors. The pharmacophore portion of the ligand usually has a restricted conformation when bound to the receptor to accommodate its size and charge within the binding pocket. In terms of change in the free energy of binding of a ligand to its receptor, the free energy of the ligand in the bound form is more favorable to binding than is the unbound ligand state. That means very flexible peptide ligands need more free energy change to bind to the receptor as the ligand has to change the conformation from free to bound state. If we introduce the conformational constraints in the designed peptide that are ready to bind to the receptor, then these designed peptides can bind to the receptor with high affinity (small K_d value). Another way to look at this is when an epitope of a protein binds to the receptor, it has a secondary structure that is suitable for binding. When we use such peptide epitopes and synthesize them, the synthetic peptide may or may not acquire suitable conformation in solution. Hence, to increase the binding affinity, conformational constraints in the peptide can be introduced. Conformational constraints also help to stabilize the peptide under in vivo conditions.

N- and C-Termini Modification Linear and cyclic peptides with cyclization by disulfide bond have free N- and C-termini. These terminals have charges depending on the pH condition. In vivo, such peptides are susceptible to degradation via aminopeptidases and carboxypeptidases (for a detailed description of these enzymes present in vivo, readers should refer to Vlieghe et al. (2010)). Peptidases can be classified either as exopeptidases, which specifically hydrolyze the C- or N-termini of a peptide, or as endopeptidases that are capable of hydrolyzing amide bonds within a peptide (Bottger et al. 2017; Cudic and Stawikowski 2007). N-termini of amino peptides can be acetylated, and C-termini can be amidated in peptides to make the termini stable. Other methods, such as the substitution of D-amino acids at the N- and C-termini, will help to protect the peptide from enzymatic degradation. Some of the modifications are seen in examples of the hormonal drugs abarelix and ganirelix that are GnRH antagonists and in the case of the HIV virus fusion inhibitor enfuvirtide (Fig. 1.5). The hormonal peptide GnRH has pyroglutamic acid at the N-terminal, and the C-terminal is amidated. N- and C-termini are also modified for attaching imaging agents to the target peptide (theranostic approach).

Cyclization (Main Chain or Side Chain) The global conformation of peptides can be restricted using a cyclization of the peptide chain that is widely accepted and employed in peptide-based drug design. Usually, such cyclization applies to peptides chains of <15 amino acids length as increasing the chain length may not result in conformational constraints if the peptide chain is long. The main reason for this is that cyclization helps to stabilize the peptide in vivo as termini are not available

Abarelix



Enfuvirtide

Acetyl-YTSLIHSLIEESQNQQEKNEQELLELDKQWASLWNWF-NH₂

Octreotide

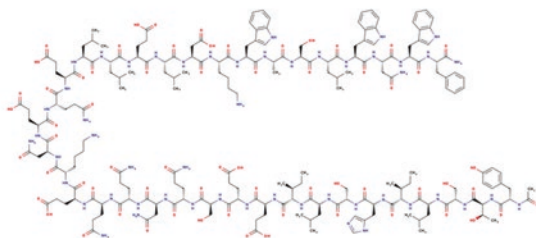
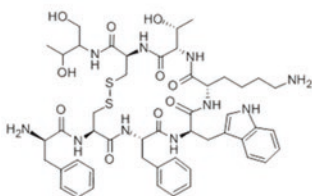


Fig. 1.5 Modification of peptides for in vivo delivery. The FDA-approved peptide drug abarelix is modified at the N-terminal by acetylation and at the C-terminal by amidation. Octreotide is cyclized by a disulfide bond. The sequence of the HIV fusion inhibitor peptide enfuvirtide was modified at the N- and C-termini along with its structure. Structures of abarelix and octreotide were obtained from PubChem. Enfuvirtide was from <https://go.drugbank.com/drugs/DB00109>

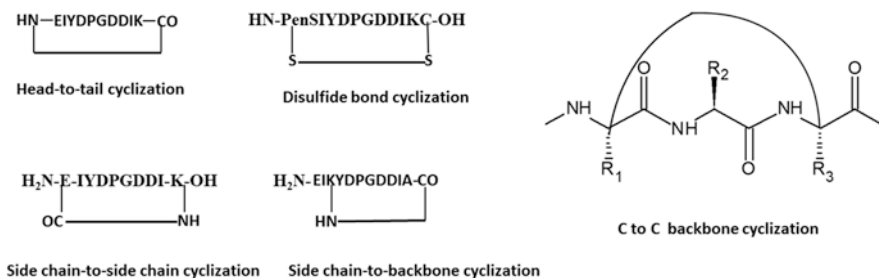


Fig. 1.6 Different strategies for peptide cyclization

to amino and carboxypeptidases. Cyclization is also known to induce and stabilize the secondary structure of peptides by restricting the phi and psi values of some amino acids. There are different methods of cyclization: (a) backbone head to tail cyclization, (b) disulfide bond cyclization, (c) side-chain cyclization, and (d) formation of lactam bridges (Fig. 1.6). In terms of side-chain cyclization, there are different possibilities such as backbone-to-side chain and side chain-to-side chain. While cyclizing a peptide backbone, a few key points must be considered. At the point of cyclization, N- or C-termini should not play a major role in binding to the receptor. The secondary structure of the peptide should not be completely different than the secondary structure that is required for binding to the receptor. In some cases, cyclization can cause a complete loss of activity. Hence, structural studies related to a peptide using NMR or circular dichroism should be carried out on a linear active peptide before cyclization. When cyclization is carried out using a side chain, one must make sure that the pharmacophore element of the side chain used for cyclization does not have any importance in binding to the receptor. Alanine scanning or truncation of the peptide provides information about the pharmacophore element and amino acids needed for the activity. One of the easiest ways to evaluate whether cyclization will have any effect on the activity of the peptide is to introduce the Cys amino acids at the N- and C-termini of the peptide sequence and cyclize the peptide by disulfide bond cyclization. If the peptide becomes more active than the linear peptide, then backbone cyclization can be carried out (Fig. 1.6). One of the earlier attempts to achieve new cyclization methods was described by Gilon et al. (1991). They described the backbone type of cyclization without altering the carboxy or amino terminal groups. They proposed the concept of C-backbone cyclization where the NH or C^αH hydrogens were replaced by ω-functional alkylidene chains. Design of cyclic peptides from epitope of protein is described in a review by Rubin et al. (2018). During cyclization, conformational constraints can be used. There are several reports of design of cyclic peptides and converting them to peptidomimetics (Dishon et al. 2018; Talhami et al. 2020; Ling et al. 2015). One of the best-known and widely used methods of cyclization and conformational constraints is the induction of beta-turns/beta-hairpin bends in the peptide structure. Beta-turn is stabilized after cyclization if the peptide has beta-turn propensity amino acids such as Pro-Gly, Pro-Pro, and Asn-Pro (Gibbs et al. 2002; Shapovalov et al. 2019; Rotondi and

Gierasch 2006; Wilmot and Thornton 1990). In case such amino acids are not present in the peptide, Pro-Gly or Pro-Pro sequences should be introduced to obtain stable turn structures in peptides and cyclic peptides (Chatterjee et al. 2008a; Hutchinson and Thornton 1994). This is very useful in peptide epitopes that have beta-sheet structures in the 3D structure, but the sequences are not continuous (examples will be provided later). Similarly, alpha helix inducing amino acid residues and unusual amino acid residues can be used (Li et al. 2020; Fisher et al. 2017; Karle et al. 1989; Hurevich et al. 2013).

1.5 Peptidomimetic

The definition of peptidomimetic has changed over the years. Peptidomimetics are compounds whose pharmacophore mimics a natural peptide or protein in 3D space with the ability to interact with the biological target and produce the same biological effect (Vagner et al. 2008). Then the definition of peptidomimetic was classified into three types (type I to type III). In 2015, Grossman et al. classified peptidomimetics into four classes and named them Class A–D (Pelay-Gimeno et al. 2015) (Table 1.1).

1.5.1 Backbone or Side-Chain Modification

Peptidomimetics can be designed by incorporating changes at particular amino acids in the peptide backbone or side chain. This targeted chemical modification of peptides may increase the in vivo stability. However, identification of potential degradation sites in the peptide by an enzymatic process is important in the design of peptidomimetics. Experiments can be conducted to identify the degradation of peptides using proteolytic enzymes (Miranda et al. 2008). In general, based on the enzymes in vivo, the route of administration, and the sequence of peptides, enzymatic stability of a peptide can be proposed. Data can be obtained from the literature described by Li et al. and recently by Bottger et al. (Li et al. 1995; Bottger et al. 2017). There are established methods for backbone modification of peptides, including isosteric or isoelectronic substitution. A peptide bond can be replaced with its

Table 1.1 Peptidomimetic definition and classes (Pelay-Gimeno et al. 2015)

Class	Description
A	Minimal alterations to the peptide backbone and side chain
B	Major backbone alteration-unnatural amino acids and small molecule building blocks
C	Significant modification of the peptide structure replacement of peptide backbone with small-molecule character
D	A small molecule mimics the mode of action of a bioactive peptide without a direct link to its side-chain functionality

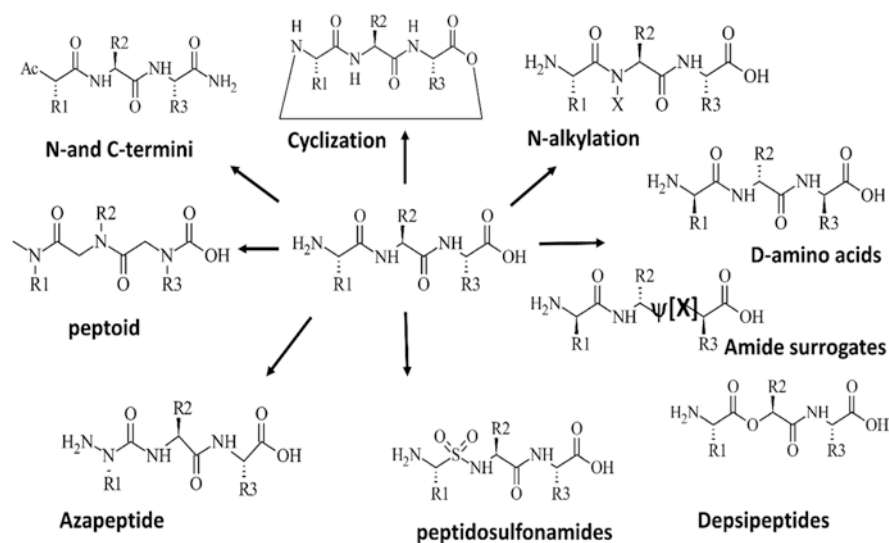


Fig. 1.7 Peptide main chain modification by different strategies and surrogates of peptide bonds

surrogates to protect against protease cleavage of the amide bond and significantly improve the stability of peptides *in vivo*. However, such modifications can have an effect on the physicochemical and conformational properties of the peptide (Cudic and Stawikowski 2007). Thus, amide bonds at the selected position should be replaced with surrogates of amide bonds. Some of the surrogates used in peptide-based drug design are shown in Fig. 1.7. When such surrogates are used in the peptidomimetic synthesis, the standard solid-phase synthesis procedure requires modifications in the synthesis steps, and appropriate coupling and protecting strategies are used.

A simpler approach to peptide modification for the stability of peptides is to modify the side chains involved in the protease recognition site. Natural amino acids in the peptide can be replaced with nonnatural amino acids with modified side chains but may retain the physicochemical properties of amino acid side chains (Wisniewski et al. 2011; Frey et al. 2008; Henninot et al. 2018). Several amino acid side-chain surrogates are available on the market, some of which are prepared for use in standard solid-phase synthesis using Fmoc chemistry. Examples of modifications of arginine and lysine side chain include ornithine, citrulline, homoarginine, and N-isopropylornithine. Apart from the side chain, some of the unnatural amino acids such as beta-amino acids (Seebach and Gardiner 2008; Cabrele et al. 2014; Del Borgo et al. 2017; Kudo et al. 2014) have backbone modifications incorporated within the molecule. In addition to binding site residues, other amino acid residues that are not important in binding to the receptor or enzyme can be modified for solubility or hydrophobicity to cross the intestinal barrier.

1.5.2 Secondary Structure Mimics

The interaction between proteins or proteins and peptides depends on the 3D structure of the peptide, the disposition of amino acids side chain at the site of interaction, and physicochemical properties of amino acid side chain. Most of the ligand-receptor interactions involve secondary structure elements that present proper orientation of amino acids from the ligand to the receptor. Thus, the secondary structure of peptides is important in the design. Since peptides designed as ligands for therapeutic peptides are often a short sequence of amino acids (4–15 amino acids), they do not acquire a stable secondary structure in solution. Thus, imposing the secondary structure elements in such peptides is important for enhancing the activity of the peptide or binding to the receptor with high affinity. Secondary structures in short peptides can be incorporated using the peptidomimetic approach where an organic functional group is introduced in the peptide to induce secondary structure. β -turn is one of the most widely used secondary structures in peptides for stability and conformational constraints as well as for exposing the amino acid residues for receptor binding (pharmacophore). Several β -turn mimetics that are organic functional groups are available for incorporation into the peptide. Similarly, mimetics of alpha-helix and β -sheet are also available for the design. For details and examples of secondary structure mimetics, readers can refer to the literature (Fuller et al. 2009; Tan et al. 2016; Crecente-Garcia et al. 2020; Eckhardt et al. 2010; Deike et al. 2020; Jayatunga et al. 2014; Lanning and Fletcher 2015; Lao et al. 2014; Mabonga and Kappo 2020; Davis et al. 2007; Loughlin et al. 2010; Khakshoor and Nowick 2008; Nowick 2008; Ross et al. 2010; Lenci and Trabocchi 2020; Gokhale et al. 2015).

In addition to the modification of side chains or backbones, other methods such as lipidation of peptides including myristoylation are used for increasing the in vivo efficacy of peptides and their possible oral bioavailability (Dishon et al. 2019).

1.6 Protein-Protein Interaction and Peptide Drug Design

Protein-protein interaction (PPI) is a major player in almost all biological processes. Nearly 650,000 PPI interactions have been predicted to be involved in maintaining the perfect balance for the functioning of cells in humans (Stumpf et al. 2008). PPI involves the binding of two or more proteins with each other for carrying out specific cellular function. Almost all metabolic processes involve PPI, and several studies suggest that various metabolic diseases like cancer, autoimmune diseases, cardiovascular diseases, and neurodegeneration involve complex interactions of various PPI, affecting pathogenesis and disease progression (Fischer et al. 2015). During the development of a disease such as cancer, certain proteins are overexpressed by the cancer cells resulting in an abnormal and undesired PPI and in the

collapse of the natural and perfectly engineered cellular signaling (Ryan and Matthews 2005). Due to this critical nature of the PPI, there is a vast open field of therapeutic opportunity for researchers to explore new treatments for different diseases such as cancer, autoimmune diseases, and viral and bacterial infections by modulating specific PPIs. Inhibitors of PPI and peptide-based drug design have applications in three main areas, metabolic diseases, autoimmune diseases, and cancer.

Targeting the abnormal and aberrant PPI is of great clinical importance, but it has also been a significant challenge. Historically, PPIs are considered unattractive targets for drug discovery; less than 0.01% of the PPIs have been targeted with an inhibitor (Thompson et al. 2012). The nature of the PPI surface makes it a very difficult site to target, and unlike conventional drug targets such as enzymes and ligand-binding sites of GPCR, most of the binding surface between proteins are usually large, 1500–3000 Å², and involve many polar and hydrophobic interactions. Because the flat nature of the binding interface provides no definite binding cavity, PPI inhibition is challenging; in addition, the binding surface is not well defined like small molecule binding surfaces (Cunningham et al. 2017). Hence, targeting the PPI with small molecules with high specificity has been incredibly difficult. Due to the diverse nature of the PPI, a thorough and detailed knowledge of the interaction at the molecular level is required for a successful drug design and discovery approach (Ryan and Matthews 2005). The drug discovery process has taken a significant leap in targeting difficult PPI sites, and scientists have been successful in targeting so-called undruggable sites with the help of advanced drug screening techniques and rational drug design (Rognan 2015). Recent advances in structure-based drug design have led to the identification and prediction of binding surface between the proteins, as well as details of binding, including the key amino acids that dominate the PPI binding site. This region of key residues that determines the fate of PPI is termed a hotspot region. Specific and selective modulation of PPI can be achieved by studying these key residues and the nature and role of their interactions in binding (Modell et al. 2016). Analysis of the PPI binding site has shown that 15–40% of PPI is mediated by short linear peptides (London et al. 2013). Small synthetic molecules, peptides, and proteins are designed to specifically target PPI, and each class of these targeting agents used has advantages and disadvantages with respect to efficacy, specificity, bioavailability, and synthesis process. Targeting PPI with peptides and peptidomimetics that can bind to hotspot regions of a binding interface had been studied extensively (Laxio Arenas et al. 2019; Cunningham et al. 2017). Certain PPIs that are sufficiently exposed can be effectively targeted with specific antibodies; however, antibodies generally cannot bind to intracellular PPIs and deeply embedded PPIs and also have substantial stability issues. Peptides and peptidomimetics are explored as simple and effective alternatives for targeting PPI as they have high affinity and specificity (Wu and Gellman 2008; Jo et al. 2012) (Zhao and Chmielewski 2005).

1.7 Peptides from Natural Sources and Modification for Drug Design

The human body produces several peptides, including hormones, growth factors, and anti-infective agents. More than 5000 naturally occurring peptides have been identified from different natural sources (Padhi et al. 2014). Many of these peptides are produced for physiological processes in humans and other species, as well as defense mechanisms, and thus, a variety of peptides are produced. Among the naturally occurring sources, much of the literature concerns antimicrobial peptides and marine peptides.

Peptides from the venom of spiders, snails, and snakes have evolved as a means of predators to kill prey. Some of these are being used to treat neurological or cardiovascular diseases or to alleviate pain (Uhlig et al. 2014). One of the best examples of such natural peptides is exenatide, a glucagon-like peptide-1 agonist used in the treatment of diabetes mellitus type 2. This peptide was originally discovered in a hormone found in the saliva of the Gila monster. A synthetic version of this peptide is now produced as a commercial drug (Bond 2006). Apart from natural amino acids, these peptides have D amino acids and modification of main or side chains in the amino acids. Thus, natural resources can be used to explore the chemical space that has evolved for millions of years.

Naturally occurring peptides can be used directly for therapeutic purposes or can be modified for synthetic purposes or for therapeutic use. They can also be used as a template for drug design, as in the case of cyclotides (de Veer et al. 2019). Antimicrobial peptides (AMPs) are produced by multicellular organisms as a defense mechanism against competing microbes. These peptides exhibit their antimicrobial activity by interacting with membrane leading to disruption of membrane structure and membrane-associated physiological events. Most of these AMPs have positive charges, and they interact with negatively charged lipids of the membrane. These peptides have a combination of hydrophobic and hydrophilic amino acids placed in a particular pattern in the sequence and secondary structure that enables them to insert into the hydrophobic bilayer of the membrane. AMPs do not necessarily have helical structures, and they exhibit different folded structures (Fjell et al. 2011). There are AMP databases available that will help to screen the peptide for therapeutic activity (Wang et al. 2009). These peptides also exhibit toxicity, and hence, AMPs may have to be modified for human therapeutic purposes. Several synthetic AMPs have entered clinical trials, and several of these are immunomodulatory agents. There are more than 3000 AMPs that have been characterized and evaluated for their potency for therapeutic purposes. However, most of these failed to enter clinical trials (Chen and Lu 2020). Currently, there are seven AMPs that are approved by FDA, six of which are small peptides, including gramicidin D, daptomycin, vancomycin, oritavancin, dalbavancin, and telavancin (Usmani et al. 2017). In terms of the design of AMPs as peptide drugs, one of the important aspects is to make the peptide less toxic to humans while maintaining their antimicrobial (antibacterial) properties. This approach includes rational design, high-throughput screening, structure-activity relationship (SAR) approaches, and computational

methods (Torres et al. 2019; Ballantine et al. 2018; Krauson et al. 2012; Porto et al. 2018; Lopez-Perez et al. 2017).

Other peptides from natural sources are from venom peptides of spiders, snails, and snakes, and these include eptifibatide (Saab et al. 2012), linaclotide (Love et al. 2014), ziconotide (Pope and Deer 2013), and pramlintide (“Drugs for type 2 diabetes” 2019; Jones 2007). Most of the naturally occurring peptides must be extracted and purified for characterization and activity. Once the sequence of the peptide and structure is elucidated, analogs can be designed and synthesized. The SAR of these naturally occurring peptides will be similar to those of the designed peptides described in this chapter.

1.8 Screening Assays

Whether peptides are designed by a rational design approach or extracted from natural resources by high-throughput screening, they must be screened for potency in terms of pharmacological or biological activity. Since initial screening is performed on several peptides, the assay developed should be relatively simple, specific, sensitive, rapid, and reliable. In addition to this, it should be inexpensive as screening a large number of peptides for activity may require a large number of reagents and become very expensive. Usually, a dose-response curve is needed to establish the 50% inhibition or pharmacological activity. Establishment of a biological assay that rapidly determines pharmacological or enzymatic activity is essential. For example, in the case of anticancer activity of compounds, the antiproliferative activity of compounds in cells using MTT (Kumar et al. 2018) or CellTiter-Glo assay (Tolliday 2010) is preferred. These assays can be performed on a large number of a library of compounds in a relatively short time. In the case of RGD peptide design for anti-thrombotic activity, platelet aggregation using a simple light scattering instrument can be used. In the case of enzyme-based drug design, a simple competitive binding assay based on luminescence or fluorescence change upon the addition of inhibitors is evaluated. Other bioanalytical techniques such as surface plasmon resonance for binding of a ligand to a specific protein can be used (Fang 2011).

In peptide-based drug design, controls should be used to validate the potency of the peptide. Typically, a peptide with a scrambled sequence is generated once the sequence of one of the potent peptides is known. This control could be linear or cyclic, depending on whether the potent peptide is linear or cyclic.

1.9 Binding Assays

When the pharmacological activity of a peptide is established, one needs to establish the molecular mechanism in terms of binding and whether the peptide designed binds to the protein of interest and exhibits its activity. If the target protein is an

enzyme, cellular or pure enzyme activity assays can be carried out. If the target is a receptor or protein-protein interaction, a particular binding assay needs to be established. Examples include ELISA, enzyme activity inhibition assay, and binding by radioactive or fluorescent-labeled peptides. Currently, non-labeled protein assays such as surface plasmon resonance are available. Some of these assays use pure protein, and pure protein assay is different than assays that are carried out in the cellular condition. Each has its advantages and disadvantages.

Once a hit compound is obtained in these assays, the most potent compound is studied in an animal model of the appropriate disease state. If the peptide shows potency in a preclinical animal model, further optimization for the stability of the peptide *in vivo* is investigated. Since the subject of this chapter is the design of peptides and peptidomimetics, we restrict our descriptions to the design part and provide some examples of peptide-based drug design. Four examples are provided here that include enzyme-based drug design, receptor-based drug design, peptides from natural resources, and protein-protein interaction and drug design.

1.10 Examples of Peptide-Based Drug Design

1.10.1 *Enzyme-Based Drug Design*

In the rational drug design process, details of biological pathways of a particular pathological condition and structures of receptors that are important in biochemical pathways are studied initially. Following that, which pathways should be modified and what protein receptors or enzymes should be targeted will be decided. For example, in the case of high-blood pressure control, pathways related to the renin-angiotensin system are studied (Fig. 1.8) (Foye 2008). Angiotensinogen, a protein released by the liver, is converted into angiotensin I, a peptide, by the plasma renin enzyme. Angiotensin I is subsequently converted to angiotensin II, a shorter peptide, by an angiotensin-converting enzyme (ACE) found on the surface of vascular endothelial cells. Angiotensin II binds to angiotensin receptors found in the heart, blood vessels, and kidney. The binding of angiotensin II to receptors causes blood vessels to narrow (vasoconstrictive effect) and increase blood pressure (Masuyer et al. 2012). This is a general biochemical pathway that is well controlled with the help of the circulatory system, kidney, and brain.

In disease states, conversion and release of angiotensinogen and angiotensin II peptide are increased and, hence, increase blood pressure or hypertension. From the schematic diagram (Fig. 1.8), it can be seen that the production of angiotensin I peptide can be modulated by blocking the enzyme renin, blocking the angiotensin-converting enzyme, or blocking the binding of angiotensin II to receptors (AT1). Once the biochemical pathway was discovered, scientists proposed that rationally designed peptides could be used as inhibitors of enzymes or peptide molecules that could block the AT1 receptors and could be used for therapeutic purposes. The

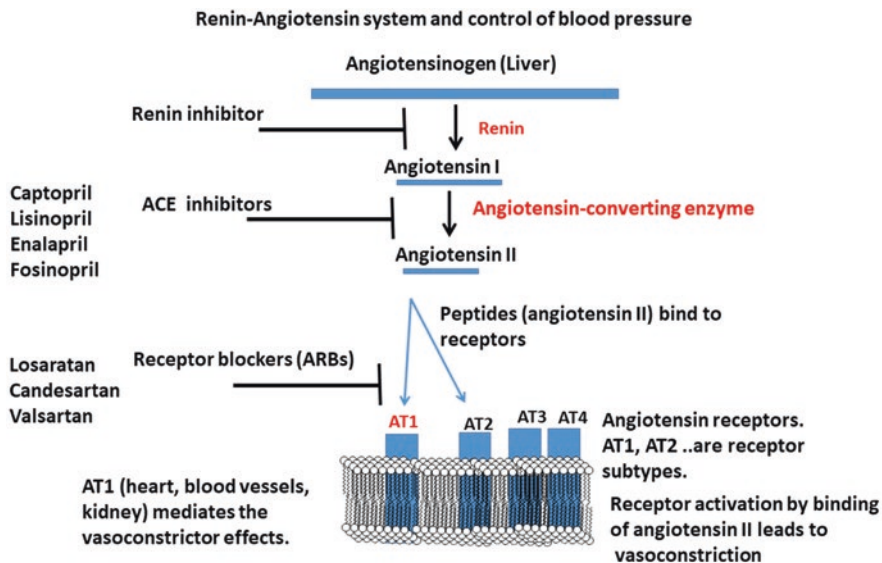
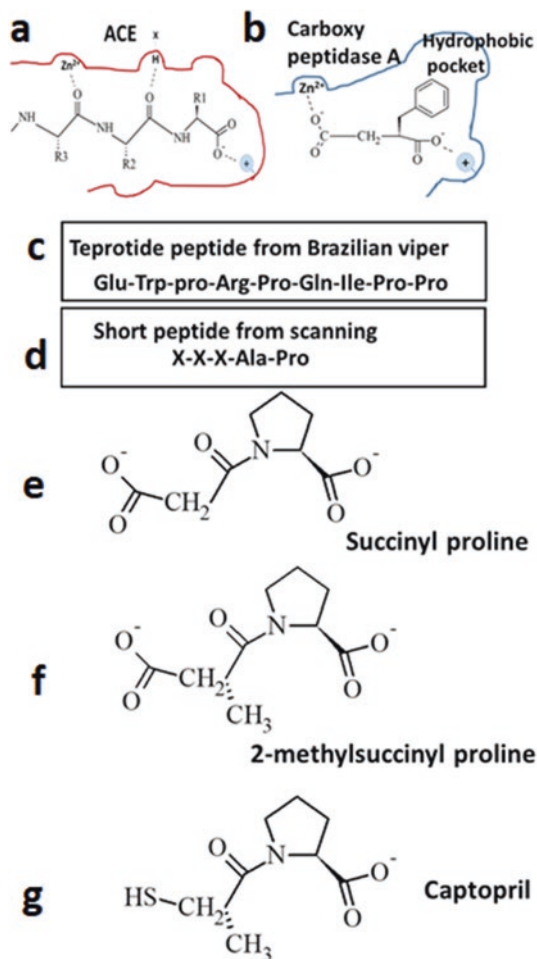


Fig. 1.8 Biochemical pathway depicting the renin-angiotensin system and inhibition strategy. This pathway can be inhibited or modulated at different stages. Angiotensinogen is a protein, and angiotensin I and II are peptides. Drugs that inhibit the pathways are shown at the left

second step in this process was the choice of pathway to be blocked. From Fig. 1.8, it is clearly seen that there are three possible places in the renin-angiotensin biochemical pathways that can be modulated to control blood pressure. Once the targeting pathway was established, the structures of proteins that were involved in the biochemical pathway were studied in detail. In the example above, the angiotensin-converting enzyme (ACE) structure was studied to design a peptide that binds to an ACE and slows the conversion of angiotensin I to angiotensin II. The molecule designed was an enzyme inhibitor.

When peptide-based ACE inhibitions were designed, the crystal structure of ACE was not available. However, ACE is similar to carboxypeptidase in its structure and function. The difference was that carboxypeptidase cuts one amino acid at a time from the C-terminal, whereas ACE cuts two amino acids at a time. Hence, it was proposed that inhibitors of carboxypeptidase could be used for the design of ACE inhibition. It was also known that peptides from the venom of South American pit vipers were known to be inhibitors of ACE that convert ang I to ang II. However, these peptides lacked oral activity (Ondetti et al. 1971). The binding site of carboxypeptidase and ACE were compared (Fig. 1.9). ACE is a metalloenzyme that hydrolyzes the peptide bond from the C-terminus. The metal ion zinc participates in the enzyme reaction for hydrolysis of the peptide bond that converts angiotensin I to angiotensin II. The binding of the substrate to the enzyme involves a negatively charged C-terminal of the substrate interacting with a positively charged amino acid from the enzyme. A zinc ion interacts with the carbonyl carbon of the peptide bond,

Fig. 1.9 Design of ACE inhibitors for control of blood pressure. (a) Binding site of ACE, (b) carboxypeptidase A with benzyl succinic acid, (c) a peptide from Brazilian viper venom that could control the blood pressure and a peptide derived from scanning and modification that can inhibit ACE, (d) succinyl proline designed based on a peptide from Brazilian viper venom and benzyl succinic acid to inhibit ACE, (e) and (f) modification of succinyl proline resulting in (g) captopril (Acharya et al. 2003), Foye's principles of med chem (Foye 2008)



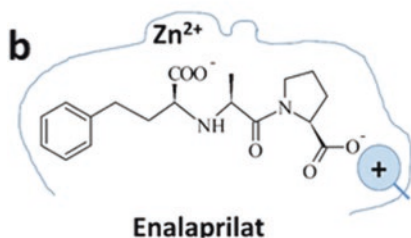
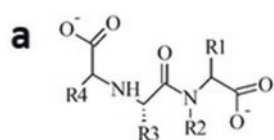
holding the substrate tight in the enzyme cavity. The hydrophobic groups proline and phenylalanine help to increase the binding affinity of the substrate to the enzyme. In addition to this, a hydrogen-bonding interaction between substrate and enzyme increases the affinity of the substrate for the enzyme. Based on this, three different approaches were used: (a) sulfhydryl-containing inhibitors, (b) dicarboxylate-containing inhibitors, and (c) phosphonate-containing inhibitors (Lemke and Williams 2012).

It is also known that benzyl succinic acid is an inhibitor of carboxypeptidase and, hence, can also be used as an inhibitor of ACE (Fig. 1.9). By combining the C-terminal part of pit viper venom (proline) and benzyl succinic acid, 2-methylsuccinyl-L-proline was designed. However, the potency of this compound was well below that of peptides derived from a pit viper. Amino acid proline was replaced with other amino acids, but structure-activity (SAR) studies indicated that

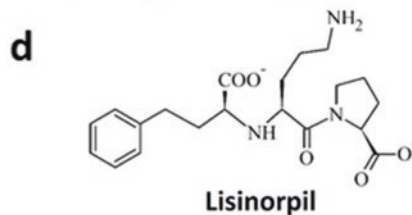
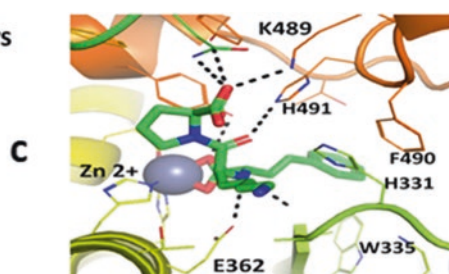
proline was necessary for the activity (Fig. 1.9). When the succinyl carboxylate was replaced with sulfur, which binds to zinc with high affinity, the compound showed a potency that was 10 to 20-fold greater than that of the viper peptide. The addition of the 2-methyl group further enhanced the activity, and the resulting compound was called captopril (Fig. 1.9). Captopril became a successful drug in controlling high blood pressure in patients. However, because of the presence of the sulfhydryl group in its chemical structure, captopril had a metallic taste and produced rashes in some patients. Thus, next-generation peptidomimetics were designed to overcome these effects.

The idea was to design compounds that lack the sulfhydryl group yet bind to zinc with high affinity. The compounds designed were tripeptide analogs. The C-terminal two amino acids were retained (Ala-Pro), but the N-terminal amino acid was replaced isosterically by the N-carboxyl methyl group. The use of a phenylethyl group at the N-terminal resulted in enalaprilat (Fig. 1.10). The compound showed excellent activity via IV administration, but oral bioavailability was poor, and hence, a prodrug enalapril was used. Eight other analogs were designed to improve the activity and oral bioavailability. Among them, lisinopril has a basic amino acid lysine, and this peptidomimetic is orally available. The crystal structures of ACE complexed with substrate peptide as well as inhibitors such as lisinopril have been elucidated, and details of the interaction of the enzyme with its substrate have been described (Masuyer et al. 2012; Sturrock et al. 2004). The interaction of angiotensin II peptide bound to ACE is shown in Fig. 1.10 with hydrogen bonding and

Dicarboxylate-containing inhibitors of ACE



Enalaprilat



Lisinopril

Fig. 1.10 Design of dicarboxylate-containing inhibitors of ACE. (a) Template structure for the design, (b) enalaprilat, (c) crystal structure of ACE complexed with lisinopril (PDB ID:2c6n), (e) terminal hydrogen bonding/electrostatic interaction by Lys (positive charge), hydrogen bonding by E362 and H491, and hydrophobic interaction of phenyl ethyl group with F490 H331 and W335 are shown, (d) structure of lisinopril

electrostatic and hydrophobic interactions. Lisinopril is a peptide-based drug designed to decrease the production of angiotensin II (Corradi et al. 2006; Masuyer et al. 2012). From Fig. 1.10, it is clear that lisinopril binds to the enzyme with zinc metal ion with an electrostatic interaction and phenyl and prolyl groups with hydrophobic interactions. Notice that the backbone of the peptide and side chain is modified in the case of lisinopril so that the drug does not undergo hydrolysis caused by the enzyme inhibiting the enzymatic reaction. Lisinopril competitively binds to ACE and slows the production of angiotensin II, thus controlling the high blood pressure.

1.10.2 Receptor Ligand-Based Drug Design

An example of receptor-based peptide design is cilengtide, an FDA-approved cyclic peptide drug used to treat glioblastomas, a form of brain cancer. The structural design and development aspect of this peptide as a drug candidate is described by Prof. Kessler's group in an article by Mas-Moruno et al. (2010). The peptide was designed before the details of molecular biology, and the crystal structures of the receptors and ligand proteins were elucidated. The designs described here are not in chronological order; rather, they are described from biochemical and structural aspects to make it easier to understand the design of peptides in a sequential process.

1.10.2.1 Biochemical Basis of the Design

Integrins play a key role in cell-cell and cell-extracellular matrix (ECM) adhesion and control the cell adhesion through the signaling process. Integrins are a family of proteins that perform a multifaceted function by a combination of heterodimers. Each integrin structure has an alpha and beta subunit that non-covalently interacts to form 24 transmembrane dimers by a combination of 18 alpha and 8 beta-integrin subunits. The alpha and beta subunits have a large extracellular domain, a short cytoplasmic tail, and a single transmembrane domain (Fig. 1.11) (Desgrosellier and Cheresh 2010). The protein-protein interaction of integrins with ECM proteins results in cell adhesion and migration, signaling pathways for cell proliferation, survival, and differentiation, as well as cytoskeletal organization in cells (Hood and Cheresh 2002). Detailed structural studies revealed that integrins exist in different conformational states in ligand-bound and unbound forms and, thus, control the signaling process from the outside of the cells to the inside. Apart from outside-in signaling, they are also known to respond to intracellular signaling, which controls the cell adhesion to ECM, and cell invasion and migration; hence, they also perform the "inside-out signaling" process (Eilken and Adams 2010). Thus, blocking integrin interaction with its ligand will have an effect on cell invasion, migration, and angiogenesis and the detachment of cells that leads to apoptosis. In vertebrates, depending on the type of alpha and beta subunits, the 24-subunit family can be

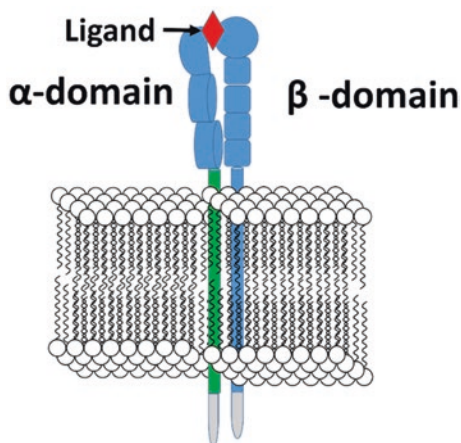


Fig. 1.11 A schematic diagram of a proposed model of integrin extracellular domains, transmembrane (TM), and cytoplasmic domains. The alpha and beta subunits are in activation and deactivation state and bind to the ligand. The ligand-binding site has an RGD sequence in the ligand. Peptides based on RGD sequences were designed to modulate integrin-ECM interactions (Zheng and Leftheris 2020; Mas-Moruno et al. 2010)

classified into collagen receptors, leukocyte-specific receptors, laminin receptors, and RGD receptors. Among these, $\alpha V\beta 3$, $\alpha V\beta 5$, and $\alpha 5\beta 1$ are involved in metastasis and angiogenesis in cancer tumors (Desgrosellier and Cheresh 2010).

The groundbreaking work of determining how the integrin interacts with some of its ligands using a short sequence of peptide stretch, namely, Arg-Gly-Asp (RGD motif), was done by Ruoslahti and Pierschbacher (Pierschbacher et al. 1981, 1983; Pierschbacher and Ruoslahti 1984a). The RGD motif is recognized by integrins in fibronectin, fibrinogen, vitronectin, and osteopontin. As mentioned earlier, integrin heterodimers are involved in different functions, and we will focus our attention on their involvement in cancer and the discovery of cilengitide, an RGD peptide for cancer therapy. The integrins that are expressed on epithelial and endothelial cells and contribute to tumor progression are $\alpha V\beta 3$, $\alpha V\beta 5$, and $\alpha 5\beta 1$. Integrin expression can vary between normal and tumor tissue, and this fact makes it a targetable molecule for cancer therapy. Tumor cell expression of the integrins $\alpha v\beta 3$, $\alpha v\beta 5$, $\alpha 5\beta 1$, $\alpha 6\beta 4$, $\alpha 4\beta 1$, and $\alpha v\beta 6$ is correlated with disease progression in various cancer tumor types, and in particular, $\alpha v\beta 3$ and $\alpha v\beta 5$ have implications in tumor invasion (Bello et al. 2001). Positron emission tomography (PET) images and the immunohistochemistry of patients with solid tumors has indicated that $\alpha v\beta 3$ expression in cancer tissue but not in normal tissues (Bello et al. 2001). Due to their primary expression on activated endothelial cells, the integrins $\alpha v\beta 3$, $\alpha v\beta 5$, and $\alpha 5\beta 1$ are attractive targets for cancer therapy and the treatment of nonmalignant angiogenic disorders. Especially in the case of solid tumors, antiangiogenic molecules represent a new potent concept of therapy. The inhibition of integrin-ligand interactions suppresses cellular growth and induces apoptotic cell death (Kerr et al. 2002; Mas-Moruno et al. 2010).

1.10.2.2 Identification of a Binding Epitope

The idea was to block the integrin-receptor interaction using drug-like molecules and, in this case, by peptides since the interacting molecules are proteins and they make contact via peptide epitopes. The first step in the drug design was to identify peptide epitopes involved in protein-protein interactions. Since the structure of the complex of integrin receptor was not known at that time, the entire protein is needed to be scanned for binding epitopes. One of the approaches used was directing antibodies to disrupt the interaction of two proteins and hence cell adhesion. As described earlier, using antibody mapping a large region of protein fibronectin that is important in binding to the integrin was identified (Knudsen et al. 1985). The cell-binding region of fibronectin was identified in a protein fragment of 170–200 kD and contained within an 11.5 kD fragment. To pinpoint the residues, Pierschbacher et al. prepared synthetic peptides of around 30 amino acids covering the 108 amino acids of fibronectin. Among the four synthetic peptides, one from the C-terminal stretch of 11.5 kD protein fragments was responsible for cell adhesion (Pierschbacher et al. 1983). Using mapping and antibody approach, it was determined that a small region of protein, namely, Arg-Gly-Asp-Ser, was the adhesion site of the integrin-ligand interactions. Thus, a small fragment of a peptide that was important in binding to the integrin was identified. If one could mimic the RGDS fragment in protein fibronectin by synthesizing a short peptide, that peptide could bind to integrins and inhibit the integrin-ligand interaction and, hence, cell adhesion. The molecules that could be designed based on RGDS peptide could be designed as drugs to inhibit the biochemical pathways of angiogenesis, invasion, and migration and thus used for cancer treatment. At this stage, a simple rapid assay is needed to screen the peptides designed. In their work, Pierschbacher et al. employed a cell attachment assay using human fibroblast or rat kidney cells and a flat-bottom microplate reader (Pierschbacher et al. 1983). Tranqui et al. (1989) used a platelet aggregation assay. Once the compounds were screened for cell adhesion or aggregation inhibition activity, they were then screened for binding to specific integrin receptors.

Integrins bind to different ligands of ECM proteins, and it was discovered that the RGD motif was found in vitronectin, osteopontin, collagens, von Willebrand factor, fibrinogen, thrombospondin, and laminin. However, integrins were able to distinguish different ECM proteins with the same binding motif. How was that possible? Detailed analysis of binding of different proteins to integrins, the amino acid sequence around the RGD motif, and the conformation of the RGD motif in different ECM provided information about the ability of integrins to bind to different ECM with the RGD motif but distinguish between different proteins (Ruoslahti and Pierschbacher 1987; Humphries 1990). The same analysis also indicated that the Ser residue in RGDS motif can be replaced with different amino acids for binding to ligands (in this case ligands refer to different ECM protein ligands) (Pierschbacher and Ruoslahti 1984b). Initial work on RGD peptides was related to inhibition of platelet aggregation to evaluate the ability of RGD peptides for cell adhesion inhibition and enzyme-linked immunosorbent assay (ELISA) to evaluate the specificity of RGD peptide to bind to a particular integrin. Modifications of RGD peptides and

Table 1.2 Platelet aggregation assay. Binding of platelets to endothelial cells and their inhibition by peptides (Tranqui et al. J. Cell Biol. 1989). Amino acid residue S in RGDS was replaced with different amino acids, and platelet aggregation inhibition was evaluated

Peptides	IC ₅₀
RGD	>1000
RGDS	55
RGDQ	45
RGDC	35
RGDV	18
RGDF	9
NIMEILRGDF	160
ILRGDFSSAN	>500

Table 1.3 Cyclic RGD peptides designed with conformational constraint and D amino acids binding to vitronectin (Aumailley et al. 1991; Mas-Moruno et al. 2010; Muller et al. 1994; Dechantsreiter et al. 1999)

Peptides	Comments
GRGDSPK	Linear peptide
RGDFv	D-amino acid substitution
Cyclo(rGDFV)	Cyclic, D-Arg
Cyclo(RGdFV)	D-Asp
Cyclo(RGDfV)	D-Phe
Cyclo(RGDFv)	D-val
Cyclo(N(Me)RGDfV)	N-methyl-Arg, D-Phe
Cyclo(R-N(Me)GDfV)	N-methyl-gly, D-Phe
Cyclo(RG(N-Me)DfV)	N-methyl-Asp, D-Phe
Cyclo(RGD(N-Me)fV)	N-methyl-D-Phe
Cyclo(RGDf-N(Me)V)	D-Phe, N-methyl-Val

flanking residues were done with different amino acids. Only six amino acid peptides were shown to be potent with cyclic structure. Tranqui et al. (1989) used RGD peptides along with fibrinogen γ -chain peptides to evaluate the adhesion and inhibition of platelets and endothelial cells. They modified the amino acid residue next to Asp in the RGD sequence and generated tetrapeptides and determined that the RGDFV sequence is important in binding to integrins (Tranqui et al. 1989) (Tables 1.2 and 1.3). Pierschbacher and Ruoslahti studied the influence of stereochemistry on Arg-Gly-Asp-X-Pro-Cys peptide based on a peptide from a fibronectin cell attachment site. When L-Arg was replaced with D-Arg, there was no change in the inhibition activity of the peptide. When L-Asp was replaced with D-Asp, activity was lost. The same group worked on the substitution of different amino acids at position X in the peptide Arg-Gly-Asp-X-Pro-Cys and found that the substitution of different amino acids has an effect on selectivity to receptors (Pierschbacher and

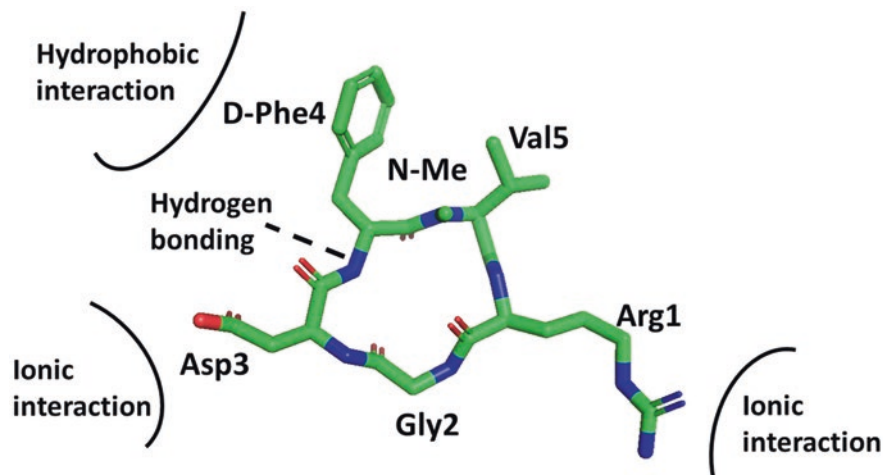


Fig. 1.12 Crystal structure of cilengitide (PDB ID: 115g) with proposed interaction with the receptor. Structure-activity relationships were established based on the proposed interaction with the receptor as shown in the figure

Ruoslahti 1987). The Kessler group cyclized the peptide with RGDfV sequence (Table 1.2) that had exhibited potent inhibition activity and generated several peptides to find the conformation necessary for binding to different ligands (Mas-Moruno et al. 2010). Linear peptides have a flexible structure, but cyclic peptides have restricted conformation due to backbone cyclization. A concept called “spatial screening” was used to investigate different conformations of cyclic penta- and hexapeptides of the RGD motif in which chirality of the amino acids was changed and the amino acids flanking the RGD motif were replaced with different amino acids. Different peptides designed are shown in Table 1.3. Additional D-amino acids were introduced to evaluate the effect of chirality of amino acids on binding to integrin receptors. This comprehensive analysis indicated that in peptide cyclo (RGDfV) with a D-Phe, the RGD motif formed a kink around Gly, and this peptide has an optimal conformation to bind to $\alpha V\beta 3$ but is selective toward $\alpha IIb\beta 3$ integrin receptor (Mas-Moruno et al. 2010). Several modifications are described in a series of publications (see references in Mas-Moruno et al. (2010)). Structure activity studies on cyclo (RGDfV) suggested that D-Phe along with Gly and Asp is important for activity, but amino acid at position 5 has no effect on the activity (Fig. 1.12). This structure was further modified with amide bond substitution, beta-turn mimetics, sugars, and retro-inverso analogues (Geyer et al. 1994; Haubner et al. 1996; Lohof et al. 2000; Wermuth et al. 1997).

1.10.2.3 N-Methylation of Cyclic Peptide

N-methylation of the peptide bond has been used to increase the stability and pharmacokinetic properties as well as the potency of peptide-based drugs. Furthermore, N-methylation restricts the conformation of peptides as the methyl group at the nitrogen restricts the rotation of the psi bond (Chatterjee et al. 2008b; Biron et al. 2008; Chatterjee et al. 2006). Thus, N-methylation scanning was done on the cyclic RGDfV peptide. The peptide containing N-methyl Val had antagonist activity with an IC_{50} value of 0.58 nM for $\alpha V\beta 3$ and showed 1500 times more selectivity in inhibiting vitronectin to $\alpha V\beta 3$ than fibrinogen (Fg) to $\alpha IIb\beta 3$ (Fig. 1.13). Detailed structural analysis of this compound was carried out with NMR and molecular modeling (Mas-Moruno et al. 2010). Later, the crystal structure of this compound in a complex with integrin was elucidated (Xiong et al. 2002). The compound was named cilengitide (Fig. 1.13).

1.10.2.4 Preclinical and Clinical Evaluation

Several in vitro and in vivo studies were carried out on cilengitide to characterize this molecule (Mas-Moruno et al. 2010; Reardon et al. 2011). Cilengitide is known to be a potent inhibitor of the interaction between integrins and their ECM ligands, resulting in inhibition of angiogenesis and, hence, inducing apoptosis of growing endothelial cells. It was shown to influence cellular adhesion to $\alpha v\beta 3$ ligands, to

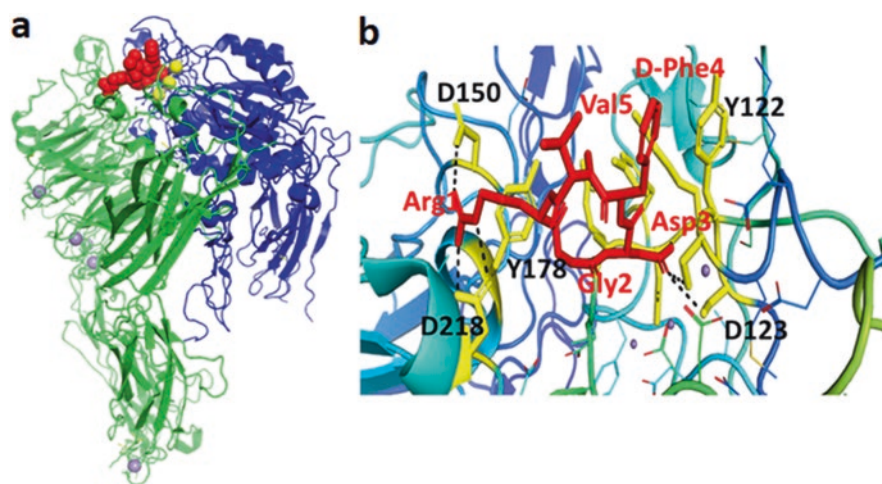


Fig. 1.13 Crystal structure of the complex of integrin $\alpha 5\beta 3$ with cilengitide. (a) Extracellular segment of the protein with cilengitide (red spheres) is shown. Subunits $\alpha 5$ and $\beta 3$ are shown in green and blue. (b) Details of the interaction of cilengitide (red sticks) with the receptor (PDB ID: 115g). Amino acids from the protein are shown with single letter amino acid codes. Amino acids from cilengitide are shown with three-letter code

induce increased apoptosis after detachment of $\alpha v\beta 3$ and $\alpha v\beta 5$ expressing cells in vitro, and to block the growth of cancer tumors in a xenograft mice model. The inhibition of αv integrins resulted in a significant reduction of functional vessel density and retardation of tumor growth and metastasis in vivo (Paolillo et al. 2009; MacDonald et al. 2001). Cilengitide was formulated as a sterile aqueous solution and administered as an IV injection. Pharmacokinetic studies using the radioactively labeled compound in animal studies showed that 90% of the drug was unchanged in urine for 72 h. In clinical trials, the drug reached maximal plasma concentration in 1 h after injection and elimination half-life of 3–5 h (Reardon et al. 2011). Cilengitide is used for the treatment of glioblastoma, and the [European Medicines Agency](#) granted it [orphan drug](#) status. However, a large phase III clinical trial of cilengitide failed to show improvement in patients with newly diagnosed glioblastoma, and hence, the development of cilengitide was halted and targeting $\alpha v\beta 3$ for cancer therapy is still under investigation (Alday-Parejo et al. 2019; Stupp et al. 2014; Chinot 2014).

1.10.3 Example of a Peptide from Natural Resources: Polymyxin B

Polymyxins are cationic antimicrobial peptides used as the last-line therapy to treat multidrug-resistant gram-negative bacterial infections. Polymyxin B is a lipopeptide antibiotic isolated from *Bacillus polymyxa*, a gram-positive bacterium capable of fixing nitrogen. It is found in soil, plant tissues, marine sediments, and hot springs. In nature, it is produced to protect plants as bacteria produce diverse antimicrobial peptides that suppress the growth and fitness of human and plant pathogenic microbes. These types of lipopeptides were discovered more than 50 years ago (Ainsworth et al. 1947). Polymyxin is produced by fermentation by inoculating *Bacillus polymyxa* into a fermentation medium. Several methods of synthesis of the lipopeptide have also been reported (Vogler et al. 1965; Sharma et al. 1999; de Visser et al. 2003).

Polymyxins have a narrow therapeutic window due to their toxicity and are used as the last line of therapy for multidrug-resistant pathogens. Early approval of polymyxin E for gram-negative bacteria was discontinued in 1970 because of nephrotoxicity (Koch-Weser et al. 1970; Poirel et al. 2017). However, the use of polymyxin was reintroduced as a last line of therapy (Li et al. 2006; Biswas et al. 2012) due to the prevalence of multidrug-resistance gram-negative bacteria. Polymyxins are cationic polypeptides containing a cyclic heptapeptide with a tripeptide side chain acylated at the N-terminal by a fatty acid tail. Several polymyxins have been isolated and identified, and since they are from natural sources, each has a difference in their chemical structure of the fatty acyl group. Among these groups, polymyxin B and E groups are extensively studied due to their therapeutic effect against gram-negative bacteria. Commercially available polymyxins contain polymyxin B1 and B2 and polymyxin E1 and E2 as their major components (Orwa et al. 2001a, b).

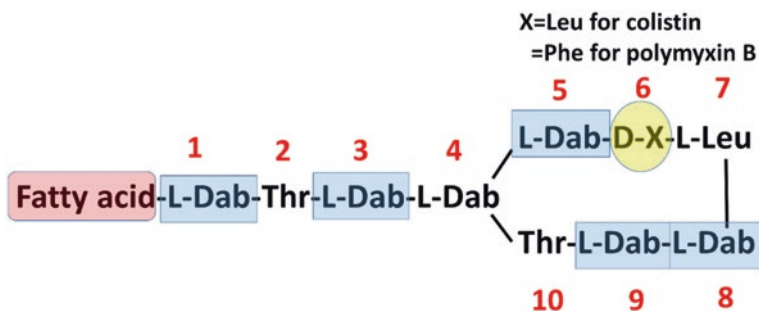


Fig. 1.14 The general structure of polymyxins with a numbering scheme. The fatty acid at the N-terminal is different for different polymyxins: 6-methyloctanoic acid for colistin A and polymyxin B1 and 6-methylheptanoic acid for colistin B and polymyxin B2. Dab, diaminobutyric acid

Polymyxins are non-ribosomal cyclic lipopeptides having a general structure shown in Fig. 1.14. These are cyclic decapeptides with an N-terminal fatty acid, a linear tripeptide peptide sequence, and a cyclic heptapeptide. The peptide structure has five diaminobutyric acid (Dab) residues, which imparts a negative charge to the molecule at physiological pH (polycationic at pH 7.4). There are hydrophobic motifs at positions 6 and 7, which is important in binding to serum proteins; position 6 has a D-amino acid, and all other amino acids have L-configuration. Polymyxin B peptides have D-Phe amino acid at position 6, and seven individual polymyxin B components have been identified; these differ from one another by branched and non-branched fatty acyl groups at the N-terminal, varying in length from 7 to 9 carbons, and are labeled polymyxin B1–B6 (Velkov et al. 2010).

The polymyxin E peptides have D-Leu at position 6, and the individual components of the polymyxin E group contain structurally distinct branched and non-branched N-terminal fatty acyl groups that vary in length from 7 to 9 carbons and are labeled polymyxin E1, E2, E3, E4, E7, and E8-Ile (Velkov et al. 2010).

1.10.3.1 Mechanism of Action

Polymyxins are antimicrobial peptides (AMPs), and the activity of AMPs stems from their interaction with microbial membranes to cause disruption of the physical integrity of the membrane. Once the membrane integrity is disrupted, AMPs translocate into the cytoplasm and act upon intracellular targets (Hancock and Sahl 2006). The main difference between the mammalian cell membrane and the bacterial cell membrane is the reason for their selectivity toward microbes. Most of the AMPs are positively charged. In gram-positive and gram-negative bacteria, the cytoplasmic membranes are rich in lipids such as the phospholipids phosphatidylglycerol, cardiolipin, and phosphatidylserine, which results in a negative charge. However, in gram-positive bacteria, the cytoplasmic membrane is covered with a thick peptidoglycan layer, and hence, negative charge is masked. In gram-negative

bacteria, the cytoplasmic membrane is covered by a thin peptidoglycan layer, and the negative charge is exposed. Thus, gram-negative bacteria can attract positively charged AMPs. A mammalian cell has phospholipids phosphatidylethanolamine, phosphatidylcholine, and sphingomyelin, providing a membrane with a neutral net charge (zwitterionic). In addition to this, the mammalian cell membrane has an asymmetric distribution of lipids, with outer leaflet zwitterionic lipids and inner leaflet negatively charged phospholipids, making the inner layer negatively charged and outer layer neutral. Thus, AMPs are not attracted to mammalian cell membranes. AMPs can interact with mammalian cell membranes via hydrophobic interaction since AMPs have hydrophobic amino acid residues, but hydrophobic interactions are weak compared to electrostatic interactions (Fjell et al. 2011). Mammalian cells contain cholesterol, which stabilizes the membrane compared to the microbial membrane (Lai and Gallo 2009).

Detailed interactions of AMPs with the cell membrane are proposed, and their translocation process into the membrane is being elucidated (Fjell et al. 2011; Brogden 2005; Nguyen et al. 2011). AMPs with charged and hydrophobic amino acids in their sequence make contact with negative charges of the membrane, forming a helical structure that exposes the hydrophobic groups outside and hydrophilic groups inside the helical structure. The peptide inserts into the membrane hydrophobic core as a helical structure and goes deep into the membrane. It perturbs the membrane structure and forms a pore in the membrane that leaks the ions and metabolites from inside the cell; thus, the microbial cell is lysed. The peptide, once inside the microbial cell, also targets protein synthesis, enzyme activity, and DNA/RNA activity.

Polymyxin has a positively charged α , γ -diaminobutyric acid (Dab) residue. Five of these γ -diaminobutyric acid (Dab) residues can interact with the membrane phospholipids, which are negatively charged in gram-negative bacteria. The lipopolysaccharide (LPS) in microbes is stabilized by divalent cations Ca^{2+} PLUS_SPI and Mg^{2+} PLUS_SPI. Polymyxin positively charged Dab interaction with the membrane displaces these divalent cations and causes LPS destabilization, leading to leaking of microbial cells and cell lysis. In the case of polymyxin, LPS of gram-negative bacteria is the proposed target. Further, polymyxin is known to target respiratory enzymes such as NADH-quinone oxidoreductase (Dixon and Chopra 1986; Deris et al. 2014).

1.10.3.2 SAR

Length of Fatty Acid Chain and Substitution

SAR data indicate that a hydrophobic substituent at the N-terminus (length of chain and bulkiness of the substituent) is important for antimicrobial activity. The LPS-binding affinity of the polymyxin correlates with the hydrophobicity of the N-terminal substituent. Modified polymyxin cubist with a shorter N-terminal group

had *in vitro* and *in vivo* antibacterial activity comparable to that of polymyxin B and colistin (Velkov et al. 2010; Okimura et al. 2007).

The Positive Charge of Dab Side Chains

Dab side chain with its cationic nature is known to be important in binding to LPS in microbe membranes and exerting the action of polymyxins. There are five Dab residues in the molecule. Overall, positive charges can be reduced on the molecule, and it can still retain the activity. Dab1 is important in polymyxin, and those within the cyclic peptide are essential for polymyxin antibacterial activity (Vaara et al. 2008).

The Linear Tripeptide Segment

The linear tripeptide segment (Dab1-Thr2-Dab3) links the heptapeptide cyclic core of the polymyxin molecule to the N-terminal fatty acyl chain. As indicated above, Dabs contribute to positive charges and interaction with LPS. Deletion of Dabs in this region suggested that the tripeptide segment can only be truncated by deletion of Dab1 with a negligible loss of antibacterial activity (Vaara et al. 2008).

The Hydrophobic Motif of the Heptapeptide Ring

The hydrophobic segment in the cyclic structure D-Phe6-L-Leu7 is highly conserved across the naturally occurring polymyxins. These residues seem to be important in binding to plasma protein, increasing the stability of the peptide and its residence time *in vivo*. The introduction of fatty acyl amino acid derivatives at these positions is shown to improve antimicrobial activity, LPS binding, and plasma protein binding, whereas the introduction of hydrophilic groups or β -turn mimetics appears to negatively impact antimicrobial activity (Kanazawa et al. 2009).

The Heptapeptide Backbone

Deleting the amino acids in the heptapeptide ring or increasing the size of the ring by addition of amino acids/functional groups resulted in a decrease in the antimicrobial activity of polymyxin. The molecular model of the polymyxin B-LPS complex indicated that the heptapeptide backbone with 23 atoms forming a ring is the most suitable ring size for binding to LPS and provides the most ideal structural configuration for potent antimicrobial activity (Tsubery et al. 2000).

Formulations and Administration

Polymyxin B and E (polymyxin is also called colistin) are used in clinical practice. Colomycin and Coly-Mycin M are the parenteral formulations of colistin. Colistin sulfate can be used for different routes of administration, including oral tablets, liquid syrup, and topical form. Colistin sulfate and sodium colistin methanesulfonate (CMS) are also formulated for inhalation therapy (Tsubery et al. 2000; Biswas et al. 2012; Poirel et al. 2017; Yahav et al. 2012).

Resistance

Polymyxin is used as a therapeutic agent for treating multidrug-resistant bacterial infections. Does polymyxin develop resistance? There are some microbes that are resistant to polymyxin, and a majority of the studies are performed on *P. aeruginosa*. Bacteria develop resistance to polymyxin by reducing the negative charge in the membrane. For example, in *P. aeruginosa*, salmonella, the modifications in the lipid A with 4-amino-4-deoxy-L-arabinose (L-Ara4N), and/or phosphoethanolamine (PEtn) is known to reduce the negative charge and cause weaker interaction of the drug with the microbial membrane and, hence, resistance. Other mechanisms of resistance, including genetic modification, are proposed (Helander et al. 1994; Mathur and Waldor 2004; Brodsky et al. 2002).

Toxicity and Modification of Polymyxin

Polymyxin B and E both exhibit toxicity. The currently available polymyxins (polymyxin E (colistin) and polymyxin B) have the potential to cause nephrotoxicity and neurotoxicity, which limits the dosing of polymyxin and, thus, the last line of therapy for patients infected with antibiotic-resistant microbes. Some of the toxicity is observed during the therapy and can be reversed after the completion of the therapy. To reduce the toxicity, several modifications of polymyxin were suggested. These include reduction in positive charge on the molecule, modification of the N-terminal region, substitution of Dab at a particular position with diaminopropionic acid, and modification in the cyclic heptapeptide (Velkov et al. 2010; Vaara 2013; Brown and Dawson 2017; Brown et al. 2019).

1.10.4 Protein-Protein Interaction and Drug Design

Here, we provide an example of peptide and grafted peptide drug design targeted at proteins involved in an immune response. The cells of the immune system are responsible for the ability of vertebrate animals to combat invading microbes. The human immune system has specialized cells to distinguish self from nonself. Self is

preserved, and nonself is destroyed. This process is carried out by a highly orchestrated process using the cells of the immune system and protein molecules such as antibodies and cytokines generated by the immune cells. In the adaptive immune system, foreign antigens are destroyed with the help of T cells, which are the major players. In some cases, T cells attack the body cells, creating a disease state. This phenomenon of T cells attacking the body's tissues and destroying their results in several disease states is called "autoimmune disease." Thus, T cells should be able to distinguish between "self" and "nonself" before neutralizing or destroying other cells (Kamradt and Mitchison 2001; Jois et al. 2006). In the case of autoimmune diseases, T cells attack the healthy cells and destroy them, resulting in a disease state. For example, in the case of diabetes mellitus (type I), T cells attack the beta cells, and insulin production is significantly decreased. Multiple sclerosis is caused by the destruction of the myelin membrane, and systemic lupus is caused by T cells attacking different organs. Another major disease caused by T cells is rheumatoid arthritis, where T cells attack the joints and destroy the cartilage, causing bone erosion. In such cases, modulation of unwanted immune response is needed; thus, protein-protein interactions involved in the immune response are modulated by peptides or peptidomimetics as possible therapeutic agents.

T cells recognize antigen-presenting cells (APC) with the help of T-cell receptors (TCR) on T cells and major histocompatibility complex (MHC) on antigen-presenting cells such as dendritic cells (DC). T-cell activation requires two signals, recognition of an antigen by the TCR (signal 1) and a concomitant signal provided by adhesion/co-stimulatory molecules (signal 2) to achieve full activation of the T cells (Chen and Flies 2013; Bretscher 1999). This co-stimulation is vital for optimal immune response and involves multiple PPI between receptors of APC and T cells. Among these multiple PPIs for co-stimulation, LFA-1/ICAM-1, CD2/LFA-3 (CD2/CD58), CD28/B7 (CD28/CD80), and (CD11a-CD18)/CD54 are some of the major co-stimulatory interactions between T cells and APC (Chen and Flies 2013; Bakdash et al. 2013; Chambers and Allison 1997) (Fig. 1.15). Any of these protein-protein interactions can be modulated by peptides, small molecules, or antibodies to alter the unwanted immune response and make them suitable for use as therapeutic agents to treat autoimmune diseases.

1.10.4.1 Biochemical Pathway

Among the adhesion/co-stimulatory molecules, CD2 and CD58 molecules have been shown to be important in inflammatory diseases, and there is an upregulation of CD58 in inflammatory diseases (Mrowietz 2002; Chamian et al. 2005; Balanescu et al. 2002). CD2 is a transmembrane protein on T cells that binds to its ligand CD58 on APC (CD48 in mice) (Chen and Flies 2013; Davis et al. 2003; van der Merwe and Davis 2003). CD48 has a high degree of homology to CD58 (~60%) and is similar in 3D structure CD58 (Wang et al. 1999; Velikovskiy et al. 2007). Upon binding of CD2 to CD58, CD2 generates co-stimulatory signals via its cytoplasmic tail, resulting in calcium flux and induction of the release of cytokines. Subsequently,

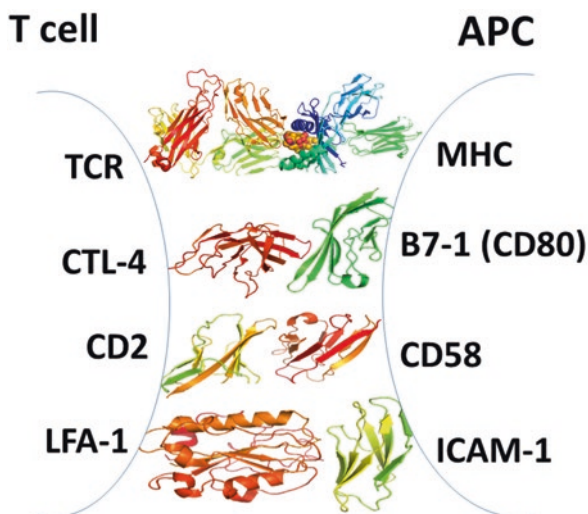


Fig. 1.15 A schematic diagram of the TCR-MHC complex along with co-stimulatory molecules that help the adhesion of T cell to APC to generate a signal for immune response. CD2-CD58 interactions are known to be involved in the early stage of immune response, bringing the cells close, after which the other co-stimulatory molecules come in contact with each other. TCR-MHC (PDB ID: 4grl), CTL-4-B7 (CD80) (PDB ID: 1i8l), CD2-CD58 (PDB ID: 1qa9), (PDB ID: 1t0p)

other adhesion molecules such as LFA-1 and ICAM-1 bind to each other, strengthening the cell adhesion and, hence, T-cell APC contact (Espagnolle et al. 2007; Kaizuka et al. 2009). Alefacept is a recombinant human CD58-Ig fusion protein that effectively binds to CD2 and prevents CD2 interaction with CD58. It was successfully used clinically to treat plaque psoriasis but has been withdrawn (why?) (Chamian et al. 2005; Rustin 2012). These findings make CD2 and CD58 molecules attractive targets for developing therapeutic agents for inflammatory diseases.

1.10.4.2 Design of Peptides

First-Generation Peptides

Since the biochemical pathway for CD2 and CD58 interaction is known and the crystal structure of the individual protein molecules, as well as the complex, is elucidated, we will next look at the details of PPI of CD2 and CD58 using the crystal structure. The CD58 binding domain of CD2 consists of β -strands with charged residues (Kim et al. 2001; Wang et al. 1999) (Fig. 1.16a and b). On examining CD2 crystal structure in the CD2-CD58 complex, the CD58 contact areas in CD2 involve β -strands and loops. The CD2 epitopes are mapped in F strands and two turns (FG loop and C'C'' loop) (Fig. 1.16b).

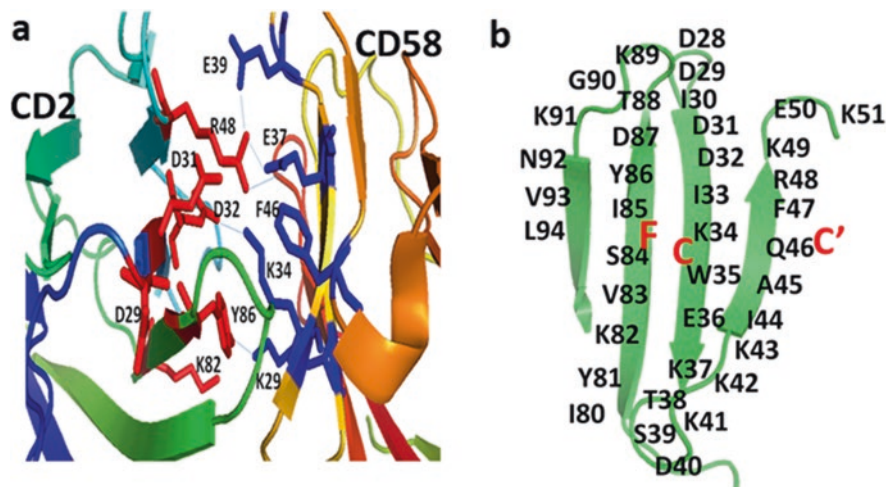


Fig. 1.16 (a) Crystal structure of CD2-CD58 complex showing adhesion domain and interface residues that are involved in hydrogen bonding and hydrophobic interactions. (b) Secondary structure elements that are important in binding to CD58 are labeled (F, C, C') with residue numbers (PDB ID: 1q9)

The design of the peptides was based on the following observations from previous studies. (i) CD2 and CD4 have similar 3D structures. Peptides from CD4 CC' loop β -turn have been shown to be important in binding to its receptor (Huang et al. 1997). (ii) Rat CD2 and human CD2 have similar 3D structures (backbone rmsd 0.9 Å) and have sequence homology (>60%). (iii) Mutagenesis studies on CD2/CD58 suggested that residues around the β -turn, β -strand (Fig. 1.16 and 1.17), and flanking residues of the β -turn at the interface between CD2 and CD58 are important for cell-cell adhesion (Wang et al. 1999). In human CD2 (hCD2), β -turns are at Arg48-Lys49-Glu50-Lys51, Asp87-Thr88-Lys89-Gly90 (Fig. 1.17), and Thr38-Ser39-Asp40-Lys41. Alanine scanning and mutagenesis studies indicated that in hCD2, replacement of Arg48 and Lys51 by alanine resulted in complete loss of binding of hCD2 to CD58. Gly90 in CD58 is also important in hydrogen bonding. The flanking residue to the β -turn, Asp87-Thr88-Lys89-Gly90, Tyr-86 residue in CD2 has been identified as the energetic hotspot region (Kim et al. 2001). Replacement of Tyr86 by alanine reduced the binding of hCD2 to CD8 by 1000-fold (Fig. 1.16 and 1.17). Therefore, it was hypothesized that stable β -turn structures containing these residues mimic the native CD2 and that these epitopes could be designed to modulate CD2-CD58 interaction. To start with, human and rat CD2 sequences were compared, and CD2-CD58 interface residues with beta-strand and beta-turn were chosen. Linear as well as cyclic peptides were designed. Pen and Cys were used for cyclization with Pen and Cys as flanking residues on either side of the β -turn sequence for the design of cyclic peptide. Cyclic constraint with disulfide bonds or peptide bonds at the two ends of the peptide sequence is known to stabilize

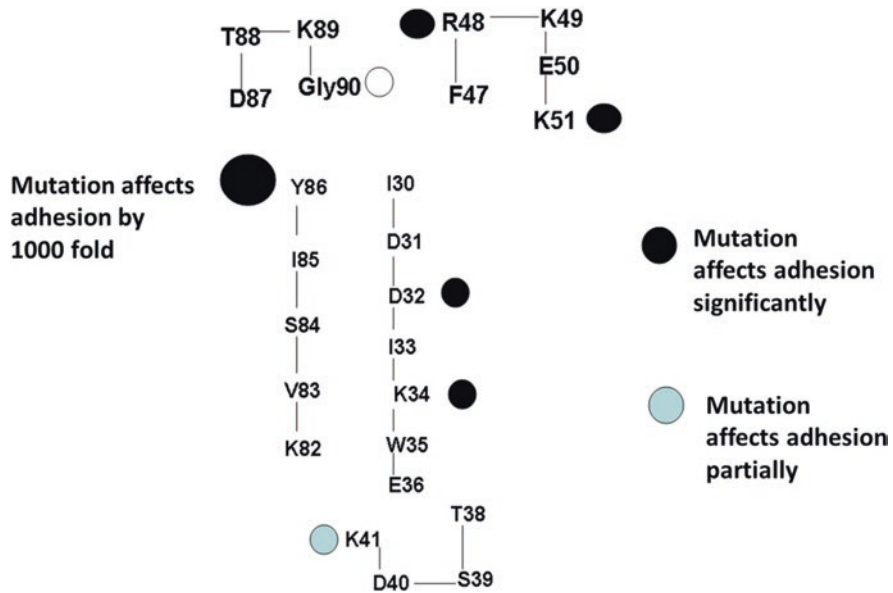


Fig. 1.17 The sequences of fragments of the secondary structure of CD2 that are important in binding to CD58 (F and C β -strands, β -turns in the FG loop, CC' loop, and C'C'' loop) are shown with residue numbers. Mutation of residues that affect the binding of CD2 to CD58 significantly is indicated by open and closed circles. Peptides were designed based on these results, as discussed in the text

the peptide conformation (Hruby 1993). For the control peptide, a 12-amino acid residue sequence was chosen from the hotspot region of CD2 (containing Tyr86) (Kim et al. 2001), and the sequence was reversed. Tyr86 and Tyr81 were replaced with Ala to generate the control peptide (Table 1.4).

Screening Assay

Next, an assay is needed to evaluate how the peptides we designed inhibit CD2 and CD58 interaction and adhesion of cells. E-rosetting is the most widely used method to identify T cells by CD2-CD58 interaction (Albert-Wolf et al. 1991). Sheep red blood cells (SRBC) express CD58 protein, while Jurkat leukemic T cells express CD2 protein on their surface. The ability of Jurkat cells to express CD2 was measured by flow cytometry assay. Binding of Jurkat cells to SRBC due to CD2 and CD58 interaction results in the formation of E-rosettes. Furthermore, assays based on fluorescence were also established (Liu et al. 2004; Satyanarayanajois et al. 2011): 1) Caco-2 cells and Jurkat cells (Caco-2 cells express CD58, while Jurkat cells express CD2 protein) and 2) OVCAR cells and Jurkat cells. The inhibitory activities of designed CD2 peptides were measured using fluorescently labeled Jurkat cells and a microplate fluorescence analyzer. Initial screening using cell

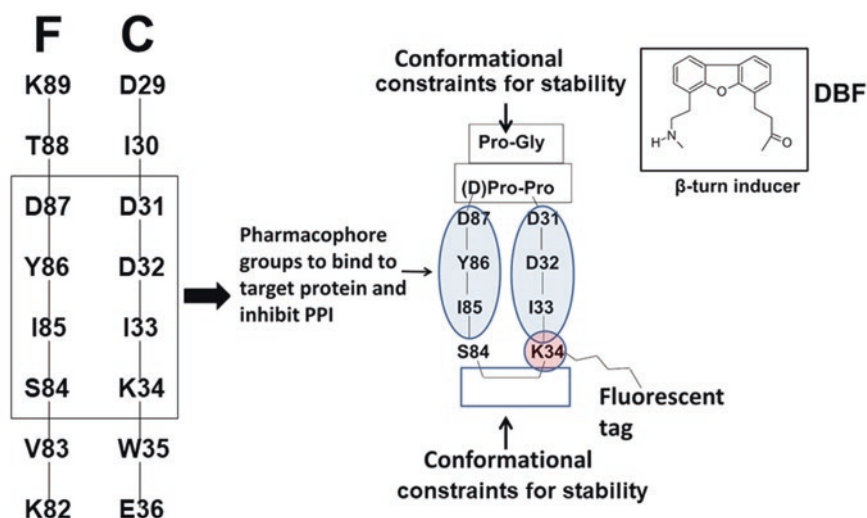
Table 1.4 First-generation peptides derived from CD2 protein. For residue numbers and secondary structures involved, please refer to Figs. 1.16 and 1.17. These peptides exhibited cell adhesion inhibition activity with $IC_{50} > 10 \mu M$

Name	Sequence	Comments
DS	DDIDDIKWEKTS-NH ₂	C strand (D31-S39)
SE	SDKKKIAQFRKE-NH ₂	C' strand (S39-E49)
IK	IYKVSIIYDTGK-NH ₂	F strand (I79-K91)
cYK	Cyclo(1,12)PenIYDTKKGKNVLC-OH	Cyclic version (85–94) pen at N-terminal Cys at C-terminal for disulfide bond
Control	KGKTDIAISVKAI-NH	Sequence reversed replaced Y by ala
IDD12	30IDDIKW35-V83SIYDT88	Peptide from F and C strands discontinuous region Linear (30 to 35; 83 to 88)
cDD12	Cyclo(1,12)-pen-DDIKW35-VSIYD87-C-OH	Cyclic version (31–35; 83–87)
IDD14	I30DDIKWE36-K82VSIYDT88	Linear (30 to 36; 86 to 88) sequences joined by clockwise orientation
cDD14	Cyclo(1,14)-pen-DDIKWE36-K82VSIYD87-C-OH	Cyclic (31 to 36; 82–87) sequences joined by clockwise orientation
cVW12c	Cyclo(1,12)-pen-S84IYDT88-I30DDIK34-C-OH	Cyclic (84 to 88; 30 to 34) sequences joined by anticlockwise orientation
cVW14c	Cyclo(1,14)-pen-S84IYDTK89-D29IDDIK34-C-OH	Cyclic (84 to 89; 29–34). Anticlockwise orientation

adhesion assay suggested that cyclic 12 mer peptides derived from the beta-turn region had better inhibition activity than the linear peptides. Truncation of these peptides resulted in the loss of activity of the peptides, but all of these peptides had inhibition activity with an IC_{50} value in the 10–20 μM range (Liu et al. 2005). Using the CD2-CD58 structure, the beta-strand region was used and the direction of the peptide chain was varied since the peptides are from the discontinuous region of the CD2 epitope (Liu et al. 2007). The ring size and chain direction of the beta-strands were optimized at this stage.

Second-Generation Peptides

In the first screening, several peptides were designed and evaluated for inhibition of cell adhesion activity (Gokhale et al. 2011; Gokhale et al. 2013; Gokhale et al. 2015; Sable et al. 2016). In the second-generation peptides, discontinuous epitopes from CD2 F and C strand were used, but conformational constraints using the Pro-Gly sequence were employed to induce beta-hairpin structure (Giddu et al. 2009). Instead of a disulfide bond, we used the backbone cyclization of the peptide (Satyanarayanajois et al. 2010). Furthermore, the same peptides were modified by the Pro-Pro sequence. Pro-pro sequences with different chirality are known to induce beta-hairpin structure and are more rigid compared to Pro-Gly beta-turn inducers (Fig. 1.18 and Table 1.5). A ten amino acid peptide with Pro-Pro sequence



β -strand region from protein CD2

Fig. 1.18 Schematic diagram of conformation constraints used for the design of second-generation peptides from F and C strands of CD2 protein. Important residues from F and C strands were joined by peptide bonds. Conformational constraints such as beta-turn could be introduced in the peptide by a Pro-Gly or D-Pro-Pro sequence or by an organic moiety dibenzofuran (DBF)

and F and C strands from CD2 with backbone cyclization showed cell adhesion inhibition activity with an IC_{50} value of 7 nM. Thus, a huge increase in activity was seen compared to that of the first peptides we designed. We used this peptide (namely peptide 6) to evaluate the in vivo activity in an animal model of collagen-induced arthritis (CIA). Details of this model and how to evaluate the peptide in the CIA model are available in the literature (Brand et al. 2007; Miyoshi and Liu 2018). Peptide 6 was able to suppress the T-cell immune response in cells derived from transgenic mice that develop arthritis similar to human arthritis (Gokhale et al. 2015; Gokhale et al. 2013). Peptide 6 was considered a “hit” compound and used for further improvement. We decided to investigate the importance of amino acids in peptide 6 by alanine scanning. Alanine scanning is widely used for the evaluation of the importance or functional role of amino acid residues present in peptides/proteins. From the alanine scanning studies, mutation at position Ile2 of the peptide resulted in an increase in potency to inhibit cell adhesion interaction. The most potent peptide from alanine scanning was further studied for its detailed three-dimensional (3D) structure using NMR and binding to CD58 protein using surface plasmon resonance and flow cytometry.

Structural Studies To confirm our hypothesis that the β -turn structures are important for the inhibitory activities of the peptides, the structures of the cyclic peptides

Table 1.5 Second and third-generation peptides from F and C discontinuous epitopes with conformational constraints. Amino acids/organic moiety introduced in the CD2 sequence as conformational constraints are highlighted. IC₅₀ values are obtained from cell adhesion inhibition activity of peptides between T cells and epithelial cells

Second-generation peptides	The sequence of the peptide	Comments	IC ₅₀ (μM)
1	Cyclo(1,12) PenSIYDPGDDIKC-OH	Disulfide bond cyclization and pro-Gly sequence used as a conformational constraint	
2	Cyclo(1,11) NH ₂ -E1-SIYDPGDDI-K11-OH	Cyclized by the side chain of E1 and K11	
3	Cyclo(1,10)EIYDPGDDIK	Main chain cyclization	
4	Cyclo(1,10) NH ₂ -E1-IYDPGDDI-K10-OH	Cyclized by the side chain of E1 and K10	
5	Cyclo (1,8) IYDpPDDK	D-pro-pro for β-hairpin, truncated form of 6	
6	Cyclo (1,10) SIYDpPDDIK	D-pro-pro for β-turn, S1-D4 to D7-K10 residues from CD2 protein sequence	0.0069
7	Cyclo(1,6)KDD-DBF-DYI	DBF for β-hairpin, sequence from protein CD2 used in an anticlockwise direction, truncated form of 8	0.0011
8	Cyclo(1,8) KIDD-DBF-DYIS	DBF for β-hairpin, sequence from protein CD2 used in an anticlockwise direction	>100
Third-generation peptides			
SFTI-a	Cyclo(<u>CKASAPPSCYDGDD</u>)	Peptide 6 grafted to SFTI-1	0.043
SFTI-b	Cyclo(<u>CKAEAKPSCYDGDD</u>)	Peptide 6 grafted to SFTI-1	>50
SFTI-a1	Cyclo[<u>CKSAPPSCAYDGDD</u>]	Peptide 6 grafted to SFTI-1 and based on alanine scanning	0.023
SFTI-DBF		SFTI-a1 replaced prolines with DBF moiety	
Control (SFTI-1)	<u>CTKSIPPICFPDGR</u>	SFTI-1 peptide	>100

were determined by NMR and molecular modeling. Cyclic peptides that were cyclized by backbone cyclization exhibited a beta-hairpin structure, thus reinforcing our hypothesis that the “active core” located in this turn region or β-turn exposes the important residues to the receptor. Molecular modeling studies predicted that the β-turn in the cyclic peptides closely mimics the conformational features of the β-turn in CD2 protein (Giddu et al. 2009; Gokhale et al. 2011; Parajuli et al. 2021).

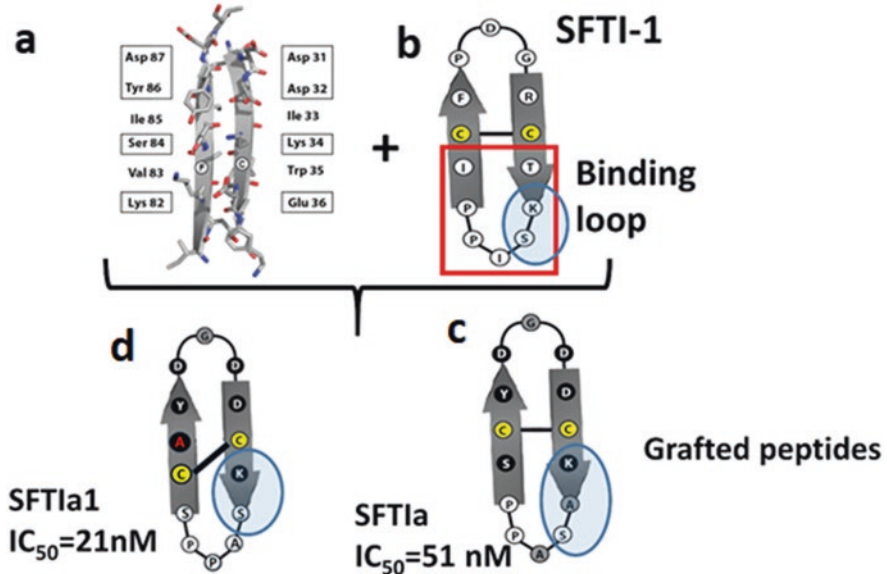


Fig. 1.19 (a) F and C strands of CD2 protein with adhesion domain and residues that are important in binding to CD58 (PDB ID: 1qa9). (b) Schematic diagram of SFTI-1 peptide used for grafting the F and C strands of CD2 and peptide 6. The trypsin binding loop is highlighted. (c) and (d) Grafted peptides SFTIa and SFTIa1. These grafted peptides exhibited cell adhesion inhibition activity with IC₅₀ values of 51 and 21 nM

Third-Generation Design of Grafted Peptides

To improve the stability of this active peptide (peptide **6**), we adopted a novel strategy of multicyclic peptide grafting. In the grafted peptide, amino acids in the loop, disulfide bond, and prolines are important in retaining the stability of the peptide. This peptide was used to graft to the sunflower trypsin inhibitor to improve the stability of the peptide (Fig. 1.19 and Table 1.5). The overall design of grafted peptides starting from peptide **6** is shown in Fig. 1.19. The grafted peptide exhibited cell adhesion inhibition activity of 43 nM and was able to suppress RA in the CIA model. The grafted peptide SFTI-a1 was further studied for its potency and its thermal, chemical, and enzymatic stability. The grafted peptide exhibited improved activity compared to our previous grafted peptide and was stable against thermal, chemical, and enzymatic degradation. Furthermore, the grafted peptide was able to inhibit the PPI of CD2 and CD58 and modulate the CD2-generated signaling for T-cell proliferation (Sable et al. 2016; Parajuli et al. 2021). Further modification of the grafted peptide for oral bioavailability is in progress.

1.11 Summary

Peptide-based therapeutics is gaining momentum, with nearly 10% of the pharmaceutical market and several in the pipeline. In 2019, the peptide therapeutic market was valued at around \$26 billion. It is estimated that peptide-based drugs will have more than \$40 billion/year on the market, and this market is expected to continue to grow substantially over the next 5–10 years. Attractive pharmacological profiles and high affinity make them attractive lead candidates for drug discovery. Although oral delivery of peptides is challenging, numerous technologies that enable the oral delivery of peptide therapeutics are currently in development. Another challenge of peptide therapeutics is the manufacturing cost. While they have lower production costs than protein-based drugs, synthesis of peptides on a small scale for clinical studies can be time-consuming, as large-scale productions are established only after the drug can be shown to be therapeutically efficacious and marketable. The cost of manufacturing peptide drugs was high 10 years ago due to raw materials, solvent requirements, biotechnology production, and limitations on the manufacturing scale of peptides. During the last decade, several improvements were made in peptide synthesis for large-scale production. Raw materials for peptide synthesis became available from different countries at cheaper prices, and manufacturing scales were increased from 10 to 1000 Kg levels. Thus, peptide therapeutics is close to expanding into a large market.

Acknowledgments This research work was supported by funding from the National Cancer Institute of the National Institutes of Health grant 1R01CA255176 to SJ.

References

- Aapro M, Krendyukov A, Schiestl M, Gascon P. Epoetin biosimilars in the treatment of chemotherapy-induced anemia: 10 years' experience gained. *BioDrugs*. 2018;32:129–35.
- Acharya KR, Sturrock ED, Riordan JF, Ehlers MR. Ace revisited: a new target for structure-based drug design. *Nat Rev Drug Discov*. 2003;2:891–902.
- Ainsworth GC, Brown AM, Brownlee G. Aerosporin, an antibiotic produced by *Bacillus aerosporus* Greer. *Nature*. 1947;159:263.
- Albert-Wolf M, Meuer SC, Wallich R. Dual function of recombinant human CD58: inhibition of T cell adhesion and activation via the CD2 pathway. *Int Immunol*. 1991;3:1335–47.
- Alday-Parejo B, Stupp R, Ruegg C. Are integrins still practicable targets for anti-cancer therapy? *Cancers (Basel)*. 2019;11
- Arts EJ, Hazuda DJ. HIV-1 antiretroviral drug therapy. *Cold Spring Harb Perspect Med*. 2012;2:a007161.
- Aumailley M, Gurrath M, Muller G, Calvete J, Timpl R, Kessler H. Arg-Gly-Asp constrained within cyclic pentapeptides. Strong and selective inhibitors of cell adhesion to vitronectin and laminin fragment P1. *FEBS Lett*. 1991;291:50–4.
- Bakdash G, Sittig SP, van Dijk T, Figdor CG, de Vries IJ. The nature of activatory and tolerogenic dendritic cell-derived signal II. *Front Immunol*. 2013;4:53.

- Balanescu A, Radu E, Nat R, Regalia T, Bojinca V, Predescu V, Predeteanu D. Co-stimulatory and adhesion molecules of dendritic cells in rheumatoid arthritis. *J Cell Mol Med.* 2002;6:415–25.
- Ballantine RD, Li YX, Qian PY, Cochrane SA. Rational design of new cyclic analogues of the antimicrobial lipopeptide tridecaptin A1. *Chem Commun.* 2018;54:10634–7.
- Bello L, Francolini M, Marthyn P, Zhang J, Carroll RS, Nikas DC, Strasser JF, Villani R, Cheresh DA, Black PM. Alpha(v)beta3 and alpha(v)beta5 integrin expression in glioma periphery. *Neurosurgery.* 2001;49:380–9. discussion 90
- Biron E, Chatterjee J, Ovadia O, Langenegger D, Brueggen J, Hoyer D, Schmid HA, Jelinek R, Gilon C, Hoffman A, Kessler H. Improving oral bioavailability of peptides by multiple N-methylation: somatostatin analogues. *Angew Chem Int Ed Engl.* 2008;47:2595–9.
- Biswas S, Brunel JM, Dubus JC, Reynaud-Gaubert M, Rolain JM. Colistin: an update on the antibiotic of the 21st century. *Expert Rev Anti-Infect Ther.* 2012;10:917–34.
- Bond A. Exenatide (Byetta) as a novel treatment option for type 2 diabetes mellitus. *Proc (Baylor Univ Med Cent).* 2006;19:281–4.
- Bottger R, Hoffmann R, Knappe D. Differential stability of therapeutic peptides with different proteolytic cleavage sites in blood, plasma and serum. *PLoS One.* 2017;12:e0178943.
- Brand DD, Latham KA, Rosloniec EF. Collagen-induced arthritis. *Nat Protoc.* 2007;2:1269–75.
- Bretscher PA. A two-step, two-signal model for the primary activation of precursor helper T cells. *Proc Natl Acad Sci U S A.* 1999;96:185–90.
- Brodsky IE, Ernst RK, Miller SI, Falkow S. mig-14 is a Salmonella gene that plays a role in bacterial resistance to antimicrobial peptides. *J Bacteriol.* 2002;184:3203–13.
- Brogden KA. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat Rev Microbiol.* 2005;3:238–50.
- Brown P, Dawson MJ. Development of new polymyxin derivatives for multi-drug resistant Gram-negative infections. *J Antibiot (Tokyo).* 2017;70:386–94.
- Brown P, Abbott E, Abdulle O, Boakes S, Coleman S, Divall N, Duperchy E, Moss S, Rivers D, Simonovic M, Singh J, Stanway S, Wilson A, Dawson MJ. Design of next generation polymyxins with lower toxicity: the discovery of SPR206. *ACS Infect Dis.* 2019;5:1645–56.
- Cabrele C, Martinek TA, Reiser O, Berlicki L. Peptides containing beta-amino acid patterns: challenges and successes in medicinal chemistry. *J Med Chem.* 2014;57:9718–39.
- Chakrabarti P, Pal D. The interrelationships of side-chain and main-chain conformations in proteins. *Prog Biophys Mol Biol.* 2001;76:1–102.
- Chambers CA, Allison JP. Co-stimulation in T cell responses. *Curr Opin Immunol.* 1997;9:396–404.
- Chamian F, Lowes MA, Lin SL, Lee E, Kikuchi T, Gilleaudeau P, Sullivan-Whalen M, Cardinale I, Khatcherian A, Novitskaya I, Wittkowski KM, Krueger JG. Alefacept reduces infiltrating T cells, activated dendritic cells, and inflammatory genes in psoriasis vulgaris. *Proc Natl Acad Sci U S A.* 2005;102:2075–80.
- Chatterjee J, Mierke D, Kessler H. N-methylated cyclic pentaalanine peptides as template structures. *J Am Chem Soc.* 2006;128:15164–72.
- Chatterjee B, Saha I, Raghothama S, Aravinda S, Rai R, Shamala N, Balam P. Designed peptides with homochiral and heterochiral diproline templates as conformational constraints. *Chemistry.* 2008a;14:6192–204.
- Chatterjee J, Gilon C, Hoffman A, Kessler H. N-methylation of peptides: a new perspective in medicinal chemistry. *Acc Chem Res.* 2008b;41:1331–42.
- Chen L, Flies DB. Molecular mechanisms of T cell co-stimulation and co-inhibition. *Nat Rev Immunol.* 2013;13:227–42.
- Chen CH, Lu TK. Development and challenges of antimicrobial peptides for therapeutic applications. *Antibiotics (Basel).* 2020;9
- Chinot OL. Cilengitide in glioblastoma: when did it fail? *Lancet Oncol.* 2014;15:1044–5.
- Corradi HR, Schwager SL, Nchinda AT, Sturrock ED, Acharya KR. Crystal structure of the N domain of human somatic angiotensin I-converting enzyme provides a structural basis for domain-specific inhibitor design. *J Mol Biol.* 2006;357:964–74.
- Craik DJ. *NMR in drug design.* CRC Press; 1996.

- Crecente-Garcia S, Neckebroek A, Clark JS, Smith BO, Thomson AR. beta-turn mimics by chemical ligation. *Org Lett*. 2020;22:4424–8.
- Cudic P, Stawikowski M. Pseudopeptide synthesis via Fmoc solid-phase synthetic methodology. *Mini Rev Organ Chem*. 2007;4:268–80.
- Cunningham AD, Qvit N, Mochly-Rosen D. Peptides and peptidomimetics as regulators of protein-protein interactions. *Curr Opin Struct Biol*. 2017;44:59–66.
- Davis SJ, Ikemizu S, Evans EJ, Fugger L, Bakker TR, van der Merwe PA. The nature of molecular recognition by T cells. *Nat Immunol*. 2003;4:217–24.
- Davis JM, Tsou LK, Hamilton AD. Synthetic non-peptide mimetics of alpha-helices. *Chem Soc Rev*. 2007;36:326–34.
- de Veer SJ, Kan MW, Craik DJ. Cyclotides: from structure to function. *Chem Rev*. 2019;119:12375–421.
- de Visser PC, Kriek NM, van Hoof PA, Van Schepdael A, Filippov DV, van der Marel GA, Overkleef HS, van Boom JH, Noort D. Solid-phase synthesis of polymyxin B1 and analogues via a safety-catch approach. *J Pept Res*. 2003;61:298–306.
- Dechantsreiter MA, Planker E, Matha B, Lohof E, Holzemann G, Jonczyk A, Goodman SL, Kessler H. N-Methylated cyclic RGD peptides as highly active and selective alpha(V)beta(3) integrin antagonists. *J Med Chem*. 1999;42:3033–40.
- Deike S, Rothmund S, Voigt B, Samantray S, Strodel B, Binder WH. beta-Turn mimetic synthetic peptides as amyloid-beta aggregation inhibitors. *Bioorg Chem*. 2020;101:104012.
- Del Borgo MP, Kulkarni K, Aguilar MI. Using β -amino acids and β -peptide templates to create bioactive ligands and biomaterials. *Curr Pharm Des*. 2017;23:3772–85.
- Deris ZZ, Akter J, Sivanesan S, Roberts KD, Thompson PE, Nation RL, Li J, Velkov T. A secondary mode of action of polymyxins against Gram-negative bacteria involves the inhibition of NADH-quinone oxidoreductase activity. *J Antibiot (Tokyo)*. 2014;67:147–51.
- Desgrosellier JS, Cheresch DA. Integrins in cancer: biological implications and therapeutic opportunities. *Nat Rev Cancer*. 2010;10:9–22.
- Dishon S, Schumacher A, Fanous J, Talhami A, Kassis I, Karussis D, Gilon C, Hoffman A, Nussbaum G. Development of a novel backbone cyclic peptide inhibitor of the innate immune TLR/IL1R signaling protein MyD88. *Sci Rep*. 2018;8:9476.
- Dishon S, Schumacher-Klinger A, Gilon C, Hoffman A, Nussbaum G. Myristoylation confers oral bioavailability and improves the bioactivity of c(MyD 4-4), a cyclic peptide inhibitor of MyD88. *Mol Pharm*. 2019;16:1516–22.
- Dixon RA, Chopra I. Leakage of periplasmic proteins from *Escherichia coli* mediated by polymyxin B nonapeptide. *Antimicrob Agents Chemother*. 1986;29:781–8.
- Drugs for type 2 diabetes. *Med Lett Drugs Ther*. 2019;61:169–78.
- Dyson HJ, Rance M, Houghten RA, Lerner RA, Wright PE. Folding of immunogenic peptide fragments of proteins in water solution. I. Sequence requirements for the formation of a reverse turn. *J Mol Biol*. 1988;201:161–200.
- Eckhardt B, Grosse W, Essen LO, Geyer A. Structural characterization of a beta-turn mimic within a protein-protein interface. *Proc Natl Acad Sci U S A*. 2010;107:18336–41.
- Eilken HM, Adams RH. Turning on the angiogenic microswitch. *Nat Med*. 2010;16:853–4.
- Espagnolle N, Depoil D, Zaru R, Demeur C, Champagne E, Guiraud M, Valitutti S. CD2 and TCR synergize for the activation of phospholipase Cgamma1/calcium pathway at the immunological synapse. *Int Immunol*. 2007;19:239–48.
- Fang Y. Label-free receptor assays. *Drug Discov Today Technol*. 2011;7:e5–e11.
- Fischer G, Rossmann M, Hyvonen M. Alternative modulation of protein-protein interactions by small molecules. *Curr Opin Biotechnol*. 2015;35:78–85.
- Fisher BF, Hong SH, Gellman SH. Helix propensities of amino acid residues via thioester exchange. *J Am Chem Soc*. 2017;139:13292–5.
- Fjell CD, Hiss JA, Hancock RE, Schneider G. Designing antimicrobial peptides: form follows function. *Nat Rev Drug Discov*. 2011;11:37–51.

- Fosgerau K, Hoffmann T. Peptide therapeutics: current status and future directions. *Drug Discov Today*. 2015;20:122–8.
- Foye WO. Foye's principles of medicinal chemistry. Lippincott Williams & Wilkins; 2008.
- Frey V, Viaud J, Subra G, Cauquil N, Guichou JF, Casara P, Grassy G, Chavanieu A. Structure-activity relationships of Bak derived peptides: affinity and specificity modulations by amino acid replacement. *Eur J Med Chem*. 2008;43:966–72.
- Fuller AA, Du D, Liu F, Davoren JE, Bhabha G, Kroon G, Case DA, Dyson HJ, Powers ET, Wipf P, Gruebele M, Kelly JW. Evaluating beta-turn mimics as beta-sheet folding nucleators. *Proc Natl Acad Sci U S A*. 2009;106:11067–72.
- Geyer A, Mueller G, Kessler H. Conformational analysis of a cyclic RGD peptide containing a psi.[CH2-NH] bond: a positional shift in backbone structure caused by a single dipeptide mimetic. *J Am Chem Soc*. 1994;116:7735–43.
- Gibbs AC, Bjorn Dahl TC, Hodges RS, Wishart DS. Probing the structural determinants of type II' beta-turn formation in peptides and proteins. *J Am Chem Soc*. 2002;124:1203–13.
- Giddu S, Subramanian V, Yoon HS, Satyanarayanajois SD. Design of beta-hairpin peptides for modulation of cell adhesion by beta-turn constraint. *J Med Chem*. 2009;52:726–36.
- Gilon C, Halle D, Chorev M, Selinger Z, Byk G. Backbone cyclization: a new method for conferring conformational constraint on peptides. *Biopolymers*. 1991;31:745–50.
- Gokhale A, Weldegiorghis TK, Taneja V, Satyanarayanajois SD. Conformationally constrained peptides from CD2 to modulate protein-protein interactions between CD2 and CD58. *J Med Chem*. 2011;54:5307–19.
- Gokhale A, Kanthala S, Latendresse J, Taneja V, Satyanarayanajois S. Immunosuppression by co-stimulatory molecules: inhibition of CD2-CD48/CD58 interaction by peptides from CD2 to suppress progression of collagen-induced arthritis in mice. *Chem Biol Drug Des*. 2013;82:106–18.
- Gokhale AS, Sable R, Walker JD, McLaughlin L, Kousoulas KG, Jois SD. Inhibition of cell adhesion and immune responses in the mouse model of collagen-induced arthritis with a peptidomimetic that blocks CD2-CD58 interface interactions. *Biopolymers*. 2015;
- Hancock RE, Sahl HG. Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. *Nat Biotechnol*. 2006;24:1551–7.
- Haubner R, Schmitt W, Hölzemann G, Goodman SL, Jonczyk A, Kessler H. Cyclic RGD peptides containing β -turn mimetics. *J Am Chem Soc*. 1996;118:7881–91.
- Hauser AS, Attwood MM, Rask-Andersen M, Schiøth HB, Gloriam DE. Trends in GPCR drug discovery: new agents, targets and indications. *Nat Rev Drug Discov*. 2017;16:829–42.
- Helander IM, Kilpeläinen I, Vaara M. Increased substitution of phosphate groups in lipopolysaccharides and lipid A of the polymyxin-resistant pmrA mutants of *Salmonella typhimurium*: a 31P-NMR study. *Mol Microbiol*. 1994;11:481–7.
- Henninot A, Collins JC, Nuss JM. The current state of peptide drug discovery: back to the future? *J Med Chem*. 2018;61:1382–414.
- Hood JD, Cheresh DA. Role of integrins in cell invasion and migration. *Nat Rev Cancer*. 2002;2:91–100.
- Hruby VJ. Conformational and topographical considerations in the design of biologically active peptides. *Biopolymers*. 1993;33:1073–82.
- Hruby VJ. Designing peptide receptor agonists and antagonists. *Nat Rev Drug Discov*. 2002;1:847–58.
- Hruby VJ, Cai M. Design of peptide and peptidomimetic ligands with novel pharmacological activity profiles. *Annu Rev Pharmacol Toxicol*. 2013;53:557–80.
- Huang Z, Li S, Korngold R. Immunoglobulin superfamily proteins: structure, mechanisms, and drug discovery. *Biopolymers*. 1997;43:367–82.
- Humphries MJ. The molecular basis and specificity of integrin-ligand interactions. *J Cell Sci*. 1990;97(Pt 4):585–92.
- Hurevich M, Ratner-Hurevich M, Tal-Gan Y, Shalev DE, Ben-Sasson SZ, Gilon C. Backbone cyclic helix mimetic of chemokine (C-C motif) receptor 2: a rational approach for inhibiting dimerization of G protein-coupled receptors. *Bioorg Med Chem*. 2013;21:3958–66.

- Hutchinson EG, Thornton JM. A revised set of potentials for beta-turn formation in proteins. *Protein Sci.* 1994;3:2207–16.
- Jayatunga MK, Thompson S, Hamilton AD. alpha-Helix mimetics: outwards and upwards. *Bioorg Med Chem Lett.* 2014;24:717–24.
- Jo H, Meinhardt N, Wu Y, Kulkarni S, Hu X, Low KE, Davies PL, DeGrado WF, Greenbaum DC. Development of alpha-helical calpain probes by mimicking a natural protein-protein interaction. *J Am Chem Soc.* 2012;134:17704–13.
- Jois SD, Jining L, Nagarajarao LM. Targeting T-cell adhesion molecules for drug design. *Curr Pharm Des.* 2006;12:2797–812.
- Jones MC. Therapies for diabetes: pramlintide and exenatide. *Am Fam Physician.* 2007;75:1831–5.
- Kaizuka Y, Douglass AD, Vardhana S, Dustin ML, Vale RD. The coreceptor CD2 uses plasma membrane microdomains to transduce signals in T cells. *J Cell Biol.* 2009;185:521–34.
- Kamradt T, Mitchison NA. Tolerance and autoimmunity. *N Engl J Med.* 2001;344:655–64.
- Kanazawa K, Sato Y, Ohki K, Okimura K, Uchida Y, Shindo M, Sakura N. Contribution of each amino acid residue in polymyxin B(3) to antimicrobial and lipopolysaccharide binding activity. *Chem Pharm Bull(Tokyo).* 2009;57:240–4.
- Karle IL, Flippen-Anderson JL, Uma K, Balaram H, Balaram P. Alpha-helix and mixed 3(10)/alpha-helix in cocrystallized conformers of Boc-Aib-Val-Aib-Aib-Val-Val-Val-Aib-Val-Aib--OMe. *Proc Natl Acad Sci U S A.* 1989;86:765–9.
- Kaspar AA, Reichert JM. Future directions for peptide therapeutics development. *Drug Discov Today.* 2013;18:807–17.
- Kerr JS, Slee AM, Mousa SA. The alpha v integrin antagonists as novel anticancer agents: an update. *Expert Opin Investig Drugs.* 2002;11:1765–74.
- Khakshoor O, Nowick JS. Artificial beta-sheets: chemical models of beta-sheets. *Curr Opin Chem Biol.* 2008;12:722–9.
- Kim M, Sun ZY, Byron O, Campbell G, Wagner G, Wang J, Reinherz EL. Molecular dissection of the CD2-CD58 counter-receptor interface identifies CD2 Tyr86 and CD58 Lys34 residues as the functional “hot spot”. *J Mol Biol.* 2001;312:711–20.
- Knudsen KA, Horwitz AF, Buck CA. A monoclonal antibody identifies a glycoprotein complex involved in cell-substratum adhesion. *Exp Cell Res.* 1985;157:218–26.
- Koch-Weser J, Sidel VW, Federman EB, Kanarek P, Finer DC, Eaton AE. Adverse effects of sodium colistimethate. Manifestations and specific reaction rates during 317 courses of therapy. *Ann Intern Med.* 1970;72:857–68.
- Krauson AJ, He J, Wimley WC. Gain-of-function analogues of the pore-forming peptide melittin selected by orthogonal high-throughput screening. *J Am Chem Soc.* 2012;134:12732–41.
- Kudo F, Miyanaga A, Eguchi T. Biosynthesis of natural products containing beta-amino acids. *Nat Prod Rep.* 2014;31:1056–73.
- Kumar P, Nagarajan A, Uchil PD. Analysis of cell viability by the MTT assay. *Cold Spring Harb Protoc.* 2018;2018
- Lai Y, Gallo RL. AMPed up immunity: how antimicrobial peptides have multiple roles in immune defense. *Trends Immunol.* 2009;30:131–41.
- Lanning ME, Fletcher S. Multi-facial, non-peptidic alpha-helix mimetics. *Biology (Basel).* 2015;4:540–55.
- Lao BB, Drew K, Guarracino DA, Brewer TF, Heindel DW, Bonneau R, Arora PS. Rational design of topographical helix mimics as potent inhibitors of protein-protein interactions. *J Am Chem Soc.* 2014;136:7877–88.
- Lau JL, Dunn MK. Therapeutic peptides: historical perspectives, current development trends, and future directions. *Bioorg Med Chem.* 2018;26:2700–7.
- Laxio Arenas J, Kaffy J, Onger S. Peptides and peptidomimetics as inhibitors of protein-protein interactions involving beta-sheet secondary structures. *Curr Opin Chem Biol.* 2019;52:157–67.
- Lemke TL, Williams DA. Foye’s principles of medicinal chemistry. Wolters Kluwer Health/Lippincott Williams & Wilkins; 2012.
- Lenci E, Trabocchi A. Peptidomimetic toolbox for drug discovery. *Chem Soc Rev.* 2020;49:3262–77.

- Li S, Schoneich C, Borchardt RT. Chemical instability of protein pharmaceuticals: Mechanisms of oxidation and strategies for stabilization. *Biotechnol Bioeng.* 1995;48:490–500.
- Li J, Nation RL, Turnidge JD, Milne RW, Coulthard K, Rayner CR, Paterson DL. Colistin: the re-emerging antibiotic for multidrug-resistant Gram-negative bacterial infections. *Lancet Infect Dis.* 2006;6:589–601.
- Li X, Chen S, Zhang WD, Hu HG. Stapled helical peptides bearing different anchoring residues. *Chem Rev.* 2020;120:10079–144.
- Lindsay R, Kregge JH, Marin F, Jin L, Stepan JJ. Teriparatide for osteoporosis: importance of the full course. *Osteoporos Int.* 2016;27:2395–410.
- Ling S, Liu Y, Fu J, Colletta A, Gilon C, Holoshitz J. Shared epitope-antagonistic ligands: a new therapeutic strategy in mice with erosive arthritis. *Arthritis Rheumatol.* 2015;67:2061–70.
- Liu J, Chow VT, Jois SD. A novel, rapid and sensitive heterotypic cell adhesion assay for CD2-CD58 interaction, and its application for testing inhibitory peptides. *J Immunol Methods.* 2004;291:39–49.
- Liu J, Ying J, Chow VT, Hraby VJ, Satyanarayanajois SD. Structure-activity studies of peptides from the “hot-spot” region of human CD2 protein: development of peptides for immunomodulation. *J Med Chem.* 2005;48:6236–49.
- Liu J, Li C, Ke S, Satyanarayanajois SD. Structure-based rational design of beta-hairpin peptides from discontinuous epitopes of cluster of differentiation 2 (CD2) protein to modulate cell adhesion interaction. *J Med Chem.* 2007;50:4038–47.
- Lohof E, Planker E, Mang C, Burkhardt F, Dechantsreiter MA, Haubner R, Wester HJ, Schwaiger M, Holzemann G, Goodman SL, Kessler H. Carbohydrate derivatives for use in drug design: cyclic alpha(v)-selective RGD peptides this work was supported by the Fonds der Chemischen Industrie, the Deutsche Forschungsgemeinschaft, and the Sanderstiftung. The authors thank M. Urzinger, B. Cordes, M. Kranawetter, M. Wolff, and A. Zeller for technical assistance. *Angew Chem Int Ed Engl.* 2000;39:2761–4.
- London N, Raveh B, Schueler-Furman O. Druggable protein-protein interactions – from hot spots to hot segments. *Curr Opin Chem Biol.* 2013;17:952–9.
- Lopez-Perez PM, Grimsey E, Bourne L, Mikut R, Hilpert K. Screening and optimizing antimicrobial peptides by using SPOT-synthesis. *Front Chem.* 2017;5:25.
- Loughlin WA, Tyndall JD, Glenn MP, Hill TA, Fairlie DP. Update 1 of: beta-strand mimetics. *Chem Rev.* 2010;110:PR32–69.
- Love BL, Johnson A, Smith LS. Linaclotide: a novel agent for chronic constipation and irritable bowel syndrome. *Am J Health Syst Pharm.* 2014;71:1081–91.
- Mabonga L, Kappo AP. Peptidomimetics: a synthetic tool for inhibiting protein–protein interactions in cancer. *Int J Pept Res Ther.* 2020;26:225–41.
- MacDonald TJ, Taga T, Shimada H, Tabrizi P, Zlokovic BV, Cheresch DA, Laug WE. Preferential susceptibility of brain tumors to the antiangiogenic effects of an alpha(v) integrin antagonist. *Neurosurgery.* 2001;48:151–7.
- Martin JL, Begun J, Schindeler A, Wickramasinghe WA, Alewood D, Alewood PF, Bergman DA, Brinkworth RI, Abbenante G, March DR, Reid RC, Fairlie DP. Molecular recognition of macrocyclic peptidomimetic inhibitors by HIV-1 protease. *Biochemistry.* 1999;38:7978–88.
- Mas-Moruno C, Rechenmacher F, Kessler H. Cilengitide: the first anti-angiogenic small molecule drug candidate design, synthesis and clinical evaluation. *Anti Cancer Agents Med Chem.* 2010;10:753–68.
- Masuyer G, Schwager SL, Sturrock ED, Isaac RE, Acharya KR. Molecular recognition and regulation of human angiotensin-I converting enzyme (ACE) activity by natural inhibitory peptides. *Sci Rep.* 2012;2:717.
- Mathur J, Waldor MK. The *Vibrio cholerae* ToxR-regulated porin OmpU confers resistance to antimicrobial peptides. *Infect Immun.* 2004;72:3577–83.
- Mavromoustakos T, Durdagi S, Koukoulitsa C, Simcic M, Papadopoulos MG, Hodoscek M, Grdadolnik SG. Strategies in the rational drug design. *Curr Med Chem.* 2011;18:2517–30.
- McGregor MJ, Islam SA, Sternberg MJ. Analysis of the relationship between side-chain conformation and secondary structure in globular proteins. *J Mol Biol.* 1987;198:295–310.

- Miranda LP, Holder JR, Shi L, Bennett B, Aral J, Gegg CV, Wright M, Walker K, Doellgast G, Rogers R, Li H, Valladares V, Salyers K, Johnson E, Wild K. Identification of potent, selective, and metabolically stable peptide antagonists to the calcitonin gene-related peptide (CGRP) receptor. *J Med Chem.* 2008;51:7889–97.
- Miyoshi M, Liu S. Collagen-induced arthritis models. *Methods Mol Biol.* 2018;1868:3–7.
- Modell AE, Blosser SL, Arora PS. Systematic targeting of protein-protein interactions. *Trends Pharmacol Sci.* 2016;37:702–13.
- Morrison KL, Weiss GA. Combinatorial alanine-scanning. *Curr Opin Chem Biol.* 2001;5:302–7.
- Mrowietz U. Treatment targeted to cell surface epitopes. *Clin Exp Dermatol.* 2002;27:591–6.
- Muller G, Gurrath M, Kessler H. Pharmacophore refinement of gpIIb/IIIa antagonists based on comparative studies of antiadhesive cyclic and acyclic RGD peptides. *J Comput Aided Mol Des.* 1994;8:709–30.
- Muttenthaler M, King GF, Adams DJ, Alewood PF. Trends in peptide drug discovery. *Nat Rev Drug Discov.* 2021;20:309–25.
- Nguyen LT, Haney EF, Vogel HJ. The expanding scope of antimicrobial peptide structures and their modes of action. *Trends Biotechnol.* 2011;29:464–72.
- Nowick JS. Exploring beta-sheet structure and interactions with chemical model systems. *Acc Chem Res.* 2008;41:1319–30.
- Okimura K, Ohki K, Sato Y, Ohnishi K, Sakura N. Semi-synthesis of polymyxin B (2-10) and colistin (2-10) analogs employing the Trichloroethoxycarbonyl (Troc) group for side chain protection of alpha, gamma-diaminobutyric acid residues. *Chem Pharm Bull(Tokyo).* 2007;55:1724–30.
- Ondetti MA, Williams NJ, Sabo EF, Pluscec J, Weaver ER, Kocy O. Angiotensin-converting enzyme inhibitors from the venom of *Bothrops jararaca*. Isolation, elucidation of structure, and synthesis. *Biochemistry.* 1971;10:4033–9.
- Orwa JA, Govaerts C, Busson R, Roets E, Van Schepdael A, Hoogmartens J. Isolation and structural characterization of colistin components. *J Antibiot (Tokyo).* 2001a;54:595–9.
- Orwa JA, Govaerts C, Busson R, Roets E, Van Schepdael A, Hoogmartens J. Isolation and structural characterization of polymyxin B components. *J Chromatogr A.* 2001b;912:369–73.
- Padhi A, Sengupta M, Sengupta S, Roehm KH, Sonawane A. Antimicrobial peptides and proteins in mycobacterial therapy: current status and future prospects. *Tuberculosis (Edinb).* 2014;94:363–73.
- Paolillo M, Russo MA, Serra M, Colombo L, Schinelli S. Small molecule integrin antagonists in cancer therapy. *Mini Rev Med Chem.* 2009;9:1439–46.
- Parajuli P, Sable R, Shrestha L, Dahal A, Gauthier T, Taneja V, Jois S. Modulation of co-stimulatory signal from CD2-CD58 proteins by a grafted peptide. *Chem Biol Drug Des.* 2021;97:607–27.
- Pelay-Gimeno M, Glas A, Koch O, Grossmann TN. Structure-based design of inhibitors of protein-protein interactions: mimicking peptide binding epitopes. *Angew Chem Int Ed Engl.* 2015;54:8896–927.
- Pierschbacher MD, Ruoslahti E. Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature.* 1984a;309:30–3.
- Pierschbacher MD, Ruoslahti E. Variants of the cell recognition site of fibronectin that retain attachment-promoting activity. *Proc Natl Acad Sci U S A.* 1984b;81:5985–8.
- Pierschbacher MD, Ruoslahti E. Influence of stereochemistry of the sequence Arg-Gly-Asp-Xaa on binding specificity in cell adhesion. *J Biol Chem.* 1987;262:17294–8.
- Pierschbacher MD, Hayman EG, Ruoslahti E. Location of the cell-attachment site in fibronectin with monoclonal antibodies and proteolytic fragments of the molecule. *Cell.* 1981;26:259–67.
- Pierschbacher M, Hayman EG, Ruoslahti E. Synthetic peptide with cell attachment activity of fibronectin. *Proc Natl Acad Sci U S A.* 1983;80:1224–7.
- Poirel L, Jayol A, Nordmann P. Polymyxins: antibacterial activity, susceptibility testing, and resistance mechanisms encoded by plasmids or chromosomes. *Clin Microbiol Rev.* 2017;30:557–96.
- Pope JE, Deer TR. Ziconotide: a clinical update and pharmacologic review. *Expert Opin Pharmacother.* 2013;14:957–66.

- Porto WF, Irazazabal L, Alves ESF, Ribeiro SM, Matos CO, Pires AS, Fensterseifer ICM, Miranda VJ, Haney EF, Humblot V, Torres MDT, Hancock REW, Liao LM, Ladram A, Lu TK, de la Fuente-Nunez C, Franco OL. In silico optimization of a guava antimicrobial peptide enables combinatorial exploration for peptide design. *Nat Commun*. 2018;9:1490.
- Qiu Y, Taichi M, Wei N, Yang H, Luo KQ, Tam JP. An orally active bradykinin B1 receptor antagonist engineered as a bifunctional chimera of sunflower trypsin inhibitor. *J Med Chem*. 2017;60:504–10.
- Rader AFB, Weinmuller M, Reichart F, Schumacher-Klinger A, Merzbach S, Gilon C, Hoffman A, Kessler H. Orally active peptides: is there a magic bullet? *Angew Chem Int Ed Engl*. 2018;57:14414–38.
- Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. *J Mol Biol*. 1963;7:95–9.
- Reardon DA, Neyns B, Weller M, Tonn JC, Nabors LB, Stupp R. Cilengitide: an RGD pentapeptide alphanubeta3 and alphanubeta5 integrin inhibitor in development for glioblastoma and other malignancies. *Future Oncol*. 2011;7:339–54.
- Rognan D. Rational design of protein–protein interaction inhibitors. *Med Chem Commun*. 2015;6:51–60.
- Ross NT, Katt WP, Hamilton AD. Synthetic mimetics of protein secondary structure domains. *Philos Trans A Math Phys Eng Sci*. 2010;368:989–1008.
- Rotondi KS, Gierasch LM. Natural polypeptide scaffolds: beta-sheets, beta-turns, and beta-hairpins. *Biopolymers*. 2006;84:13–22.
- Rubin SJS, Tal-Gan Y, Gilon C, Qvit N. Conversion of protein active regions into peptidomimetic therapeutic leads using backbone cyclization and cycloscan – how to do it yourself! *Curr Top Med Chem*. 2018;18:556–65.
- Ruoslahti E, Pierschbacher MD. New perspectives in cell adhesion: RGD and integrins. *Science*. 1987;238:491–7.
- Rustin MH. Long-term safety of biologics in the treatment of moderate-to-severe plaque psoriasis: review of current data. *Br J Dermatol*. 2012;167(Suppl 3):3–11.
- Ryan DP, Matthews JM. Protein-protein interactions in human disease. *Curr Opin Struct Biol*. 2005;15:441–6.
- Saab F, Ionescu C, Schweiger MJ. Bleeding risk and safety profile related to the use of eptifibatide: a current review. *Expert Opin Drug Saf*. 2012;11:315–24.
- Sable R, Durek T, Taneja V, Craik DJ, Pallerla S, Gauthier T, Jois S. Constrained cyclic peptides as immunomodulatory inhibitors of the CD2:CD58 protein-protein interaction. *ACS Chem Biol*. 2016;11:2366–74.
- Sable R, Parajuli P, Jois S. Peptides, peptidomimetics, and polypeptides from marine sources: a wealth of natural sources for pharmaceutical applications. *Mar Drugs*. 2017;15
- Satyanarayanajois SD, Buyuktimkin B, Gokhale A, Ronald S, Siahaan TJ, Latendresse JR. A peptide from the beta-strand region of CD2 protein that inhibits cell adhesion and suppresses arthritis in a mouse model. *Chem Biol Drug Des*. 2010;76:234–44.
- Satyanarayanajois SD, Ronald S, Liu J. Heterotypic cell adhesion assay for the study of cell adhesion inhibition. *Methods Mol Biol*. 2011;716:225–43.
- Seebach D, Gardiner J. Beta-peptidic peptidomimetics. *Acc Chem Res*. 2008;41:1366–75.
- Shapovalov M, Vucetic S, Dunbrack RL Jr. A new clustering and nomenclature for beta turns derived from high-resolution protein structures. *PLoS Comput Biol*. 2019;15:e1006844.
- Sharma SK, Wu AD, Chandramouli N, Fotsch C, Kardash G, Bair KW. Solid-phase total synthesis of polymyxin B1. *J Pept Res*. 1999;53:501–6.
- Strømgaard K, Krosgaard-Larsen P, Madsen U. Textbook of drug design and discovery. CRC Press; 2017.
- Stumpf MP, Thorne T, de Silva E, Stewart R, An HJ, Lappe M, Wiuf C. Estimating the size of the human interactome. *Proc Natl Acad Sci U S A*. 2008;105:6959–64.
- Stupp R, Hegi ME, Gorlia T, Erridge SC, Perry J, Hong YK, Aldape KD, Lhermitte B, Pietsch T, Grujcic D, Steinbach JP, Wick W, Tarnawski R, Nam DH, Hau P, Weyerbrock A, Taphoorn

- MJ, Shen CC, Rao N, Thurzo L, Herrlinger U, Gupta T, Kortmann RD, Adamska K, McBain C, Brandes AA, Tonn JC, Schnell O, Wiegel T, Kim CY, Nabors LB, Reardon DA, van den Bent MJ, Hicking C, Markivskyy A, Picard M, Weller M, Research European Organisation for, Cancer Treatment of, Consortium Canadian Brain Tumor, and Centric study team. Cilengitide combined with standard treatment for patients with newly diagnosed glioblastoma with methylated MGMT promoter (CENTRIC EORTC 26071-22072 study): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol.* 2014;15:1100–8.
- Sturrock ED, Natesh R, van Rooyen JM, Acharya KR. Structure of angiotensin I-converting enzyme. *Cell Mol Life Sci.* 2004;61:2677–86.
- Talhami A, Swed A, Hess S, Ovadia O, Greenberg S, Schumacher-Klinger A, Rosenthal D, Shalev DE, Hurevich M, Lazarovici P, Hoffman A, Gilon C. Cyclizing painkillers: development of backbone-cyclic TAPS analogs. *Front Chem.* 2020;8:532577.
- Tan YS, Lane DP, Verma CS. Stapled peptide design: principles and roles of computation. *Drug Discov Today.* 2016;21:1642–53.
- Thompson AD, Dugan A, Gestwicki JE, Mapp AK. Fine-tuning multiprotein complexes using small molecules. *ACS Chem Biol.* 2012;7:1311–20.
- Tolliday N. High-throughput assessment of Mammalian cell viability by determination of adenosine triphosphate levels. *Curr Protoc Chem Biol.* 2010;2:153–61.
- Torres MDT, Sothiselvam S, Lu TK, de la Fuente-Nunez C. Peptide design principles for antimicrobial applications. *J Mol Biol.* 2019;431:3547–67.
- Tranqui L, Andrieux A, Hudry-Clergeon G, Ryckewaert JJ, Soye S, Chapel A, Ginsberg MH, Plow EF, Marguerie G. Differential structural requirements for fibrinogen binding to platelets and to endothelial cells. *J Cell Biol.* 1989;108:2519–27.
- Tsubery H, Ofek I, Cohen S, Fridkin M. Structure activity relationship study of polymyxin B nonapeptide. *Adv Exp Med Biol.* 2000;479:219–22.
- Uhlig T, Kyprianou T, Martinelli FG, Oppici CA, Heiligers D, Hills D, Calvo XR, Verhaert P. The emergence of peptides in the pharmaceutical business: From exploration to exploitation. *EuPA Open Proteom.* 2014;4:58–69.
- Usmani SS, Bedi G, Samuel JS, Singh S, Kalra S, Kumar P, Ahuja AA, Sharma M, Gautam A, Raghava GPS. THPdb: database of FDA-approved peptide and protein therapeutics. *PLoS One.* 2017;12:e0181748.
- Vaara M. Novel derivatives of polymyxins. *J Antimicrob Chemother.* 2013;68:1213–9.
- Vaara M, Fox J, Loidl G, Siikanen O, Apajalahti J, Hansen F, Frimodt-Moller N, Nagai J, Takano M, Vaara T. Novel polymyxin derivatives carrying only three positive charges are effective antibacterial agents. *Antimicrob Agents Chemother.* 2008;52:3229–36.
- Vagner J, Qu H, Hruby VJ. Peptidomimetics, a synthetic tool of drug discovery. *Curr Opin Chem Biol.* 2008;12:292–6.
- van der Merwe PA, Davis SJ. Molecular interactions mediating T cell antigen recognition. *Annu Rev Immunol.* 2003;21:659–84.
- Velikovskiy CA, Deng L, Chlewicki LK, Fernandez MM, Kumar V, Mariuzza RA. Structure of natural killer receptor 2B4 bound to CD48 reveals basis for heterophilic recognition in signaling lymphocyte activation molecule family. *Immunity.* 2007;27:572–84.
- Velkov T, Thompson PE, Nation RL, Li J. Structure – activity relationships of polymyxin antibiotics. *J Med Chem.* 2010;53:1898–916.
- Vlieghe P, Lisowski V, Martinez J, Khrestchatsky M. Synthetic therapeutic peptides: science and market. *Drug Discov Today.* 2010;15:40–56.
- Vogler K, Studer RO, Lanz P, Lergier W, Bohni E. Syntheses in the polymyxin series. 9. Synthesis of polymyxin B-1. *Helv Chim Acta.* 1965;48:1161–77.
- Wang JH, Smolyar A, Tan K, Liu JH, Kim M, Sun ZY, Wagner G, Reinherz EL. Structure of a heterophilic adhesion complex between the human CD2 and CD58 (LFA-3) counterreceptors. *Cell.* 1999;97:791–803.
- Wang G, Li X, Wang Z. APD2: the updated antimicrobial peptide database and its application in peptide design. *Nucleic Acids Res.* 2009;37:D933–7.

- Weis WI, Kobilka BK. The molecular basis of G protein-coupled receptor activation. *Annu Rev Biochem.* 2018;87:897–919.
- Wermuth J, Goodman SL, Jonczyk A, Kessler H. Stereoisomerism and biological activity of the selective and superactive $\alpha\beta 3$ integrin inhibitor cyclo (-RGDfV-) and its retro-inverso peptide. *J Am Chem Soc.* 1997;119:1328–35.
- Wilmot CM, Thornton JM. Beta-turns and their distortions: a proposed new nomenclature. *Protein Eng.* 1990;3:479–93.
- Wilson AC, Meethal SV, Bowen RL, Atwood CS. Leuprolide acetate: a drug of diverse clinical applications. *Expert Opin Investig Drugs.* 2007;16:1851–63.
- Wisniewski K, Galyean R, Tariga H, Alagarsamy S, Croston G, Heitzmann J, Kohan A, Wisniewska H, Laporte R, Riviere PJ, Scheingart CD. New, potent, selective, and short-acting peptidic V1a receptor agonists. *J Med Chem.* 2011;54:4388–98.
- Wu YD, Gellman S. Peptidomimetics. *Acc Chem Res.* 2008;41:1231–2.
- Xiong JP, Stehle T, Zhang R, Joachimiak A, Frech M, Goodman SL, Arnaout MA. Crystal structure of the extracellular segment of integrin $\alpha V\beta 3$ in complex with an Arg-Gly-Asp ligand. *Science.* 2002;296:151–5.
- Yahav D, Farbman L, Leibovici L, Paul M. Colistin: new lessons on an old antibiotic. *Clin Microbiol Infect.* 2012;18:18–29.
- Zhao L, Chmielewski J. Inhibiting protein-protein interactions using designed molecules. *Curr Opin Struct Biol.* 2005;15:31–4.
- Zheng Y, Leftheris K. Insights into protein-ligand interactions in integrin complexes: advances in structure determinations. *J Med Chem.* 2020;63:5675–96.

Chapter 2

Peptide Synthesis: Methods and Protocols



Ted Gauthier and Dong Liu

Contents

2.1	Introduction.....	52
2.2	Fundamentals of Peptide Synthesis.....	52
2.2.1	Solution-Phase Peptide Synthesis.....	53
2.2.2	Solid-Phase Peptide Synthesis.....	54
2.3	Practical Aspects of Peptide Synthesis.....	57
2.4	Peptide Modifications.....	61
2.4.1	Cyclization.....	63
2.4.2	Disulfide Bond Formation.....	63
2.4.3	Automation of Solid-Phase Peptide Synthesis.....	64
2.4.4	Peptide Analysis and Purification.....	65
2.4.5	Sustainability in Peptide Synthesis.....	65
2.5	Examples of Peptide Syntheses Performed in the Authors' Laboratory.....	66
2.5.1	Linear Peptide EHWSY-dK-LRPG-NH ₂	66
2.5.2	Fluorescently Labeled Peptide Ac-RWVOWIO(FAM)QVR-dP-G-NH ₂	68
2.5.3	Cyclic Peptide Cyclo (SIAD-dp-PDDIK).....	70
2.5.4	PawS-Derived Peptide Cyclo (C(s-Anapa)RKSIPPR (s-Anapa) CFPDDF).....	72
	References.....	74

Abstract The development of the field of peptide synthesis has been integral in studying their physiological role in living systems as well as their use in developing therapeutics for many human diseases. From the early work of Emil Fischer to the advent of solid-phase peptide synthesis by R. Bruce Merrifield to today, peptide synthesis continues to evolve, allowing access to increasingly complex and previously unobtainable peptides in the lab. In this chapter, we give a brief history of peptide synthesis and outline some of the most used protecting groups, resins, linkers, and coupling methodologies. We end the chapter with a brief description of commonly used peptide modifications and give examples produced in the author's laboratory.

T. Gauthier (✉) · D. Liu
Department of Agricultural Chemistry, LSU AgCenter, Baton Rouge, LA, USA
e-mail: tgauthier@agcenter.lsu.edu; doliu@agcenter.lsu.edu

Keywords Solid-phase peptide synthesis · Disulfide bond · Peptide conjugate · Protecting group · HPLC method · Cyclic peptides

2.1 Introduction

Protein chemistry is one of the most important parts of biochemistry. The majority of the cellular functions are carried out by proteins and hence are fundamental components of cells. Proteins have been recognized and studied since the eighteenth century and are considered large biomolecules consisting of one or more long chains of amino acids. In contrast, peptides are much smaller than proteins but also consist of a chain of amino acids. Traditionally, peptides are defined as molecules that consist of up to 30–50 amino acids, whereas proteins are made up of more than 100 amino acids. Peptides were not discovered until the early twentieth century with the discovery of secretin by Bayliss and Starling (1902). This discovery sparked an interest in searching for other peptides and in studying their physiological function.

Because of their relatively small size compared to proteins, peptides are often treated in the same way as small organic compounds. As a result, many labs quickly began focusing on trying to synthesize these novel biomolecules. Peptides consist of amino acids that are linked together by amide bonds, which are the result of a condensation reaction between the amino group of one amino acid and the carboxylic acid group of another. The challenges surrounding peptide synthesis soon became apparent, for example, controlling the order of the coupling reaction, increasing the yield of each coupling reaction, and maintaining the solubility of the growing peptide chain, just to name a few. The first successful synthesis of a peptide was accomplished by Emil Fischer (Fischer and Fourneau 1901). Fischer successfully synthesized the dipeptide glycine-glycine. The synthesis of this simple peptide was the beginning of the field of peptide synthesis that is still evolving today.

2.2 Fundamentals of Peptide Synthesis

There are numerous review articles and textbooks that cover peptide synthesis in detail. The reader is referred to those for a more in-depth description (Jaradat 2018). There is one particularly useful book that discusses the pragmatic issues involved in peptide synthesis (Chan and White 1999). What follows is a somewhat broad description of the peptide synthesis process.

2.2.1 *Solution-Phase Peptide Synthesis*

The early work in peptide synthesis was accomplished by synthesizing peptides in solution. The controlled addition of amino acids to the growing peptide chain is critical to synthesize the desired peptide. Controlling the order of Fischer's relatively simple dipeptide was not a concern since he used the same amino acid glycine. It quickly became apparent that once different amino acids were used in the condensation reaction, the order became of utmost importance.

To address this issue, it was necessary to block or "protect" the carboxylic acid and/or amino-functional groups of each amino acid yet easily remove them when the peptide synthesis was complete. The protection of the carboxylic acid of the C-terminal amino acid was accomplished easily by esterification. An easily removable amino protecting group for the N-terminal amino acid was developed in 1932 by Bergmann and Zervas (1932). The introduction of the carbobenzoxy (Cbz or simply Z) protecting group was a huge step forward in the nascent field of peptide synthesis. The Cbz group allowed chemists to control the order of the addition of amino acids in the peptide chain. The Cbz group was removed easily by hydrogenolysis, which greatly facilitated peptide synthesis. Using the Cbz group, scientists were able to synthesize simple peptides such as glutathione (Harington and Mead 1935) and carnosine (Sifferd and du Vigneaud 1935). In 1954, du Vigneaud reported the first successful synthesis of a biologically active peptide, the hormone oxytocin, a cyclic nonamer (du Vigneaud et al. 1954). He later reported the successful synthesis of another biologically active peptide, the nonamer vasopressin. For his work, du Vigneaud was awarded the 1955 Nobel Prize in Chemistry.

As the peptide synthesis field continued to advance, scientists developed additional protecting groups for amine functional groups as well as other side-chain functional groups of many amino acids. In 1957, the acid-labile protecting group *tert*-butoxycarbonyl (Boc) was synthesized by Carpino (1957). This protecting group, used in conjunction with the Cbz group, introduced the concept of orthogonality to peptide synthesis, i.e., that one protecting group could be removed while leaving the other intact. This was important because it allowed chemists to protect amino side-chain functional groups while allowing the α -nitrogen to take part in peptide bond formation. The use of these two protecting groups in tandem allowed researchers to synthesize even larger peptides such as the 39-mer adrenocorticotrophic hormone (ACTH) reported by Schwyzer and Sieber (1963).

The process to synthesize a peptide in solution is straightforward in theory but, in practice, can be challenging. The C-terminal amino acid is esterified, while its α -nitrogen remains unprotected. The second amino acid in the sequence is protected only at its α -nitrogen. This is important as the unprotected carboxylic acid and the unprotected amino group take part in the peptide bond formation to yield the dipeptide. The reaction conditions that are used to facilitate the peptide bond formation will be discussed later in the chapter. Once the peptide bond formation (often called coupling) is complete, the α -nitrogen protecting group is removed, and the crude dipeptide ester is purified. The coupling reaction is then repeated with the next

α -nitrogen protected amino acid. Once that reaction is complete, the crude trimer peptide ester is purified. This iterative process is repeated until the desired peptide is obtained. It should be noted that any amino acids that have reactive side-chain functional groups are protected with a protecting group, which is orthogonal to the α -nitrogen protecting group. The final step in the peptide synthesis is to remove the final α -nitrogen protecting group, any side-chain protecting groups, as well as the C-terminal ester. The peptide is then purified one final time to obtain the product. Figure 2.1 depicts the stepwise synthesis of a peptide in solution.

One issue with solution-phase synthesis is the need to purify the growing peptide chain after each coupling. Often, the purification can be difficult, especially in the case of deletion sequences where amino acids are left out due to inefficient coupling. Another factor contributing to purification difficulties is the by-products from the coupling reagents, which are formed during the peptide synthesis. Frequently, it is difficult and sometimes impossible to completely remove these by-products.

As scientists sought to synthesize larger peptides, another hurdle became apparent. There is a limit to the size of the growing peptide chain before solubility of the peptide becomes a problem. In general, the larger the peptide chain becomes, the less soluble it is in solvents that are used in the coupling reaction (e.g., dimethylformamide, N-methyl-2-pyrrolidone, tetrahydrofuran, etc.). Work-arounds to this issue, such as convergent synthesis (i.e., synthesizing fragments of the large peptide, then coupling the fragments), sometimes alleviate solubility issues but can be difficult to carry out.

2.2.2 *Solid-Phase Peptide Synthesis*

Chemists searched for methods to address the shortcomings of solution-phase synthesis. In 1963, Merrifield introduced the concept of solid-phase peptide synthesis (SPPS) (Merrifield 1963). This revolutionary idea opened up the field of peptide synthesis and allowed chemists to synthesize previously unobtainable peptides. For his work, Merrifield was awarded the Nobel Prize in Chemistry in 1984.

The concept of SPPS is relatively simple. The carboxylic acid of the C-terminal amino acid is anchored to insoluble polymeric support (also known as a resin). This support also serves the role as the protecting group for the C-terminal carboxylic acid. The C-terminal amino acid is also protected at its α -nitrogen using a protecting group that can be removed easily while leaving the amino acid anchored to the resin. Once the deprotection is complete, the by-products of the deprotection reaction can be removed by filtration. The next α -protected amino acid can be coupled, and this iterative process is repeated until the peptide synthesis is complete. Figure 2.2 depicts the stepwise synthesis of a peptide on a solid support.

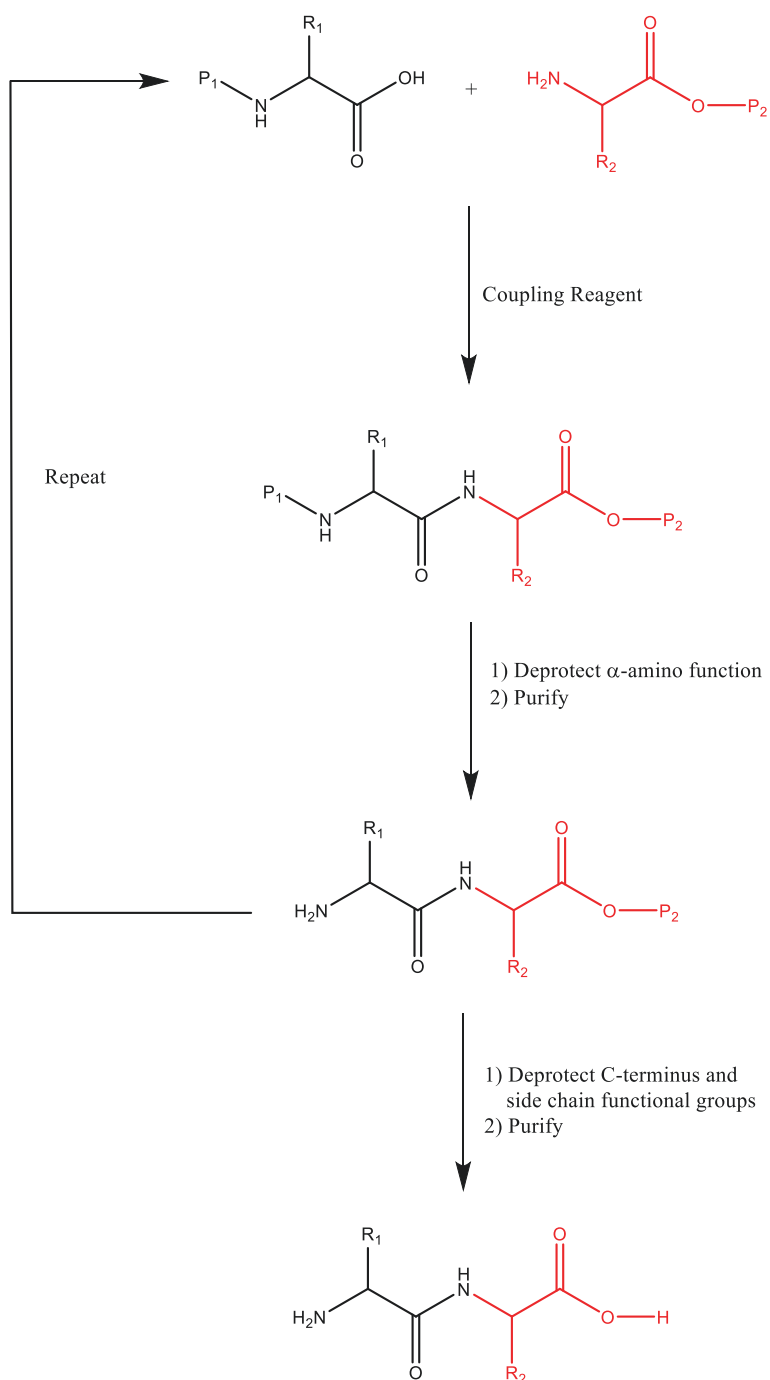


Fig. 2.1 Stepwise solution-phase peptide synthesis. P = orthogonal protecting groups; R = side-chain functional group

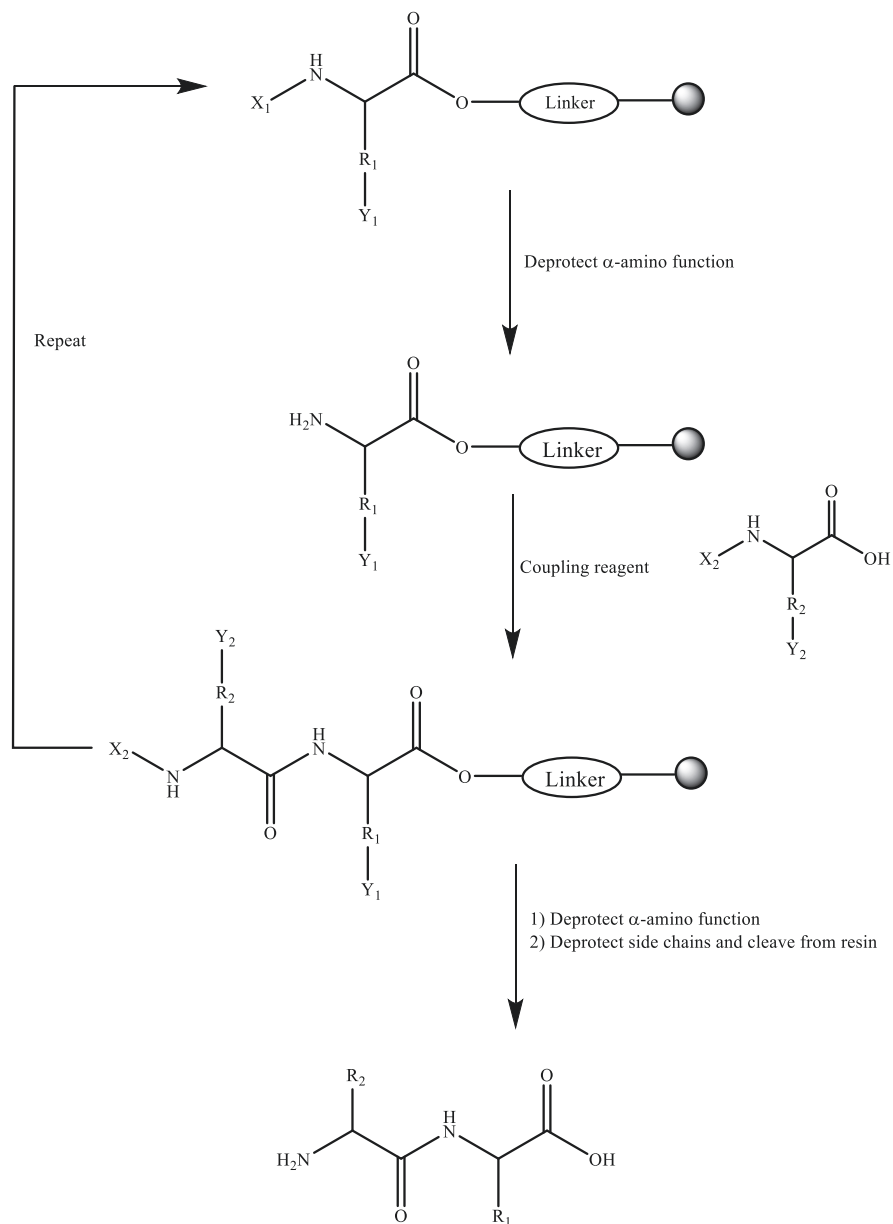


Fig. 2.2 Stepwise solid-phase peptide synthesis. R = amino acid side chain; Y = side-chain protecting group; X = α -amino protecting group

The first resin was a styrene-divinylbenzene copolymer, which was functionalized by chloromethylation. This resin came to be known as the Merrifield resin. The first amino acid was attached to the resin via an esterification linkage. At this point in time, the primary amino protecting groups were Boc (removable by mild acid) and Cbz (removable by hydrogenolysis or strong acid). Once a peptide synthesis was complete, the peptide was cleaved from the resin and simultaneously deprotected using strong acids such as anhydrous hydrogen fluoride (HF). The use of HF is a safety hazard and can lead to many unwanted side reactions.

SPPS took a leap forward in 1970 with the introduction of the base labile amino protecting group 9-fluorenylmethoxycarbonyl (Fmoc) by Carpino and Han (1970). The Fmoc protecting group is much easier to remove than the Cbz group, requiring only a mild base. Along with the use of the acid-labile Boc group, another orthogonal protecting scheme was developed. Many other protecting groups have since been developed, further increasing the usefulness of solid-phase peptide synthesis. Table 2.1 lists some of the common side-chain protecting groups used today.

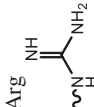
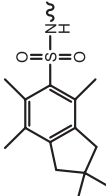
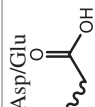
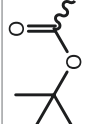
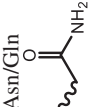
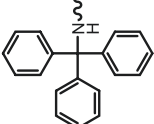
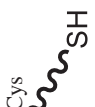
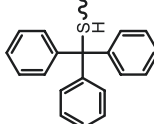
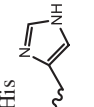
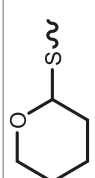
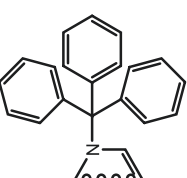
Still, strong acids such as HF were needed to cleave the peptide from the resin. In the ensuing years, many other functionalized resins were developed that eliminated the need for strong acid cleavage conditions. The most common functionalized resins will be discussed below. There are several publications and protocols related to solid-phase peptide synthesis (Behrendt et al. 2016; Hansen and Oddo 2015; Kremsmayr and Muttenthaler 2022; Hojlys-Larsen and Jensen 2013; Made et al. 2014; Guzman et al. 2021; Vanier 2013; Murray et al. 2011; Fagundez et al. 2018; Luna et al. 2016; Fields 2002; Winkler 2020; Jensen 2013; Stawikowski and Fields 2012).


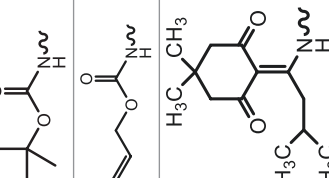


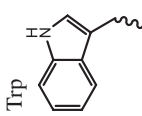
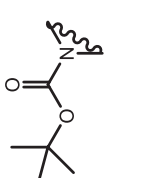
2.3 Practical Aspects of Peptide Synthesis

Before beginning a peptide synthesis, it is necessary to determine a synthesis strategy. Often, the first decision to be made is whether to carry out the synthesis in solution or on a solid support. The remainder of this chapter will focus on SPPS since that is often more relevant in a laboratory setting.

Once the decision to utilize SPPS is made, the next decision is to choose the base resin. Since the early days following Merrifield's introduction of SPPS, many resins have been developed, each with unique properties. The primary property to evaluate when choosing a resin is its swelling capacity. Most polymeric resins have pores on which the reactive sites available for peptide synthesis lie. Merrifield's resin polystyrene cross-linked with 1% divinylbenzene continues to be a popular choice today. Merrifield's resin has modest capacity to swell, which somewhat limits access to the reactive sites. In addition, the pores of this resin are relatively hydrophobic, which could cause synthesis issues if one is synthesizing a peptide that is hydrophobic and prone to self-aggregation. Today, there is a wide array of resin types with improved swelling properties. Among these resins are polystyrene cross-linked polyethylene

Table 2.1 Common side chain protecting groups used in Fmoc solid-phase peptide synthesis

Amino acid	Name	Abbreviations	Structure	Removal	Stable to
Arg 	2,2,4,6,7-pentamethyl-dihydrobenzofuran-5-sulfonyl	Pbf		95% TFA	Piperidine
Asp/Glu 	t-Butoxycarbonyl	Boc		95% TFA	1% TFA and piperidine
Asn/Gln 	Trityl	Trt		1% TFA in DCM	Piperidine
Cys 	Trityl and tetrahydropyranyl	Trt		1% TFA in DCM	Piperidine
His 	Trityl	Trt		95% TFA	Piperidine
				1% TFA in DCM	Piperidine

Amino acid	Name	Abbreviations	Structure	Removal	Stable to
Lys/Orn 	t-Butoxycarbonyl Allyloxycarbonyl 1-(4,4-Dimethyl-1,2,6-dioxocyclohexylidene)3-methylbutyl	Boc Alloc ivDde		95% TFA 3 eq. Pd (Ph ₃ P) ₄ in CHCl ₃ /AcOH/ NMM (37:2:1) 2% hydrazine in DMF	1% TFA Piperidine TFA, piperidine, and hydrazine Piperidine
Ser/Thr/Tyr 	t-butyl	tBu		95% TFA	1% TFA and piperidine
Trp 	t-Butoxycarbonyl	Boc		95% TFA	1% TFA and piperidine

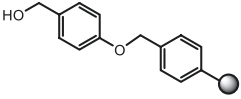
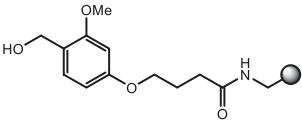
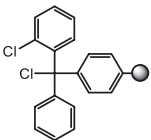
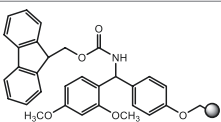
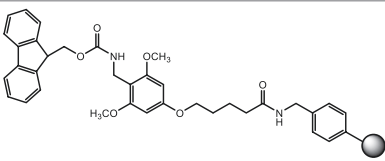
glycol (PEG) resins, polyacrylamide/PEG copolymer resins, and polyethylene cross-linked PEG resins. In addition to possessing superior swelling capacity, these types of resins tend to be relatively hydrophilic, limiting a hydrophobic peptide's ability to self-aggregate.

Another factor that must be accounted for is the resin's loading capacity, i.e., the number of reactive sites, expressed in millimoles per gram of resin. For example, a resin with a loading capacity of 0.5 mmol/g contains 0.5 mmol of reactive sites for each gram of resin. In general, for short peptides, resins with a high loading capacity (more than 0.5 mmol/g) are more commonly used, while for longer peptides, resins with low loading capacity (less than 0.5 mmol/g) are chosen.

Once a suitable polymeric support has been chosen, the next decision to be made is that of which linker to choose. The reactive sites of polymeric resins are functionalized to improve the attachment of the initial amino acid residue, allow cleavage under relatively mild conditions (sometimes without fully deprotecting the peptide), and determine what the C-terminal functional group will be. Often, the attachment of the first amino acid residue and desired C-terminal functional group go together. The most common C-terminal functional groups are C-terminal acids or C-terminal amides. For a C-terminal acid, it is recommended that the resin be purchased with the first Fmoc-protected amino acid already attached. In the cases where a C-terminal amide is preferred, the first Fmoc-protected amino acid can be added easily in the laboratory. Table 2.2 shows some of the most common linkers used today, along with their cleavage conditions and C-terminal functionality.

Once a resin and linker have been chosen, the next decision to be made is that of coupling chemistry. The peptide bond that links the amino acids together is the result of a condensation reaction. This condensation reaction requires an activation energy greater than can be supplied at room temperature alone. As a result, reagents have been developed to "activate" the carboxylic acid of one amino acid so that it can react readily with the amino group of another amino acid. The first activating reagents were simple acyl halides, acyl azides, or anhydrides. These work well and are still used today, but they have limitations such as amino acid racemization and low coupling efficiency in some sequences. Recognizing the need for better coupling reagents, scientists developed many coupling reagents in subsequent years to suit specific applications. The choice of coupling reagent is somewhat subjective, often relying on one's experience synthesizing peptides. Cost can also be a consideration with acyl halides, azides, anhydrides, and carbodiimides/additive pairs being relatively inexpensive compared to the more expensive aminium reagents. Cost of an activating agent can be less of a concern when one considers the cost of a failed synthesis. Most of these coupling reagents also require the use of a mild base (e.g., diisopropylethylamine, N-methylmorpholine, etc.) to facilitate the peptide bond formation. It should be noted that these bases are not strong enough to remove base labile protecting groups. Table 2.3 lists some of the most common coupling reagents and additives used today.

Table 2.2 Linkers commonly used for the synthesis of peptide acids and peptide amides

Resin name	Resin structure	Cleavage conditions	Peptide product
Wang resin		90–95% v/v TFA, 1–2 h	Acid
HMPB resin		1% v/v TFA in DCM, 2–5 min	Acid
2-chlorotrityl chloride resin		1% v/v TFA in DCM, 1 min	Acid
Rink amide resin		90–95% v/v TFA, 1–2 h	Amide
PAL		90–95% v/v TFA, 1–2 h	Amide

2.4 Peptide Modifications

Most peptides synthesized are simple linear peptides that require no modification. However, there are many circumstances where a peptide needs to be modified to suit a particular application. Some modifications can be addressed during the peptide synthesis with the choice of amino acid (e.g., biotin-labeled lysine). These modified amino acids tend to be quite expensive and can be difficult to couple efficiently. Frequently, performing these modifications post-synthesis can be easier to accomplish and is often the preferred method. In these instances, it is necessary to choose a protecting group that is orthogonal to all other protecting groups for the amino acid needing the modification. The most common modifications are fluorescent and biotin labeling. Such modifications of peptides can be achieved by linking the label (such as fluorescent tags or biotin) at the C-terminus, at the N-terminus, or on amino acid side-chain functional groups. These modifications have applications in enzymology, protein chemistry, immunology, and histochemistry.

Table 2.3 Common coupling reagents and additives

Name	Abbreviation	Structure	Type
N,N'-dicyclohexylcarbodiimide	DCC		Carbodiimide coupling reagent
N,N'-diisopropylcarbodiimide	DIPCDI, DIC		Carbodiimide coupling reagent
Benzotriazole-1-yloxytris(dimethyl amino) phosphonium hexafluorophosphate	BOP		Phosphonium coupling reagent
Benzotriazole-1-yloxytri(pyrrolidino)-phosphonium hexafluorophosphate	PyBOP		Phosphonium coupling reagent
[(7-azabenzotriazol-1-yl)oxy]tris-(pyrrolidino) phosphonium hexafluorophosphate	PyAOP		Phosphonium coupling reagent
1-hydroxybenzotriazole	HOBt		Carbodiimide additive
1-hydroxy-7-azabenzotriazole	HOAt		Carbodiimide additive
N-[(1H-benzotriazol-1-yl)(dimethylamino)methylene]-N-methylmethanaminium hexafluorophosphate N-oxide	HBTU		Aminium coupling reagent
N-[(dimethylamino)-1H-1,2,3-triazolo[4,5-b]pyridine-1-yl)methylene]-N-methylmethanaminium hexafluorophosphate N-oxide	HATU		Aminium coupling reagent
O-(1H-6-chlorobenzotriazole-1-yl)-1,1,1,3-tetramethyluronium hexafluorophosphate	HCTU		Aminium coupling reagent
Ethyl (2Z)-2-cyano-2-(hydroxyamino)acetate	Oxyma		Carbodiimide additive

2.4.1 Cyclization

The formation of cyclic peptides is a frequently used strategy for the development of peptides to enhance stability when compared to their linear analogs. Generally, cyclic peptides exhibit better *in vitro stability and, in some cases, in vivo stability*. *With proper conformational constraints to achieve stability and affinity, they exhibit enhanced binding toward target biomolecules and increased cell permeability* (Matsuda and Koyasu 2000; Sugita et al. 2021).

For head-to-tail (N-terminus to C-terminus) cyclic peptide synthesis, the linear peptide is usually synthesized using Fmoc chemistry on a highly acid labile resin such as 2-chlorotrityl chloride resin. Once the linear peptide synthesis is complete, the N-terminal protecting group is removed, and the peptide is cleaved from the resin with dilute TFA in DCM (usually 1% TFA), keeping the side-chain protecting groups on the linear peptide. The cyclization is completed in a solution using a coupling reagent such as HATU under basic conditions. The side-chain protecting groups are left intact to avoid undesired reactions happening during cyclization. Subsequently, deprotection of the peptide affords the fully deprotected cyclic peptide. This method requires a dilute peptide solution to minimize dimerization because of the intermolecular condensation of the linear sequences.

In addition to head-to-tail cyclic peptides, cyclic peptides also can be synthesized by utilizing side-chain cyclization strategies. The most common methods of achieving this type of cyclization include intramolecular disulfide bond formation, ring closing metathesis (stapled peptides) (Blackwell et al. 2001), and cyclic peptides that utilize click chemistry developed by K. Barry Sharpless (Li et al. 2013). Like head-to-tail cyclic peptides, these alternate cyclization strategies should be carried out using highly dilute peptide solutions to minimize undesirable intermolecular reactions.

2.4.2 Disulfide Bond Formation

Often, peptides containing multiple cysteine residues have disulfide bonds, which stabilize the secondary structure of the peptide, contributing to its biological activity. Peptides containing only two cysteine residues can only have one disulfide bond present. In general, these disulfide bonds are relatively simple to produce. When multiple disulfide bonds are required, the synthesis becomes much more difficult since control of the multiple disulfide bonds must be taken into consideration. This requires that each cysteine residue pair be protected with protecting groups that are orthogonal to each other to ensure the correct disulfide bonds are formed.

There are several methods that are commonly used to carry out the oxidation reaction that results in disulfide bond formation. The two most common methods are iodine oxidation and air oxidation. In the first example, the Cys deprotected peptide is dissolved in a suitable solvent such as 50% aqueous methanol or 50%

aqueous acetic acid. A dilute solution of I_2 (usually about 0.1 M) in the same solvent is then added slowly with vigorous stirring. After approximately 30 min, the reaction is quenched with ascorbic acid, and the disulfide peptide can be processed further. In the case of the second reaction, the Cys deprotected peptide is dissolved in a dilute solution of ammonium bicarbonate (usually about 0.1 M). The peptide solution is left stirring, exposed to the open atmosphere, until the reaction is complete. The disulfide peptide can then be processed further.

2.4.3 Automation of Solid-Phase Peptide Synthesis

The repetitive nature of SPPS makes it an ideal candidate for automation. Indeed, many companies currently sell instruments, peptide synthesizers, for this purpose. In general, there are two types of peptide synthesizers: batch synthesizers and continuous flow synthesizers.

Continuous flow synthesizers were among the first peptide synthesizers manufactured. In this instrument, the resin is packed into a glass column with filters at the top and bottom. Packing the column takes a measure of experience so that enough space is left to allow the resin to expand during the course of the synthesis. It is also necessary to pack it correctly so that channeling of the reagents and solvents does not occur, which results in reduced coupling efficiency. Once the column is packed correctly, it is placed on the system, which then directs solvents and reagents to flow through the resin on a continuous basis. These systems generally employ a pump and a series of valves to direct the correct solvent or reagent to the resin column while a peptide is being synthesized.

The other peptide synthesizer developed was the batch synthesizer. In this instrument, the resin is placed in a reaction vessel with no concern about packing the resin. All reactions and wash steps occur in a discontinuous fashion with the sequential addition of coupling and deprotection reagents. The reaction vessel is shaken, stirred, vortexed, or bubbled to agitate the resin to allow thorough mixing of the resin with the reagents. Excess reagents, solvents, and by-products are removed from the reaction vessel through a filter by the application of gas pressure or vacuum.

Batch synthesizers have taken over the industry as the preferred synthesizer type since they are more flexible and can be modified with new technologies. For example, over the last 20 years, microwave peptide synthesizers have become increasingly popular. Batch synthesizers can be easily modified to accommodate a chamber in which microwave energy is applied during coupling and deprotection steps. Many difficult peptide sequences have been synthesized using this technology (Bacsa et al. 2008). Batch synthesizers also allow for multiple channels, i.e., multiple reaction vessels, on the same instrument, thus allowing for the parallel synthesis of several peptides simultaneously. Lastly, batch synthesizers also allow for peptide syntheses to be scaled up for commercial production. Some vendors that produce synthesizers are Gyros Protein Technologies, CEM, Biotage, and AAPPTEC.

2.4.4 *Peptide Analysis and Purification*

Once the peptide synthesis is complete, the peptide will need to be analyzed for identification and purity. In most circumstances, purification will be needed to remove any by-products from the synthesis. A small aliquot of the crude peptide can be analyzed by reversed-phase HPLC equipped with a C18 column and UV detector to confirm the peptide was synthesized successfully. Normally, the detector is set to monitor a wavelength range of 215–230 nm as this is the wavelength range the peptide bonds absorb. The individual peaks can be collected and analyzed by either electrospray or matrix assisted laser desorption ionization (MALDI) mass spectrometry to determine which fraction contains the molecular weight corresponding to that of the desired peptide. At this point, the crude peptide is ready for final purification, usually by preparatory-scale HPLC using similar conditions to those used at the analytical scale. The peak corresponding to the desired peptide is collected, frozen, and lyophilized.

2.4.5 *Sustainability in Peptide Synthesis*

By far, solvents are the most heavily consumed reagent in peptide synthesis. Dimethylformamide (DMF), dichloromethane (DCM), and N-methyl-2-pyrrolidone (NMP) are among the most common solvents used in peptide synthesis and are among the most toxic solvents used in chemical synthesis. In addition, some coupling reagents, particularly carbodiimide reagents, illicit an allergic reaction in many people. Since peptides are increasingly accepted as viable pharmaceutical candidates, more attention has been focused on making their synthesis on the laboratory and commercial scale more environmentally friendly. Research has focused on ways to use solvents that are more sustainable. For example, the use of ChemMatrix® resin makes the use of water as a solvent possible (De Marco et al. 2013). Mixtures of so-called green solvents have also been explored as an alternative to DMF or NMP (Ferrazzano et al. 2019). Diisopropylcarbodiimide is used often as a substitute for dicyclohexylcarbodiimide to minimize the chances of an allergic response to the chemist. Additives, such as HOBT or HOAT, are required to minimize racemization of amino acids when using carbodiimide coupling reagents. Both additives pose explosion hazards and are sometimes difficult to obtain. As a result, OxymaPure is used as a greener additive. The reader is referred to a number of references for a more in-depth discussion of the various aspects of peptide synthesis sustainability (Varnava and Sarojini 2019; Wegner et al. 2021; Lawrenson et al. 2017; Jad et al. 2017; Pawlas and Rasmussen 2019; Pawlas et al. 2019; Isidro-Llobet et al. 2019; Al Musaimi et al. 2020; Kumar et al. 2021).

2.5 Examples of Peptide Syntheses Performed in the Authors' Laboratory

2.5.1 Linear Peptide EHWSY-dK-LRPG-NH₂

- (a) Gonadotropin-releasing hormone (GnRH) is a ten amino acid residue peptide which binds to its receptor gonadotropin receptor (GnR) and plays a key role in reproduction in vertebrates (Eidne et al. 1985). Native GnRH has the sequence pGlu¹-His²-Trp³-Ser⁴-Tyr⁵-Gly⁶-Leu⁷-Arg⁸-Pro⁹-Gly¹⁰-NH₂, where pGlu is pyroglutamyl. It is well known that the two motifs of the peptide, pGlu¹-His²-Trp³ and Arg⁸-Pro⁹-Gly¹⁰-NH₂, are essential for GnRH receptor binding. Structure activity studies suggested that when Gly amino acid at position 6 was changed to d-Lys, the peptide stability was enhanced and serves as an attachment point for macromolecular cargo such as a lytic peptide. The d-Lys⁶-GnRH peptide is used as a delivery vehicle to target the chemotherapy agents to GnRH receptors (Millar 2005; Baumann et al. 1993; Beckers et al. 2001; Nagy et al. 1996; Guo et al. 2011).
- (b) The title peptide was synthesized on a Tribute Peptide Synthesizer (Protein Technologies, Tucson, AZ) utilizing a standard Fmoc/*t*Bu strategy on a 100 μmol scale using Rink amide resin. Instead of pyroglutamic acid at the N-terminal, we used glutamic acid. Fivefold excess of Fmoc-amino acids (Fmoc-Glu (OtBu)-OH, Fmoc-His ((Trt)-OH, Fmoc-Trp (Boc)-OH, Fmoc-Ser (tBu)-OH, Fmoc-Tyr (tBu)-OH, Fmoc-d-Lys (Boc)-OH, Fmoc-Leu-OH, Fmoc-Arg (Pbf)-OH, Fmoc-Pro-OH, and Fmoc-Gly-OH) and HCTU, in the presence of 10 equivalents of N-methylmorpholine (NMM), was used for each of the amino acid coupling steps (10 min) with DMF as the solvent. Once the peptide synthesis was complete, the peptide was cleaved from the resin, and side-chain protecting groups were deprotected using 3 mL of the mixture TFA/water/TIPS (95:2.5:2.5) for 2 h and collected in a 50 mL centrifuge tube. The solution was concentrated in vacuo, and cold diethyl ether was then added to the peptide solution to precipitate the crude peptide. The peptide was centrifuged for 10 min at 4000 rpm, and the ether layer decanted. Fresh cold diethyl ether was added, and the pelleted peptide was resuspended. The peptide was centrifuged again, and the procedure was repeated five times in total. After the final ether wash, the peptide pellet was dissolved in 5 mL water containing 0.1% TFA, frozen, and lyophilized to obtain the crude target peptide as a powder.
- (c) Analysis steps: HPLC with a Waters 616 pump, Waters 2707 Autosampler, and 996 Photodiode Array Detector.
- (d) Peptide purification: Agilent Zorbax 300 SB-C18 (5 μm, 4.6 × 250 mm) with an Agilent guard column Zorbax 300 SB-C18 (5 μm, 4.6 × 12.5 mm).
- (e) Elution: A linear 5% to 55% gradient of solvent B (0.1% TFA in acetonitrile) into A (0.1% TFA in water) over 50 min at a 1 mL/min flow rate with UV detection at 215 nm.

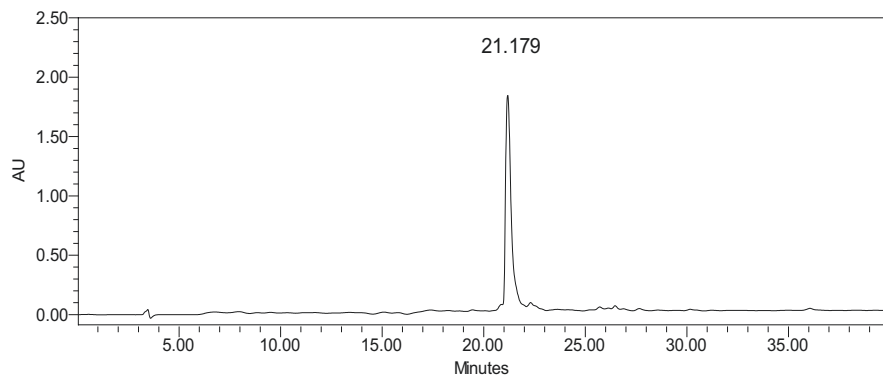


Fig. 2.3 HPLC chromatogram of EHWSY-dK-LRPG-NH₂

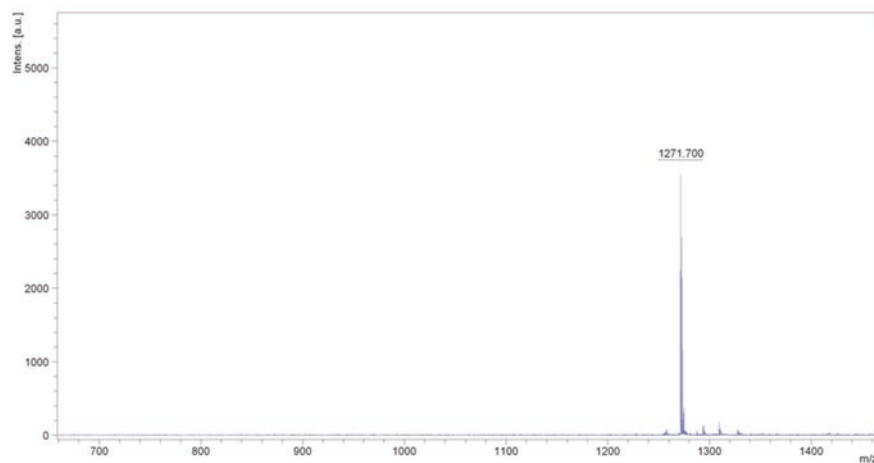


Fig. 2.4 MALDI-ToF mass spec of EHWSY-dK-LRPG-NH₂

- (f) The chromatogram showed the crude peptide contained a major peak corresponding to the correct peptide molecular weight determined by MALDI-TOF. Purification was carried out by preparatory HPLC, and fractions of high (>95%) HPLC purity with the expected mass were combined and lyophilized to obtain the pure peptide (Figs. 2.3 and 2.4).

2.5.2 *Fluorescently Labeled Peptide Ac-RWVOWIO(FAM) QVR-dP-G-NH₂*

- (a) There is much interest in cell-penetrating peptides (CPPs) as a tool for the delivery of large molecules inside cells (Bohmova et al. 2018; Hoffmann et al. 2018; Taylor and Zahid 2020). One of the criteria used in the design of cell-penetrating peptides is to use Arg amino acid along with hydrophobic amino acids to acquire the cell-penetrating properties (Walrant et al. 2020; Kauffman et al. 2015). The carboxyfluorescein (FAM) group used allows for quantification using both fluorometry (uptake) and HPLC (stability/degradation) (Squires et al. 2013). FAM group is attached to the side chain of ornithine. Such peptides are used to measure uptake kinetics and intracellular stability of CPPs (Patel et al. 2019; Cheung et al. 2009; Hallbrink et al. 2001).
- (b) The title peptide was synthesized on a Tribute Peptide Synthesizer (Protein Technologies, Tucson, AZ) utilizing a standard Fmoc/tBu strategy on a 100 μ mol scale using Rink amide resin. Fivefold excess of Fmoc-amino acids (Fmoc-Arg(Pbf)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Val-OH, Fmoc-d-Pro-OH, Fmoc-Gly-OH, Fmoc-Ile-OH, Fmoc-Gln(Trt)-OH, Fmoc-Orn(Alloc)-OH), and HCTU), in the presence of 10 equivalents of N-methylmorpholine (NMM), was used for each of the amino acid coupling steps (10 min) with DMF as the solvent. After completion, the Fmoc group of the N-terminus was removed with 20% piperidine in DMF, and the resin was washed with DMF and DCM.
- (c) Acetylation: A cocktail of acetic anhydride, NMM, and NMP (1:1:3) was added to the deprotected resin and shaken for 30 min.
- (d) FAM conjugation: Alloc group of the amino acid ornithine was removed with threefold excess of palladium (Tetrakis(triphenylphosphine) palladium(0)) in 4 ml of CHCl₃-HOAc-NMM (37:2:1) under nitrogen for 2 h. The resin was then washed with DCM followed by DMF (3 \times 30 sec), respectively.
- (e) Coupling of FAM to the delta nitrogen of the ornithine side chain was achieved with fourfold excess of FAM, HOBt, PyBOP, and DIEA in 3 mL DMF for 24 h. The same procedure was repeated for 8 h (if needed). Once again, the resin was washed with DMF and DCM (3 \times 30 sec), respectively.
- (f) Peptide cleavage: The peptide was cleaved from the resin and side chain deprotected using 4 mL of the mixture TFA/water/TIPS (95:2.5:2.5) for 3 h. After the cleavage, it was transferred to a 50 mL centrifuge tube. The solvent was concentrated in vacuo. To precipitate the peptide, chilled diethyl ether was added. To remove the ether layer, the peptide was centrifuged for 10 min (4000 rpm), and the ether layer decanted. Adding of chilled diethyl ether and centrifuging the peptide were repeated five times. After the final ether wash, the peptide pellet was dissolved in 5 mL water containing 0.1% TFA, frozen, and lyophilized to obtain the crude target peptide as a powder.
- (g) Analysis and purification: As described in Sect. 2.5.1 d and e.

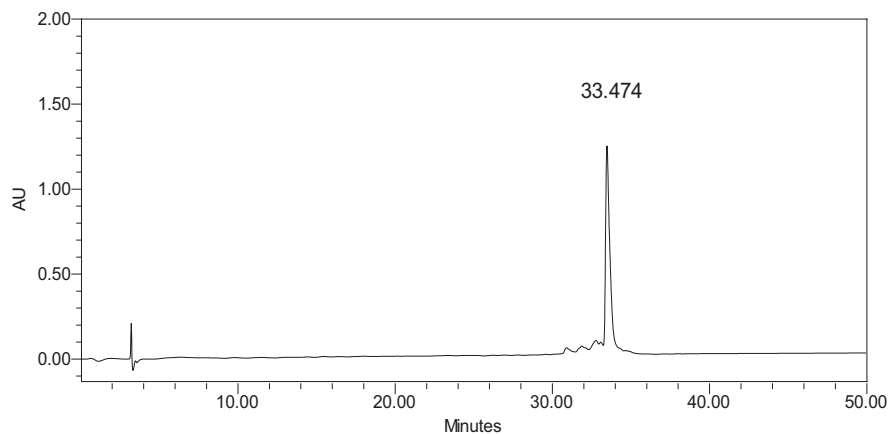


Fig. 2.5 HPLC chromatogram of Ac-RWVOWIO(FAM)QVR-dP-G-NH₂

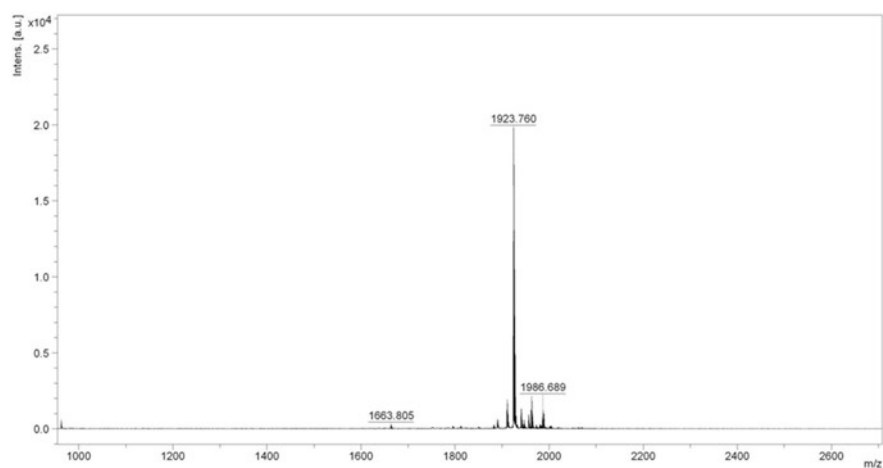


Fig. 2.6 MALDI-TOF mass spec of Ac-RWVOWIO(FAM)QVR-dP-G-NH₂

- (h) The chromatogram showed the crude peptide contained a major peak corresponding to the correct peptide molecular weight determined by MALDI-TOF. Purification was carried out by preparatory HPLC, and fractions of high (>95%) HPLC purity with the expected mass were combined and lyophilized to obtain the pure peptide (Figs. 2.5 and 2.6).

2.5.3 Cyclic Peptide Cyclo (SIAD-dp-PDDIK)

- (a) Peptides can be designed to inhibit protein-protein interactions (PPI) since inhibition of PPI can lead to therapeutic effects for autoimmune diseases and in different types of cancer. Co-stimulatory molecules CD2 and CD58 are expressed on the surface of T cells and antigen-presenting cells, respectively (van der Merwe and Davis 2003; Chen and Flies 2013). The interaction of CD2 and CD58 helps hold the two cells together and generate the signal for immune response (Davis et al. 2003). By targeting one of these proteins, the PPI of CD2-CD58 can be modulated and hence suppress the immune response. Peptides from the CD2 adhesion domain can be designed to target CD58 in autoimmune diseases such as rheumatoid arthritis (Balanesu et al. 2002; Gokhale et al. 2011; Gokhale et al. 2015). These peptides are designed from the adhesion epitopes of CD2 protein and contain a beta-strand structure. To induce the beta-strand structure in the designed peptide, Pro-Pro sequence was introduced, and the peptide was stabilized by backbone cyclization (Gokhale et al. 2015; Gokhale et al. 2013; Parajuli et al. 2021). It is well known that Pro-Gly, Pro-Pro, and d-pro-l=Pro motifs induce β -turn and β -strand structures in peptides (Chatterjee et al. 2008; Kantharaju et al. 2009; Belwal et al. 2020). The design of the structure activity of this peptide is described in the literature (Parajuli et al. 2021).
- (b) The linear peptide was synthesized with the Tribute Peptide Synthesizer (Protein Technologies, Tucson, AZ) utilizing a standard double-coupling Fmoc peptide synthesis strategy on a 50 μ mol scale using H-Pro CTC resin (0.53 mmol/g). The Fmoc group was deprotected with 20% piperidine in DMF. Fivefold excess of Fmoc-amino acids (Fmoc-Ala-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Ile-OH, Fmoc-Lys(Boc)-OH, Fmoc-d-Pro-OH, Fmoc-Ser(tBu)-OH, and Fmoc-Tyr(tBu)-OH) and HCTU, in the presence of 10 equivalents of N-methylmorpholine (NMM), was used for each of the amino acid coupling steps (10 min) with DMF as the solvent. After completion, the Fmoc group of N-terminus was removed with 20% piperidine in DMF, and the resin was washed with DMF and DCM. A 3 mL of TFE/DCM (50:50) was added to the resin and mixed for 2 h. The peptide solution was collected in a clean flask. This process was repeated one more time. The combined peptide solutions were concentrated in vacuo to remove the DCM and TFE.
- (c) The linear peptide residue was dissolved in a 15 mL solution of TFE/DMF (80:20) containing PyAOP (4 eq, 104 mg) and DIEA (8 eq, 70 μ l). The solution was agitated for 2 h for head-to-tail cyclic peptide formation. The solvent was removed in vacuo, and the residue was transferred to a 50 mL centrifuge tube. A 4 mL of TFA/EDT/water (95:2.5:2.5) was added to the residue and agitated for 2 h to remove the side-chain protecting groups.

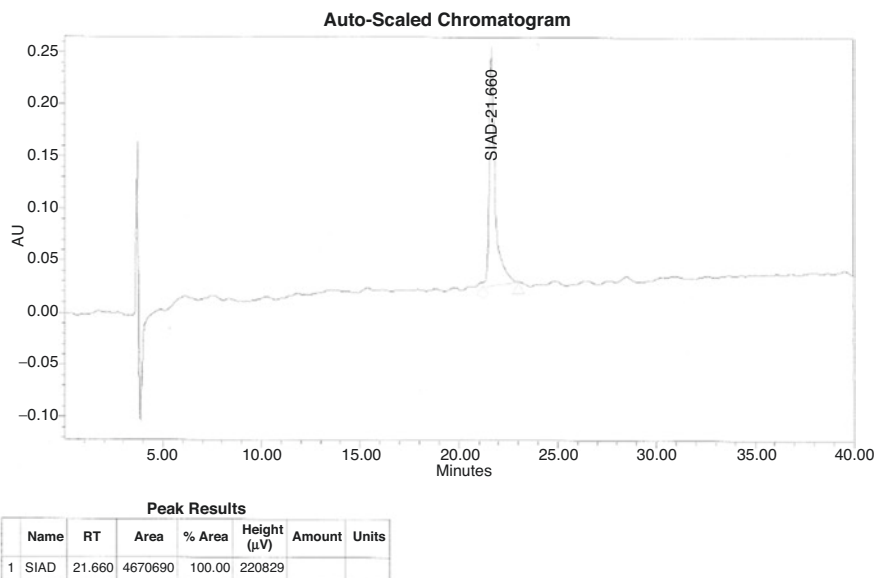


Fig. 2.7 HPLC chromatogram of cyclo (SIADpPDDIK)

Cold diethyl ether (40 mL) was then added to the solution to precipitate the crude peptide. The peptide was centrifuged for 10 min at 4000 rpm, and the ether layer decanted. The pelleted peptide was resuspended with fresh cold diethyl ether and centrifuged again. This procedure was repeated five times in total. After the final ether wash, the peptide pellet was dissolved in 5 mL water containing 0.1% TFA, frozen, and lyophilized to obtain the crude cyclic peptide.

- (d) Analysis and purification: As described above, Sect. 2.5.1 d.
- (e) The HPLC peak corresponding to the correct peptide was confirmed by MALDI-TOF. Purification was carried out by prep-HPLC on an AAPPTEC Spirit Peptide 120-C18 column (5 μm, 21.2 × 250 mm) with an AAPPTEC preparative guard column 120 C18 Peptide (5 μm, 21.2 × 15 mm). Elution was done with a linear 5–55% gradient of solvent B (0.1% TFA in acetonitrile) into A (0.1% TFA in water) over 50 min at a 1 mL/min flow rate with UV detection at 215 nm. The fractions of high (>95%) HPLC purity with the expected mass were combined and lyophilized to obtain the pure cyclic peptide, 10 mg (Figs. 2.7 and 2.8).

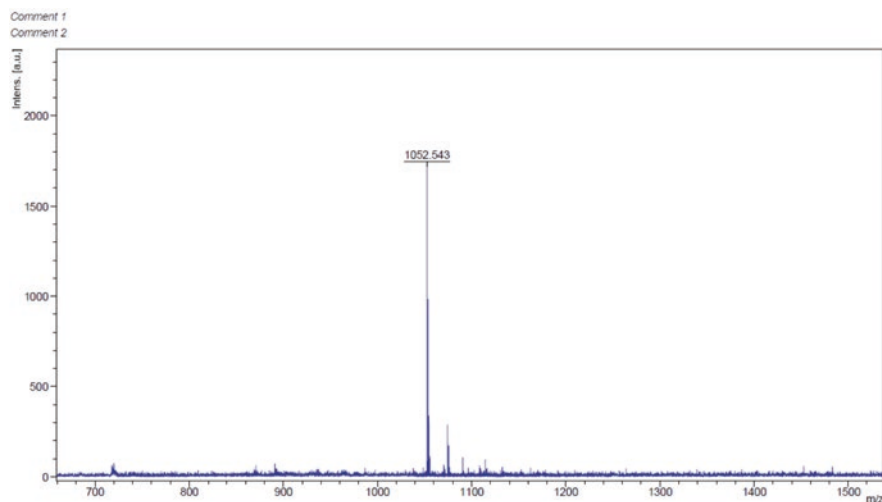


Fig. 2.8 MALDI-ToF mass spec of cyclo (SIADpPDDIK)

2.5.4 *PawS-Derived Peptide Cyclo (C(s-Anapa)RKSIPPR (s-Anapa) CFPDDF)*

- (a) PDPs are a diverse class of single disulfide-bonded peptides originating from preproalbumin with sunflower trypsin inhibitor 1 (SFTI-1) (PawS1) proteins in the seeds. These are derived from post-translational processing of seed-storage albumin genes. Sunflower trypsin inhibitor (SFTI-1), a cyclic peptide comprised of 14 amino acids found mainly in sunflower seeds and is a prototype family member of PDPs. The structure consists of a cyclic backbone and is stabilized by a disulfide bridge. Due to its small size, extensive hydrogen-bonding network, and compact rigidity, it is one of the most widely employed molecular scaffolds for drug discovery (Colgrave et al. 2010; Gentilucci et al. 2010; Lesner et al. 2011; de Veer et al. 2019). Singh et al. have shown that peptidomimetics that target HER2 protein and inhibit EGFR dimerization can be grafted onto the SFTI-1 framework (Singh et al. 2021). These peptides have main chain cyclization via head to tail cyclization, and the peptide is further stabilized by disulfide bond cyclization. The synthesis of these peptides is slightly different and involves automated synthesis as well as manual disulfide bond formation. The protecting groups used for amino acids are slightly different, and removal of peptide from the resin requires mild conditions. Synthesis of such peptides are described earlier by Dr. Craik's group and others (Cheneval et al. 2014; Singh et al. 2021). For details of the peptide design, readers can refer to Singh et al. and Kanthala

et al. (Singh et al. 2021; Kanthala et al. 2017). Here, we describe the synthesis of a grafted peptide.

- (b) The linear peptide was synthesized with Tribute Peptide Synthesizer (Protein Technologies, Tucson, AZ) utilizing a standard Fmoc/*t*Bu peptide synthesis strategy on 100 μ mol scale using H-Pro-CTC resin. The Fmoc group was deprotected with 20% piperidine in DMF. Fivefold excess of Fmoc-amino acids (Fmoc-Cys(Thp)-OH, Fmoc-(*s*-3-amino-3-(1-naphthyl) propionic acid (*s*-Anapa), Fmoc-Arg(Pbf)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Ser(*t*Bu)-OH, Fmoc-Ile-OH, Fmoc-Pro-OH, Fmoc-Phe-OH, and Fmoc-Asp(OtBu)-OH) and HCTU), in the presence of 10 equivalents of *N*-methylmorpholine (NMM), was used for each of the amino acid coupling steps (10 min) with DMF as the solvent. After completion, the Fmoc group of the N-terminus was removed with 20% piperidine in DMF, and the resin was washed with DMF and DCM. A 3 mL of 1% TFA in DCM was added to the resin and mixed for 5 min, and the liquid solution was collected in a clean flask. This step was repeated 10 times, each time collecting the solution in the same flask. The solvent was then evaporated in vacuo.
- (c) The linear peptide residue was dissolved into 50 mL of DMF to make a 2 mM peptide solution. HATU (95 mg) and DIEA (88 μ L) were added to the solution to make a HATU solution concentration of 5 mM and a DIEA solution concentration of 10 mM. The solution was agitated for 2 h to allow the head-to-tail cyclic peptide formation to occur. The solvent was transferred to a 50 mL centrifuge tube and then removed in vacuo. A 4 mL mixture of TFA/EDT/TIPS/water (94:2.5:1:2.5) was added to the peptide residue and agitated for 3 h to remove the side-chain protecting groups. Cold diethyl ether (40 mL) was then added to the peptide solution to precipitate the crude peptide. The peptide was centrifuged for 10 min at 4000 rpm, and the ether layer decanted. The pelleted peptide was resuspended with fresh cold diethyl ether and centrifuged again. This procedure was repeated five times in total. After the final ether wash, the peptide pellet was dissolved in 5 mL water containing 0.1% TFA, frozen, and lyophilized to obtain the crude cyclic peptide. The crude cyclic peptide was then dissolved into 100 mL of 50% acetonitrile/water at a concentration of 1 mM. A saturated iodine solution in 50% acetonitrile/water was added to the crude cyclic peptide solution dropwise while stirring until the solution color changed to yellow. The solution was stirred for 1 h, after which time ascorbic acid was added to stop the reaction. The solution was lyophilized to obtain the crude cyclotide.
- (d) Analysis and purification: As described above in Sect. 2.5.1 d and e.
- (e) The chromatogram showed the crude cyclotide contained a major peak corresponding to the correct molecular weight determined by MALDI-TOF. Purification was carried out by preparatory HPLC, and fractions of high (>95%) HPLC purity with the expected mass were combined and lyophilized to obtain the pure cyclotide (Figs. 2.9 and 2.10).

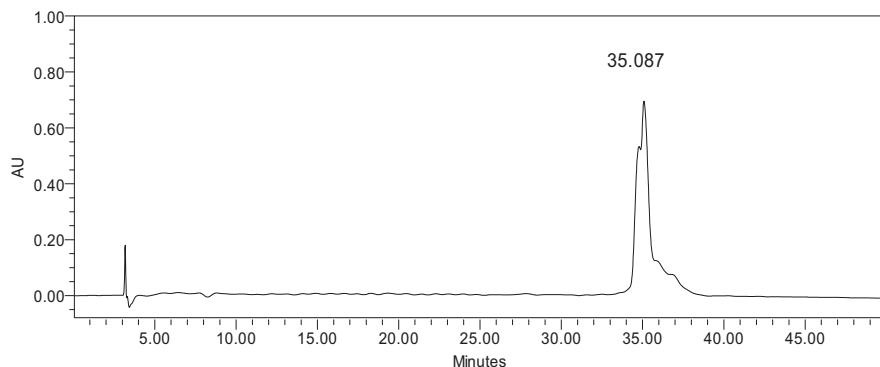


Fig. 2.9 HPLC chromatogram of cyclo (C(s-Anapa)RKSIPPR (s-Anapa)CFPDDF)

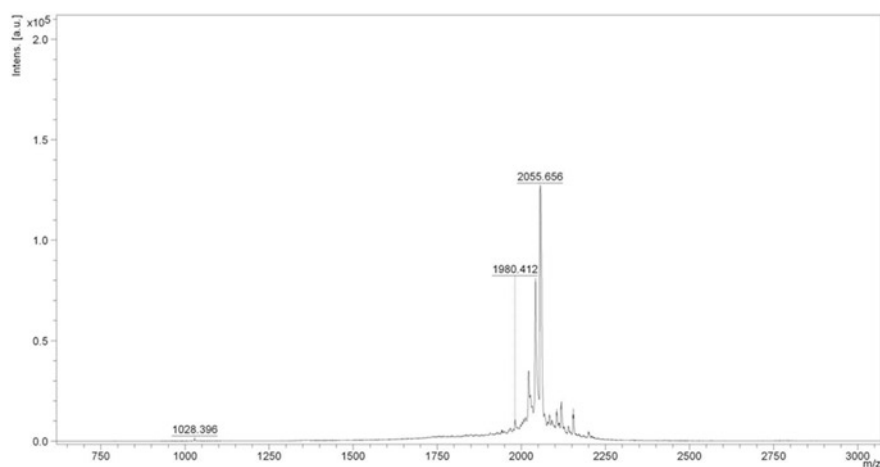


Fig. 2.10 MALDI-ToF mass spec of cyclo (C(s-Anapa)RKSIPPR (s-Anapa)CFPDDF)

References

- Al Musaimi O, de la Torre BG, Albericio F. Greening Fmoc/*t*Bu solid-phase peptide synthesis. *Green Chem.* 2020;22:996–1018.
- Bacsa B, Horvati K, Bosze S, Andreae F, Kappe CO. Solid-phase synthesis of difficult peptide sequences at elevated temperatures: a critical comparison of microwave and conventional heating technologies. *J Org Chem.* 2008;73:7532–42.
- Balanescu A, Radu E, Nat R, Regalia T, Bojinca V, Predescu V, Predeteanu D. Co-stimulatory and adhesion molecules of dendritic cells in rheumatoid arthritis. *J Cell Mol Med.* 2002;6:415–25.
- Baumann KH, Kiesel L, Kaufmann M, Bastert G, Runnebaum B. Characterization of binding sites for a GnRH-agonist (buserelin) in human breast cancer biopsies and their distribution in relation to tumor parameters. *Breast Cancer Res Treat.* 1993;25:37–46.

- Bayliss WM, Starling EH. On the causation of the so-called 'peripheral reflex secretion' of the pancreas. *Proc R Soc Lond.* 1902;69:352–3.
- Beckers T, Bernd M, Kutscher B, Kuhne R, Hoffmann S, Reissmann T. Structure-function studies of linear and cyclized peptide antagonists of the GnRH receptor. *Biochem Biophys Res Commun.* 2001;289:653–63.
- Behrendt R, White P, Offer J. Advances in Fmoc solid-phase peptide synthesis. *J Pept Sci.* 2016;22:4–27.
- Belwal VK, Datta D, Chaudhary N. The beta-turn-supporting motif in the polyglutamine binding peptide QBP1 is essential for inhibiting huntingtin aggregation. *FEBS Lett.* 2020;594:2894–903.
- Bergmann M, Zervas L. Über ein allgemeines Verfahren der Peptid-Synthese. *Ber Dtsch Chem Ges.* 1932;65:1192–201.
- Blackwell HE, Sadowsky JD, Howard RJ, Sampson JN, Chao JA, Steinmetz WE, O'Leary DJ, Grubbs RH. Ring-closing metathesis of Olefinic peptides: design, synthesis, and structural characterization of macrocyclic helical peptides. *J Org Chem.* 2001;66:5291–302.
- Bohmova E, Machova D, Pechar M, Pola R, Venclikova K, Janouskova O, Etrych T. Cell-penetrating peptides: a useful tool for the delivery of various cargoes into cells. *Physiol Res.* 2018;67:S267–S79.
- Carpino LA. Oxidative reactions of Hydrazines. II. Isophthalimides. New protective groups on nitrogen. *J Am Chem Soc.* 1957;79:98–101.
- Carpino LA, Han GY. 9-Fluorenylmethoxycarbonyl function, a new base-sensitive amino-protecting group. *J Am Chem Soc.* 1970;92:5748–9.
- Chan W, White P. Fmoc solid phase peptide synthesis: a practical approach. Oxford: OUP; 1999.
- Chatterjee B, Saha I, Raghothama S, Aravinda S, Rai R, Shamala N, Balam P. Designed peptides with homochiral and heterochiral diproline templates as conformational constraints. *Chemistry.* 2008;14:6192–204.
- Chen L, Flies DB. 'Molecular mechanisms of T cell co-stimulation and co-inhibition', *nature reviews. Immunology.* 2013;13:227–42.
- Cheneval O, Schroeder CI, Durek T, Walsh P, Huang YH, Liras S, Price DA, Craik DJ. Fmoc-based synthesis of disulfide-rich cyclic peptides. *J Org Chem.* 2014;79:5538–44.
- Cheung JC, Kim Chiaw P, Deber CM, Bear CE. A novel method for monitoring the cytosolic delivery of peptide cargo. *J Control Release.* 2009;137:2–7.
- Colgrave ML, Korsinczyk MJ, Clark RJ, Foley F, Craik DJ. Sunflower trypsin inhibitor-1, proteolytic studies on a trypsin inhibitor peptide and its analogs. *Biopolymers.* 2010;94:665–72.
- Davis SJ, Ikemizu S, Evans EJ, Fugger L, Bakker TR, van der Merwe PA. The nature of molecular recognition by T cells. *Nat Immunol.* 2003;4:217–24.
- De Marco R, Tolomelli A, Greco A, Gentilucci L. Controlled solid phase peptide bond formation using N-Carboxyanhydrides and PEG resins in water. *ACS Sustain Chem Eng.* 2013;1:566–9.
- de Veer SJ, Kan MW, Craik DJ. Cyclotides: from structure to function. *Chem Rev.* 2019;119:12375–421.
- du Vigneaud V, Ressler C, Swan JM, Roberts CW, Katsoyannis PG. The synthesis of oxytocin. *J Am Chem Soc.* 1954;76:3115.
- Eidne KA, Flanagan CA, Millar RP. Gonadotropin-releasing hormone binding sites in human breast carcinoma. *Science.* 1985;229:989–91.
- Fagundes C, Sellanes D, Serra G. Synthesis of cyclic peptides as potential anti-Malaria. *ACS Comb Sci.* 2018;20:212–9.
- Ferrazzano L, Corbisiero D, Martelli G, Tolomelli A, Viola A, Ricci A, Cabri W. Green solvent mixtures for solid-phase peptide synthesis: a Dimethylformamide-free highly efficient synthesis of pharmaceutical-grade peptides. *ACS Sustain Chem Eng.* 2019;7:12867–77.
- Fields GB. Introduction to peptide synthesis. *Curr Protoc Mol Biol.* 2002; <https://doi.org/10.1002/0471140864.ps1801s26>. Chapter 11: Unit 11 15
- Fischer E, Fournau E. Ueber einige Derivate des Glykocolls. *Ber Dtsch Chem Ges.* 1901;34:2868–77.
- Gentilucci L, De Marco R, Cerisoli L. Chemical modifications designed to improve peptide stability: incorporation of non-natural amino acids, pseudo-peptide bonds, and cyclization. *Curr Pharm Des.* 2010;16:3185–203.

- Gokhale A, Weldeghiorghis TK, Taneja V, Satyanarayanajois SD. Conformationally constrained peptides from CD2 to modulate protein-protein interactions between CD2 and CD58. *J Med Chem.* 2011;54:5307–19.
- Gokhale A, Kanthala S, Latendresse J, Taneja V, Satyanarayanajois S. Immunosuppression by co-stimulatory molecules: inhibition of CD2-CD48/CD58 interaction by peptides from CD2 to suppress progression of collagen-induced arthritis in mice. *Chem Biol Drug Des.* 2013;82:106–18.
- Gokhale AS, Sable R, Walker JD, McLaughlin L, Kousoulas KG, Jois SD. Inhibition of cell adhesion and immune responses in the mouse model of collagen-induced arthritis with a peptidomimetic that blocks CD2-CD58 interface interactions. *Biopolymers.* 2015;
- Guo H, Lu J, Hathaway H, Royce ME, Prossnitz ER, Miao Y. Synthesis and evaluation of novel gonadotropin-releasing hormone receptor-targeting peptides. *Bioconjug Chem.* 2011;22:1682–9.
- Guzman F, Gauna A, Roman T, Luna O, Alvarez C, Pareja-Barrueto C, Mercado L, Albericio F, Cardenas C. Tea bags for Fmoc solid-phase peptide synthesis: an example of circular economy. *Molecules.* 2021;26
- Hallbrink M, Floren A, Elmquist A, Pooga M, Bartfai T, Langel U. Cargo delivery kinetics of cell-penetrating peptides. *Biochim Biophys Acta.* 2001;1515:101–9.
- Hansen PR, Oddo A. Fmoc solid-phase peptide synthesis. *Methods Mol Biol.* 2015;1348:33–50.
- Harington CR, Mead TH. Synthesis of glutathione. *Biochem J.* 1935;29:1602–11.
- Hoffmann K, Milech N, Juraja SM, Cunningham PT, Stone SR, Francis RW, Anastas M, Hall CM, Heinrich T, Bogdawa HM, Winslow S, Scobie MN, Dewhurst RE, Florez L, Ong F, Kerfoot M, Champain D, Adams AM, Fletcher S, Viola HM, Hool LC, Connor T, Longville BAC, Tan YF, Kroeger K, Morath V, Weiss GA, Skerra A, Hopkins RM, Watt PM. A platform for discovery of functional cell-penetrating peptides for efficient multi-cargo intracellular delivery. *Sci Rep.* 2018;8:12538.
- Hojlys-Larsen KB, Jensen KJ. Solid-phase synthesis of phosphopeptides. *Methods Mol Biol.* 2013;1047:191–9.
- Isidro-Llobet A, Kenworthy MN, Mukherjee S, Kopach ME, Wegner K, Gallou F, Smith AG, Roschangar F. Sustainability challenges in peptide synthesis and purification: from R&D to production. *J Org Chem.* 2019;84:4615–28.
- Jad YE, Govender T, Kruger HG, El-Faham A, de la Torre BG, Albericio F. Green solid-phase peptide synthesis (GSPPS) 3. Green solvents for Fmoc removal in peptide chemistry. *Org Process Res Dev.* 2017;21:365–9.
- Jaradat D's MM. Thirteen decades of peptide synthesis: key developments in solid phase peptide synthesis and amide bond formation utilized in peptide ligation. *Amino Acids.* 2018;50:39–68.
- Jensen KJ. Solid-phase peptide synthesis: an introduction. *Methods Mol Biol.* 2013;1047:1–21.
- Kanthala SP, Liu YY, Singh S, Sable R, Pallerla S, Jois SD. A peptidomimetic with a chiral switch is an inhibitor of epidermal growth factor receptor heterodimerization. *Oncotarget.* 2017;8:74244–62.
- Kanharaju S, Raghobama U, Raghavender S, Aravinda S, Shamala N, Balaran P. Conformations of heterochiral and homochiral proline-pseudoproline segments in peptides: context dependent cis-trans peptide bond isomerization. *Biopolymers.* 2009;92:405–16.
- Kauffman WB, Fuselier T, He J, Wimley WC. Mechanism matters: a taxonomy of cell penetrating peptides. *Trends Biochem Sci.* 2015;40:749–64.
- Kremsmayr T, Muttenthaler M. Fmoc solid phase peptide synthesis of oxytocin and analogues. *Methods Mol Biol.* 2022;2384:175–99.
- Kumar A, Sharma A, de la Torre BG, Albericio F. Rhodiasolv PolarClean - a greener alternative in solid-phase peptide synthesis. *Green Chem Lett Rev.* 2021;14:545–50.
- Lawrenson SB, Arav R, North M. The greening of peptide synthesis. *Green Chem.* 2017;19:1685–91.
- Lesner A, Legowska A, Wysocka M, Rolka K. Sunflower trypsin inhibitor 1 as a molecular scaffold for drug discovery. *Curr Pharm Des.* 2011;17:4308–17.
- Li H, Aneja R, Chaiken I. Click chemistry in peptide-based drug design. *Molecules.* 2013;18:9797–817.

- Luna OF, Gomez J, Cardenas C, Albericio F, Marshall SH, Guzman F. Deprotection reagents in Fmoc solid phase peptide synthesis: moving away from Piperidine? *Molecules*. 2016;21
- Made V, Els-Heindl S, Beck-Sickinger AG. Automated solid-phase peptide synthesis to obtain therapeutic peptides. *Beilstein J Org Chem*. 2014;10:1197–212.
- Matsuda S, Koyasu S. Mechanisms of action of cyclosporine. *Immunopharmacology*. 2000;47:119–25.
- Merrifield RB. Solid phase peptide synthesis. I. the synthesis of a Tetrapeptide. *J Am Chem Soc*. 1963;85:2149–54.
- Millar RP. GnRHs and GnRH receptors. *Anim Reprod Sci*. 2005;88:5–28.
- Murray JK, Aral J, Miranda LP. Solid-phase peptide synthesis using microwave irradiation. *Methods Mol Biol*. 2011;716:73–88.
- Nagy A, Schally AV, Armatis P, Szepeshazi K, Halmos G, Kovacs M, Zarandi M, Groot K, Miyazaki M, Jungwirth A, Horvath J. Cytotoxic analogs of luteinizing hormone-releasing hormone containing doxorubicin or 2-pyrrolinodoxorubicin, a derivative 500-1000 times more potent. *Proc Natl Acad Sci U S A*. 1996;93:7269–73.
- Parajuli P, Sable R, Shrestha L, Dahal A, Gauthier T, Taneja V, Jois S. Modulation of co-stimulatory signal from CD2-CD58 proteins by a grafted peptide. *Chem Biol Drug Des*. 2021;97:607–27.
- Patel SG, Sayers EJ, He L, Narayan R, Williams TL, Mills EM, Allemann RK, Luk LYP, Jones AT, Tsai YH. Cell-penetrating peptide sequence and modification dependent uptake and sub-cellular distribution of green fluorescent protein in different cell lines. *Sci Rep*. 2019;9:6298.
- Pawlas J, Rasmussen JH. *ReGreen* SPPS: enabling circular chemistry in environmentally sensible solid-phase peptide synthesis. *Green Chem*. 2019;21:5990–8.
- Pawlas J, Nuijens T, Persson J, Svensson T, Schmidt M, Toplak A, Nilsson M, Rasmussen JH. Sustainable, cost-efficient manufacturing of therapeutic peptides using chemo-enzymatic peptide synthesis (CEPS). *Green Chem*. 2019;21:6451–67.
- Schwzyer R, Sieber P. Total synthesis of adrenocorticotrophic hormone. *Nature*. 1963;199:172–4.
- Sifferd RH, du Vigneaud V. A new synthesis of carnosine, with some observations on the splitting of the benzyl group from carbobenzoxy derivatives and from benzylthioethers. *J Biol Chem*. 1935;108:753–61.
- Singh SS, Mattheolabakis G, Gu X, Withers S, Dahal A, Jois S. A grafted peptidomimetic for EGFR heterodimerization inhibition: implications in NSCLC models. *Eur J Med Chem*. 2021;216:113312.
- Squires S, Christians E, Riedel M, Timothy D, Rodesch CK, Marvin J, Benjamin I. Effects of redox state on the efficient uptake of cell permeable peptide in mammalian cells. *Open Biochem J*. 2013;7:54–65.
- Stawikowski M, Fields GB. Introduction to peptide synthesis. *Curr Protoc Protein Sci*. 2012; <https://doi.org/10.1002/0471140864.ps1801s69>. Chapter 18: Unit 18 1
- Sugita M, Sugiyama S, Fujie T, Yoshikawa Y, Yanagisawa K, Ohue M, Akiyama Y. Large-scale membrane permeability prediction of cyclic peptides crossing a lipid bilayer based on enhanced sampling molecular dynamics simulations. *J Chem Inf Model*. 2021;61:3681–95.
- Taylor RE, Zahid M. Cell penetrating peptides, novel vectors for gene therapy. *Pharmaceutics*. 2020;12
- van der Merwe PA, Davis SJ. Molecular interactions mediating T cell antigen recognition. *Annu Rev Immunol*. 2003;21:659–84.
- Vanier GS. Microwave-assisted solid-phase peptide synthesis based on the Fmoc protecting group strategy (CEM). *Methods Mol Biol*. 2013;1047:235–49.
- Varnava KG, Sarojini V. Making solid-phase peptide synthesis greener: a review of the literature. *Chem Asian J*. 2019;14:1088–97.
- Walrant A, Bauza A, Girardet C, Alves ID, Lecomte S, Illien F, Cardon S, Chaianantakul N, Pallerla M, Burlina F, Frontera A, Sagan S. Ionpair-pi interactions favor cell penetration of arginine/tryptophan-rich cell-penetrating peptides. *Biochim Biophys Acta Biomembr*. 2020;1862:183098.
- Wegner K, Barnes D, Manzor K, Jardine A, Moran D. Evaluation of greener solvents for solid-phase peptide synthesis. *Green Chem Lett Rev*. 2021;14:153–64.
- Winkler DFH. Automated solid-phase peptide synthesis. *Methods Mol Biol*. 2020;2103:59–94.

Chapter 3

Computational Methods for Peptide Macrocycle Drug Design



Vikram Khipple Mulligan

Contents

3.1	Introduction.....	80
3.2	Available Computational Tools for Peptide Macrocycle Drug Design.....	84
3.2.1	Computational Concepts.....	84
3.2.2	Heuristic Approaches.....	87
3.2.3	Software Approaches for Peptide Drug Modelling.....	89
3.3	Pipelines for Peptide Macrocycle Drug Design.....	97
3.3.1	Target Modelling.....	98
3.3.2	Initial Stub Identification.....	100
3.3.3	Sampling Peptide Conformation.....	101
3.3.4	Designing Peptide Sequence.....	104
3.3.5	Computational Validation of Designs.....	124
3.3.6	Experimental Validation of Top Designs.....	130
3.4	Emerging Technologies.....	131
3.4.1	Deep Neural Networks and Other Machine Learning Methods.....	131
3.4.2	Quantum Computing.....	134
3.5	Conclusions.....	136
	Appendixes.....	137
	Appendix A: Running Examples.....	137
	Appendix B: Code Listings.....	139
	References.....	155

Abstract Peptide macrocycles represent a promising class of therapeutics, albeit one that is under-represented amongst existing drugs. One disadvantage of macrocycles, however, is that they can be more conformationally heterogeneous than small molecules or large, well-folded proteins. This flexibility can impede high-affinity binding. In recent years, the development of new computational tools has made possible the structure-based design of macrocycles that are able to fold into rigid structures compatible with binding to target proteins of therapeutic interest.

V. K. Mulligan (✉)

Center for Computational Biology, Flatiron Institute, New York, NY, USA

e-mail: vmulligan@flatironinstitute.org

This chapter is intended to introduce biologists, chemists, and drug developers to current computational methods for peptide macrocycle drug design. It introduces computational concepts such as parallelism and algorithmic complexity and outlines general algorithmic approaches such as Monte Carlo and simulated annealing methods. It also describes the thermodynamics of a flexible macrocycle binding to a target protein and explores molecular dynamics and Monte Carlo methods for sampling backbone conformations, deterministic and heuristic methods for designing amino acid sequences, and large-scale sampling-based methods for computationally validating and ranking designs to prioritize the likeliest candidates for chemical synthesis and experimental validation. Particular focus is given to methods implemented in the Rosetta software suite, with detailed examination of a Rosetta design protocol that was previously used to create peptide macrocycle inhibitors of an antibiotic resistance factor, the New Delhi metallo- β -lactamase 1 (NDM-1). Finally, this chapter describes new and emerging technologies that promise to enhance computational peptide macrocycle drug design, such as deep learning and quantum computing.

Keywords Peptide macrocycles · Rational drug design · Structure-guided design · Rosetta · Molecular dynamics · Simulation · Machine learning · Algorithms · Heuristics · Quantum computing

3.1 Introduction

The conventional drug discovery pipeline is resource-intensive and inefficient, typically requiring the screening of hundreds of thousands or millions of chemical compounds in order to find a handful of hits, perhaps one or two of which may be refined into leads that could be carried forward for preclinical evaluation and clinical trials. Even in phase 3 clinical trials, a majority of drugs (54%) fail. Most failures at this stage can be attributed to poor efficacy (57% of phase 3 failures) or intolerable side effects (17% of phase 3 failures) (Hwang et al. 2016). Because such late-stage failures occur after enormous investment of time, effort, and resources, any means of shifting failures to earlier phases of the drug discovery process can improve efficiency. When used well, the computation can be invaluable for ruling out more drug candidates with far less investment of time, money, or experimental resources. In addition to being useful to filter candidates out (eliminating some of the “hay” when searching for the “needle in the haystack”), rational computational design approaches can be applied to enrich the pool of hits for molecules more likely to succeed as leads and more likely to pass the preclinical evaluation and clinical trials (ensuring that more or better “needles” are present in the “haystack” in the first place) (Fig. 3.1). By enriching the candidate pool for productive molecules, computational predictions can reduce the resource cost of bringing drugs to the clinic even when such predictions have far less than 100% accuracy. Until recently, most computational tools for drug development have focused either on docking, in silico screening, and design of small molecules or on sequence design of large proteins. Very

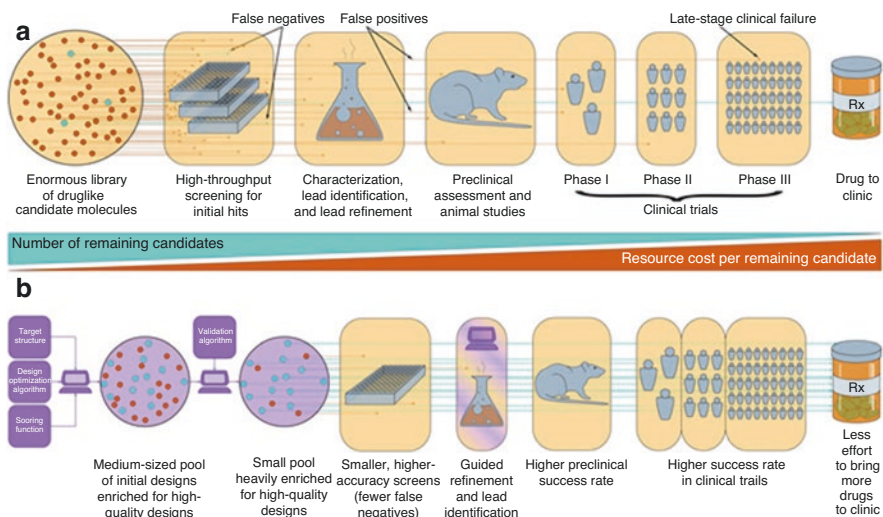
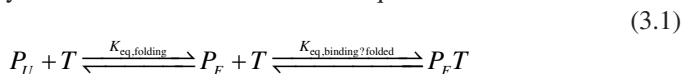


Fig. 3.1 The role of computation in drug discovery. Computational steps are shaded violet, and wet-lab experimental steps are shaded yellow. **(a)** The traditional drug discovery pipeline. Drugs are typically discovered by screening an enormous library (often 10^5 to 10^6 compounds) of candidate molecules *in vitro* to identify initial hits. This pool is presumed to contain some useful drugs (cyan) amongst many ineffective, toxic, or otherwise unusable compounds (brown). Due to its high-throughput nature, the initial screen is plagued by high false-positive and false-negative rates, as represented as brown dashed lines passing through (false positives) and cyan lines halted (false negatives) at screening steps. Initial hits are refined and characterized to identify leads in lower-throughput experiments, and leads are carried forward for preclinical evaluation, animal studies, and, ultimately, phased clinical trials (Lombardino and Lowe 2004). Even in phase III clinical trials, 54% of candidates fail (Hwang et al. 2016). As one progresses through the drug discovery pipeline, the number of candidates remaining in the test pool falls (cyan wedge), and the total resource cost per candidate rises (brown wedge), meaning that late-stage failures can be extremely costly. **(b)** The drug discovery pipeline augmented with computation. The computational design allows the production of an initial pool of rationally designed molecules that is potentially substantially enriched for usable drugs, and computational validation methods allow *in silico* screening to eliminate more of the non-productive candidates prior to wet-lab experiments. This allows a smaller pool of candidates, one that is heavily enriched for productive molecules, to be screened in lower-throughput assays with lower false-positive or false-negative rates. Hit refinement to produce leads for preclinical, animal, and clinical evaluation can also be guided by computation to further increase the likelihood of leads succeeding. The ultimate effect is to shift failures to earlier, less expensive stages of the pipeline and to allow more successes in later stages, allowing more drugs to be brought to the clinic with less cost in time, effort, or resources

recently, computational tools have been developed to facilitate the rational design of peptide macrocycles able to bind to target proteins of therapeutic interest (Mulligan 2020).

Synthetic peptides, and particularly peptide macrocycles, represent a very promising class of drugs. These molecules lie between large protein therapeutics and small molecules in size and combine many of the advantages of each class of molecule. Compared to small molecules, they present large surfaces for specific target

recognition, potentially permitting higher target affinity and fewer off-target effects. Compared to proteins, peptides offer greater potential for penetrating barriers such as the intestinal epithelium, the blood-brain barrier, or even the plasma membrane of cells (Mulligan 2020). Since peptides can be synthesized from thousands of non-canonical amino acid building blocks, they can access chemical functionalities beyond those accessible to genetically encodable proteins, which must be built from the 20 canonical amino acids. And the potential for non-canonical composition permits better evasion of proteases and of the immune system, making these molecules more usable as drugs (Dintzis et al. 1993; Weinstock et al. 2012). The Achilles' heel of peptides, however, has historically been intrinsic conformational flexibility: Where small molecules possess few degrees of conformational freedom, and proteins fold into rigid structures, peptides are often disordered in isolation, meaning that they must order themselves to bind to a target. The binding of a peptide macrocycle to a target may be considered to occur via two equilibria:



In the above, P_U represents the unfolded, unbound peptide; P_F represents the folded, unbound peptide; T represents the unbound target protein; and $P_F T$ represents the complex of the folded peptide bound to the target. The two equilibria are described by the equilibrium constants $K_{\text{eq, folding}}$ and $K_{\text{eq, binding|folded}}$. Each equilibrium has its associated change in Gibbs free energy:

$$\Delta G_{\text{folding}} = -RT \ln(K_{\text{eq, folding}}) \quad (3.2)$$

$$\Delta G_{\text{binding|folded}} = -RT \ln(K_{\text{eq, binding|folded}}) \quad (3.3)$$

The overall change in Gibbs free energy on binding, $\Delta G_{\text{binding}}$, is given by Eq. 3.4:

$$\Delta G_{\text{binding}} = \Delta G_{\text{folding}} + \Delta G_{\text{binding|folded}} \quad (3.4)$$

Conventionally, $\Delta G_{\text{binding}}$ is defined as the Gibbs free energy of the bound, ordered state minus that of the unbound, disordered state so that more negative values are considered to favour binding more. Since Gibbs free energy can be decomposed into enthalpic and entropic components, the above can be written as:

$$\Delta G_{\text{binding}} = \Delta H_{\text{binding}} - T \Delta S_{\text{binding}} = \Delta H_{\text{folding}} + \Delta H_{\text{binding|folded}} - T(\Delta S_{\text{folding}} + \Delta S_{\text{binding|folded}}) \quad (3.5)$$

In Eq. 3.5, $\Delta H_{\text{binding}}$ is the overall change in enthalpy (equal to the sum of the enthalpy changes of the two steps), T is the absolute temperature (typically the physiological temperature of 310 K), and $\Delta S_{\text{binding}}$ is the overall change in entropy

(equal to the sum of the entropy change of the two steps). As one moves to longer and more flexible peptides, the unbound, unfolded state grows more disordered relative to the bound, folded state. This means that $\Delta S_{\text{folding}}$ grows more negative, which tends to make $\Delta G_{\text{binding}}$ more positive (less favourable), resulting in lower-affinity binding. The degree of disorder of the unbound conformation can be reduced by limiting oneself to peptide macrocycles, in which two or more residues that are distant in linear sequence are covalently connected (e.g., by joining the termini with an amide bond or by connecting cysteine residues near the termini with a disulphide bond) (Wang et al. 2015). Nevertheless, macrocyclization alone is not a guarantee of a rigid structure.

Computational design tools enable the design of peptide sequences that both present maximally favourable binding interactions with a target protein (predominantly optimizing the enthalpy of binding given a folded peptide, $\Delta H_{\text{binding} | \text{folded}}$, and thereby the overall change in enthalpy, $\Delta H_{\text{binding}}$) while simultaneously promoting rigidly folded peptides that uniquely favour the binding-competent conformation (predominantly minimizing the entropic cost of folding, $\Delta S_{\text{folding}}$, and thereby the overall entropic cost of binding, $\Delta S_{\text{binding}}$) (Mulligan 2020). Indeed, it is possible to design peptides for which $\Delta G_{\text{folding}}$ is negative, that is, for which the binding-competent, folded state of the peptide is predominantly populated even when the peptide is unbound. These spontaneously folding peptide macrocycles are primarily stabilized by internal hydrogen bonds and by the intrinsic conformational preferences of amino acid residues (Bhardwaj et al. 2016; Hosseinzadeh et al. 2017), which contrasts with well-folded proteins that are primarily stabilized by the exclusion of water from a well-packed hydrophobic core (Dill et al. 1995). Of the tools available, this chapter will focus predominantly on the Rosetta software suite, though alternative and complementary software tools will also be discussed.

Thermodynamic Considerations for Peptide Macrocycle Drug Design

1. A disordered peptide macrocycle must order itself to bind to its target.
2. The ordering or folding of the macrocycle and the binding of the folded macrocycle are each described by a Gibbs free energy change. Negative Gibbs free energy changes result in spontaneous folding or binding (i.e., an equilibrium favouring the folded or bound state).
3. Computational tools can produce macrocycles that interact favourably with their targets and which bind spontaneously when folded.
4. Computational tools can also produce macrocycles for which the binding-competent fold is stabilized by internal hydrogen bonds and intrinsic conformational preferences of constituent amino acids, resulting in spontaneous folding.
5. The combination of spontaneous folding and spontaneous binding when folded yields overall spontaneous binding.

3.2 Available Computational Tools for Peptide Macrocycle Drug Design

Computational drug design and validation tools aim to make useful predictions or to generate hypotheses that can guide experiments for drug discovery; however, even with computational predictions in hand, only experiments can yield true results. The challenge is to find useful predictions that computers can make that can narrow down intractably large numbers of possibilities (such as all possible drug molecules) to pools of high-likelihood candidates that are small enough to test experimentally. Many types of computational tools exist, and all of these offer trade-offs in speed (which determines the number of drug candidates that can be considered or the number of predictions that can be made) and accuracy (which determines the number of experiments that must be performed in order to find a computational prediction that is borne out in reality). The choice of computational tool for a given problem depends on the nature of the problem, the available computing power, and the resources available for experimental validation. In this section, the available tools are reviewed, and their relative advantages and disadvantages discussed.

Important Caveats of Computation in Drug Discovery

1. Computational methods yield predictions or hypotheses, not results. Only experiments yield results.
2. Computational methods balance speed and accuracy.
3. Because computational methods have finite accuracy, a useful computational prediction narrows an intractably large set of possibilities to a set of possibilities small enough to be tested experimentally.

3.2.1 Computational Concepts

Before considering available software methods, let us consider the hardware on which these methods will run. Modern scientific computing is typically performed on computing *clusters* divided into *nodes*. Each computing node typically has its own *central processing unit* (CPU) on which computations are performed, its own *memory* in which data are stored transiently during a computation, and possibly its own *graphics processing unit* (GPU). Modern CPUs typically have multiple computing *cores*, each of which is capable of executing instructions in sequence to carry out an algorithm; at the time of this writing, 4–8 cores are common, though high-performance CPUs may have as many as 128 cores. CPU cores tend to be optimized for fast computation. In contrast, GPUs use the strategy of providing thousands of relatively slow cores, with fast communication between them, which collectively can perform a great deal of work in a short period of time.

It is convenient to have a common way to discuss the computational *expense* or *cost* of a particular task. Units of time may be intuitive for this purpose; however, because many tasks are *parallelizable*, or divisible over many computing cores, the computing cost of a task is typically measured in *core-hours*, where one core-hour is the amount of work that a single computing core can do in an hour. This is analogous to measuring human labour in person-hours. Admittedly, the concept of a core-hour is somewhat imprecise since different core architecture can result in different numbers of computations per second; however, in the past 15 years or so, most of the advancements in CPU design have increased parallel performance and decreased power consumption, while the maximum performance of a single core has not changed much. Those tasks that can be parallelized *efficiently* can be performed by two computing cores in half the time or by N cores in N-fold less time. Tasks that can be divided into smaller blocks of work that can be carried out entirely independently, in any order, and without any cross-communication can be parallelized very efficiently, and these tasks are called *embarrassingly parallelizable*. In drug discovery, one common embarrassingly parallel approach is to perform a particular calculation on many different drug candidates on each of several cores to produce many predictions at once, with no cross-communication necessary. However, there also exist many types of computational tasks that require some cross-communication between parallel blocks of work, and since there is a cost to communication or to synchronization, these show diminishing returns as more cores are applied to the task. Some tasks are not parallelizable with any efficiency at all.

It is most common for a parallelized computing task to use, at most, all the cores and all of the memory of a single computing node. When the work is divided into smaller blocks that each use one core and which can access the same memory, we describe this as *threaded parallelism*. When work is divided into smaller blocks that are each allocated their own piece of memory, we describe this as *process-based parallelism*. Since threads share memory, they have the advantage of collectively requiring less memory (just as many accountants writing to the same ledger require only one ledger) but must carefully synchronize their reads and writes to avoid overwriting one another's work. This synchronization can hinder performance. However, threading often permits very finely grained subdivision of a task, with very rapid crosstalk between parallel parts of the algorithm. Processes are analogous to many accountants each writing to his or her own ledger and each exchanging messages with the others to share information. Processes require less synchronization than threads, allowing more independent execution that often yields better performance; however, when processes do need to exchange information, this can be slower and can reintroduce synchronization issues. Since processes must each be allocated dedicated memory, the total memory requirements are higher with processes than with threads. However, processes have the advantage of being distributable across many computing nodes (which may not even be located in the same place), allowing a single algorithm to scale to use tens or even hundreds of thousands of CPUs. Hybrid approaches, in which many processes that can pass messages to one another coarsely divide work and then each process finely divides its portion of the work over many threads that can share memory, are also possible.

A final consideration is the *scaling* or *complexity* of the algorithms that will be considered here. Although complexity theory is its own field (for an overview, see Bovet and Crescenzi (1994)), we will examine some of the key concepts briefly. As the complexity of the input into an algorithm increases (e.g., the number of atoms in a particular simulation or the number of amino acids in a peptide), computing time and memory can increase linearly (the ideal case) or non-linearly. Although the precise scaling tends to have a complicated expression, computer scientists typically disregard all terms but the one that dominates as the complexity of the inputs goes to infinity. *Big-O notation* is most commonly used to compare the scaling of different algorithms. If N represents the size of the inputs to an algorithm, an algorithm that scales well will be $O(N)$ (i.e., linear) in both memory use and computing time. Quadratic scaling (i.e., $O(N^2)$) is somewhat worse but often still acceptable – indeed, computer scientists often consider any polynomial-time algorithm to scale well. Those algorithms that are $O(C^N)$ (i.e., which scale exponentially) are among the ones that are least computationally tractable.

Combinatorial search problems, such as the problem of designing an amino acid sequence for a peptide, are unfortunately often $O(C^N)$: Given C possibilities for the amino acid identity at each position and a peptide that is N amino acid residues long, there are C^N possible sequences, and an algorithm which exhaustively enumerates these sequences will have a runtime that shows $O(C^N)$ scaling. If one were designing a 10-amino acid peptide using the 20 canonical amino acids, one would have 20^{10} , or about 10^{13} (ten trillion), possible sequences to enumerate, a number that would tax the largest supercomputers in the world. Although there exist problems with exponentially scaling solution spaces that can be searched intelligently by efficient deterministic algorithms, the protein design problem is one of a class called *NP-complete* (Pierce and Winfree 2002). The question of whether there exists any deterministic algorithm that can solve *NP-complete* problems efficiently is a major unsolved problem in computer science (for a full primer, see Chapter 2 of Goldreich (2010)). In the absence of a known, efficient, deterministic approach, *heuristic* approaches are often used to help to make exponentially scaling problems computationally tractable, albeit by sacrificing strong guarantees that an optimal solution will be found. These are described in the next section.

Important Hardware and Software Considerations for the Drug Designer

1. Computational drug design and validation tasks are frequently parallelizable across computing cores or across hardware nodes.
2. Threads and processes offer advantages and disadvantages in memory requirements, degree of cross-communication between parallel workers, and ease of distribution across a large computing cluster.
3. Algorithm scaling with input size is a frequent concern. Linear scaling ($O(N)$) is a best-case scenario, but many problems in drug discovery involve exponentially scaling solution spaces. A subset of these problems is *NP-complete*, and there is no known polynomial-time deterministic algorithm that can find a best solution from the vast solution space. As such, algorithms that solve these shows non-polynomial, and usually exponential ($O(C^N)$), scaling.

3.2.2 Heuristic Approaches

Although it is attractive to have strong guarantees that an algorithm will yield an optimal solution to a problem, exponential scaling often makes such exact algorithms impractical. Heuristic methods provide “shortcuts” to finding near-optimal solutions, at the expense of dropping strong guarantees about the optimality of the solution. Although many heuristic methods exist, *Monte Carlo methods* are frequently convenient for the difficult combinatorial problems that one encounters in heteropolymer modelling and drug discovery. Such methods use random sampling to explore the space of solutions to a given problem. The random trajectory through the solution space is structured so that, as the number of samples approaches infinity, the probability of finding solutions arbitrarily close to optimality approaches 100%. Since infinitely long trajectories are impossible in actual use, practically, Monte Carlo methods may be thought of as methods that allow rapid sampling from a subset of the total solutions that is enriched in useful or acceptable (albeit suboptimal) solutions. However, convergence to optimal or even near-optimal solutions in reasonable time cannot be guaranteed with these methods. Despite this, Monte Carlo methods are popular since they work well across a wide range of problem structures and impose very few requirements on the form of the function being optimized.

The *Metropolis-Hastings algorithm* is most commonly applied to peptide and protein modelling tasks (Metropolis et al. 1953; Hastings 1970). Given some function that one is trying to optimize (e.g., an energy function), which takes some input (e.g., a peptide sequence) and produces a single value as output (e.g., an energy or score), the Metropolis-Hastings approach is to start with a random starting state and to introduce a series of random perturbation “moves”. After each move, the effect of the move on the output from the function is evaluated. If the value decreases, the move is always accepted, and the next move is performed using the result of the current move as a starting point. If the value increases, the move is accepted or rejected with conditional probability depending on the amount by which the value increases. If the move is rejected, the state is reset to the state resulting from the last move that was accepted. At the end of a trajectory of thousands to millions of moves, the state that was encountered that had the lowest value of the objective function is returned as the “best guess” for the optimal state in the entire search space. The commonly used acceptance criterion (called the *Metropolis criterion*) is given by Eq. 3.6:

$$\begin{aligned}
 & E_{\text{new}} - E_{\text{old}} \leq 0 : \text{acceptmove} \\
 E_{\text{new}} - E_{\text{old}} > 0 : & \begin{cases} e^{\frac{-(E_{\text{new}} - E_{\text{old}})}{k_B T}} \leq r : \text{rejectmove} \\ e^{\frac{-(E_{\text{new}} - E_{\text{old}})}{k_B T}} > r : \text{acceptmove} \end{cases}
 \end{aligned} \tag{3.6}$$

In the above, E_{old} and E_{new} are the values of the objective function before and after the move, respectively, and r is a value drawn randomly from a uniform distribution ranging from 0 to 1, allowing moves that increase the energy to be accepted or rejected with conditional probability that approaches 100% as the size of the increase approaches zero and 0% as the increase grows large. The temperature factor, $k_B T$, determines how the probability of rejecting a move increases as the magnitude of the increase in the objective function grows: at large values, most moves are accepted, allowing trajectories to move up out of local minima in the search space, and at small values, only those moves that decrease the objective function or produce very tiny increases are accepted, resulting in a tendency to move towards the nearest local minimum.

Given a potential function in which all barriers between wells are finite and moves that have the property of obeying the reversibility condition or detailed balance criterion (Hastings 1970) as given by Eq. 3.7, a Metropolis-Hastings search is ergodic, meaning that a sufficiently long trajectory will produce an ensemble of samples with properties representing the entire distribution of states. (This is in contrast to a process that is “trapped” searching some subset of the search space and which can only produce ensembles of samples representative of that subset. Such a process would be nonergodic.) If moves obey the detailed balance criterion, then an infinitely long Monte Carlo trajectory performed at a constant temperature factor, $k_B T$, will not only sample all points in the search space (ergodicity) but will produce a distribution matching the thermodynamic or Boltzmann distribution at that temperature. Indeed, this relation to thermodynamics gives the temperature factor its name.

$$\pi(\bar{x}_1)P(\bar{x}_2|\bar{x}_1) = \pi(\bar{x}_2)P(\bar{x}_1|\bar{x}_2) \quad (3.7)$$

In the above, $\pi(\bar{x}_1)$ and $\pi(\bar{x}_2)$ represent the probabilities of states \bar{x}_1 and \bar{x}_2 , respectively, $P(\bar{x}_2|\bar{x}_1)$ is the probability of transitioning from state \bar{x}_1 to state \bar{x}_2 , and $P(\bar{x}_1|\bar{x}_2)$ is the probability of transitioning from state \bar{x}_2 to state \bar{x}_1 . Intuitively, the detailed balance criterion may be considered to be a condition of reversibility: The sampling itself is not biasing the trajectory in one particular direction. Practically, even Monte Carlo approaches that do not strictly obey detailed balance can be useful for finding near-optimal solutions to difficult optimization problems without enormous computing resources, even if they do introduce some bias to the distribution of states sampled (Manousiouthakis and Deem 1999).

While a long Monte Carlo trajectory will ultimately produce a thermodynamic distribution in which the lowest-energy state is the most probable, a typical Monte Carlo trajectory does not come close to approximating the effect of a trajectory of infinite length; that is, Monte Carlo trajectories on large and rugged search spaces can be practically nonergodic, taking intractably long times to move from local well to local well and preventing some regions of search space from contributing to the ensemble of samples at all. To speed things up and to increase the probability of

sampling the lowest-energy state in a reasonable period of time, simulated annealing approaches are often used (Kirkpatrick et al. 1983; Černý 1985; Bertsimas and Tsitsiklis 1993). These also carry out Monte Carlo trajectories using the Metropolis-Hastings algorithm, but rather than keeping the temperature fixed, they perform one or more rounds of ramping the temperature from a high initial value (which encourages “hill climbing” to break free of high-energy local minima) to a low final value (which encourages “drilling down” to find deep local minima).

Although Monte Carlo methods are easily applied to a broad range of problems, one can conceive of cases in which one could devise a more efficient method of finding a solution that is tailored to the structure of the problem. There has been considerable research into developing specialized deterministic solvers, specialized heuristics, and specialized hybrid methods that combine deterministic steps with heuristic steps. Some of these will be discussed in greater detail when we examine the fixed-backbone peptide design problem in Sect. 3.3.4.1.

3.2.3 *Software Approaches for Peptide Drug Modelling*

Peptides able to bind to targets of therapeutic interest may be designed in many ways. In the absence of software tools, the screening of random libraries of very short peptides can yield hits. However, the number of possible sequences for a peptide increases exponentially with peptide length, rapidly dwarfing the largest possible libraries. For example, a ten-residue linear peptide made from the 20 canonical amino acids has 20^{10} (or approximately 10^{13}) possible sequences, and each additional residue increases the sequence space by a factor of 20. Even the largest genetically encoded peptide libraries cannot represent all possible sequences with so large a sequence space. Biased semi-random libraries, rationally designed by limiting the allowed amino acid types at key positions, can help to reduce the sequence space, and this has been applied successfully to the creation of α -helical binding peptides (Arslan et al. 2010), but it is difficult to generalize this to peptides that bind in an arbitrary binding conformation. For this, computational tools are an asset.

One of the simplest computational approaches to peptide design is to mine databases of known structures for amino acid sequence motifs compatible with the amino acid composition and structure of a binding site on a target and to use this information either to produce new candidate sequences in silico that can be synthesized and tested in low throughput in the laboratory or to generate probability distributions for amino acids to inform library design (Vanhee et al. 2011). Although attractive due to low computational cost, sequence-based methods such as dTER-Men (Zhou et al. 2020) rely on the existence of suitable motifs matching a desired peptide binding mode amongst the known experimentally determined protein structures. Perhaps not surprisingly, some of the greatest success at creating new peptide ligands by sequence-based methods has been with proteins that bind peptides as part of their natural function and for which many natural peptide ligands are known, such as the peptide-binding domains of the anti-apoptotic factors Bfl-1 and Mcl-1

(Frappier et al. 2019) or the major histocompatibility complex (MHC) (Honeyman et al. 1998). A hybrid method called CLASSY uses sequence motifs and an empirical energy function based on amino acid pair statistics from known protein structures, optimizing sequences using a linear programming approach. While this begins to model the problem geometry more accurately and allows greater extrapolation beyond the sequence motifs observed in the Protein Data Bank, it has also been applied primarily to creation of α -helical peptides (Grigoryan et al. 2009). In general, methods that are heavily based on database mining extrapolate poorly to peptides that incorporate non-canonical amino acid building blocks, synthetic cross-linkers, or covalent linkages such as those used to cyclize peptides. It is frequently desirable to cyclize linear peptides from genetically encoded libraries or in silico screens of protein fragments in order to improve stability and reduce the entropic cost of binding, yet naive incorporation of terminal linkages can often destabilize the binding-competent conformation and hinder binding, the opposite of the desired effect (Joshi et al. 2017). To take full advantage of the benefits of the exotic chemical building blocks and connectivities possible when making synthetic peptides, computational methods that are better able to model the physics of peptide folding and binding are important.

In this chapter, we will mainly focus on software tools for modelling molecular interactions at atomic resolution. Many software tools are now available for drug-target interaction modelling, in silico screening, or rational design. Each tool takes different inputs and produces different outputs, allowing it to answer different types of questions; in addition, each tool employs different sets of simplifying assumptions, meaning that a drug designer must be aware of the realms in which each tool can make valid predictions and where those predictions cease to be accurate or useful. This section reviews several of the commonly used computational tools, the software packages implementing these tools, and the types of predictions that each tool is best suited to produce.

3.2.3.1 Molecular Dynamics (MD)

Molecular dynamics (MD) programs simulate the motion of individual atoms within molecular systems, using an approximate energy function or force field based on the interatomic forces between pairs of atoms. This provides mechanistic insight into binding or conformational transitions and, in some cases, can allow prediction of the kinetics of binding or heteropolymer folding. Although powerful, MD is an inherently computationally expensive method.

MD force fields typically consist of a number of terms each approximating a particular physical effect, such as hydrogen bonds, the van der Waals repulsion of atoms in close proximity, or the attraction or repulsion of atoms based on charge. Also included are terms representing covalent interactions. Typically, harmonic potentials are used to constrain bond lengths, and similarly simple potentials constrain bond angles and dihedral angles. In actuality, non-covalent and covalent interatomic interactions are quantum mechanical in nature, but in MD force fields, these

are crudely approximated by simple functions of interatomic distance. This limits the accuracy of these force fields but makes them computationally tractable for large molecular systems. Parameters in MD force fields are typically tuned to reproduce known bulk properties of fluids (Patel and Brooks 2004). Empirically, this approach has been found to yield force fields able to reproduce other emergent properties of complex molecular systems, such as the folding behaviour of small, fast-folding proteins in very long MD simulations. The most commonly used MD force fields are the *Chemistry at Harvard Molecular Mechanics* (CHARMM) family (Huang et al. 2017), which have been optimized for the simulation of proteins composed of canonical amino acids, and the *Assisted Model Building with Energy Refinement* (AMBER) family (Bogetti et al. 2020), which includes the *General AMBER Force Field* (GAFF) that is well suited for simulating small-molecule drugs and other exotic chemical entities. These force fields are open-source and in the public domain, and a range of open-source and commercial software packages implement them. Popular packages include *GROMACS* (free and open-source, released under the L-GPL licence), *Nanoscale Molecular Dynamics* or NAMD (free for non-commercial use and licenced for a fee for commercial use), and the *AMBER* simulation software (commercial software with discounted rates for academics) (Case et al. 2005; Phillips et al. 2005; Pronk et al. 2013).

In a typical MD trajectory, positions and velocity vectors are stored for every atom, and atomic motions are simulated in discrete time steps. At each time step, the simulator computes all forces that all of the atoms exert on one another using the molecular mechanics force field and updates the velocity vectors for each atom according to Eq. 3.8:

$$\vec{\Delta v} = \frac{\vec{F}}{m} \Delta t \quad (3.8)$$

In Eq. 3.8, $\vec{\Delta v}$ is the change in an atom's velocity vector, Δt is the time step, \vec{F} is the force on the atom, and m is the mass of the atom. The above can be derived from the Newtonian relation $\vec{F} = m\vec{a}$, where the acceleration, \vec{a} , is equal to $\vec{\Delta v} / \Delta t$ in the limit of the time step, Δt , being very small. The step in the trajectory is then finalized by updating the position of each atom by displacing it along its velocity vector:

$$\vec{x}_{\text{new}} = \vec{x}_{\text{old}} + \vec{v}\Delta t \quad (3.9)$$

This process is repeated million or billions of times to produce an MD *trajectory*. Small adjustments are often also performed at each step to correct for numerical error and to maintain constant temperature or conservation of energy. Since the smallest atomic motions are bond vibrations that occur on the timescale of femtoseconds, Δt is generally on the order of 10^{-15} s. This means that on the order of a million steps are needed to simulate a nanosecond, a billion steps to simulate a microsecond, and a trillion steps to simulate a millisecond. The theoretical scaling of a single time step is also less than ideal, with $O(N^2)$ calculations required for a system of N atoms. In practice, however, interatomic interactions beyond a cut-off

distance can be disregarded to a good approximation. Since the density of atoms in a given volume is more or less constant, with this approximation, each atom then “feels” a constant number of other atoms within the cut-off distance, and the interatomic force calculation becomes $O(N)$ for a system of N atoms. While the computation of an individual time step can be efficiently parallelized on CPUs or GPUs, with different cores computing different interatomic interactions and updating atomic positions, each time step must be completed before the next can be computed, preventing parallelization of the trajectory as a whole. Because so many time steps must be simulated, this means that despite the roughly linear scaling for each time step, MD is still a computationally expensive technique requiring long run-times even on powerful hardware.

Due to this inherent expense, MD is difficult to use for large-scale candidate screening or for making predictions about slow processes or large molecular systems. It is also poorly suited to simulating the formation or breaking of covalent bonds (including bonds to metal ions). While MD force fields provide reasonably accurate predictions about the bulk properties of a large molecular system, they are approximate and should be viewed with suspicion on very small scales. Ultimately, this means that MD can be a useful complement to faster techniques for final computational validation of the kinetic stability of a peptide-target protein complex, or as an early step in the drug discovery process to identify flexible regions of a target protein that could complicate drug binding, but cannot be used for screening large numbers of candidates, or for interrogating the finest details of binding mode.

Finally, while MD can yield information about the *kinetic stability* of a complex, it is more poorly suited to predicting thermodynamic stability or binding affinity. This is because estimation of equilibrium properties depends on having samples that are representative of the thermodynamic distribution of states, and conventional MD can be practically nonergodic due to high potential barriers between wells in the energy landscape that prevent sampling of relevant states in reasonable computing time. Certain enhanced MD methods have been developed to try to achieve ergodicity in tractable computing time in order to address this (Bernardi et al. 2015). In cases in which there is a known reaction coordinate of interest, umbrella sampling methods, in which many simulations constrained to sample the conformational space at intervals along the reaction coordinate are performed and combined, can provide estimates of binding free energies and equilibrium constants (Torrie and Valleau 1977; Kästner 2011). When the reaction coordinate is not known, metadynamic approaches, in which the force field is altered with additional terms that drive the system away from states already explored, permit kinetic barriers to be surmounted more quickly and more extensive searches of conformation space to be performed in less time. While these are ordinarily used for qualitative exploration to find new low-energy wells, there are means of analyzing these biased trajectories to estimate the true thermodynamic distributions, again permitting computation of binding affinities (Limongelli et al. 2013). Replica exchange approaches provide an alternative that does not depend on altering the potential but instead relies on carrying out a set of parallel, independent simulations at different temperatures and permitting periodic swaps of temperature. High-temperature parts of trajectories allow broader

exploration of the conformational space by permitting kinetic barriers to be surmounted, while low-temperature parts permit very stable states to be found. Importantly, the temperature swapping can be performed in a way that ensures that the thermodynamic distribution of states is generated for any given temperature, again allowing thermodynamic properties like binding affinities to be estimated (Sugita and Okamoto 1999). Because replica exchange involves largely independent trajectories that communicate small amounts of information at broad intervals to allow temperature swaps, it is highly amenable to efficient parallelization across nodes on large computing clusters, contributing to its popularity (Eleftheriou et al. 2006). While these enhanced MD methods can bring binding affinity estimation into the realm of computational tractability, they remain computationally expensive methods for *in silico* screening of large numbers of candidate molecules to bind to a target.

3.2.3.2 Quantum Mechanics (QM)

MD force fields use simple, approximate potentials, structured for fast computation, to model the covalent and non-covalent interactions between atoms. However, these force fields only crudely approximate the quantum mechanical nature of these interactions. Quantum mechanics (QM) simulation packages are able to carry out much more accurate energy calculations, albeit at enormous computational expense. These applications find approximate solutions to the multi-electron Schrödinger equation, which cannot be solved exactly. Many approximation methods have been developed, including *Hartree-Fock methods* (Hartree 1928; Fock 1930; Slater 1930), *density functional theory* (DFT) (Kohn and Sham 1965), and *Møller-Plesset perturbation theory* (Møller and Plesset 1934). These typically show high polynomial scaling in computing time (ranging from $O(N^3)$ to $O(N^7)$ or higher) for systems of N electrons, making them very expensive for anything larger than a small molecule. (Note too that this is the scaling with the number of *valence electrons*, not the number of *atoms*. This means much steeper scaling as atoms are added to the system, and higher cost for heavier elements such as transition metals.) More approximate methods, such as the *fragment molecular orbital* (FMO) approach (Kitaura et al. 1999; Fedorov 2017), can scale to systems of thousands of atoms, but these methods also have an upper limit determined by computational expense. For the most part, although QM offers exquisite accuracy and precision over MD force fields, the computational cost precludes its use for systems of peptides bound to proteins, particularly if surrounding solvent is modelled explicitly. Additionally, although it permits very accurate calculation of enthalpies, QM poorly captures the entropic component of the Gibbs free energy, which, as shown in Eq. 3.5, is a major consideration for peptides binding to targets. However, QM calculations are useful early in the modelling process for pre-computing torsional potentials for individual amino acid building blocks (particularly for non-canonical amino acids, for which there are usually no experimentally determined structures from which to derive statistical potentials) and for precisely tuning building-block geometry (based on the assumption that an amino acid residue in the context of a larger peptide or protein

will have molecular geometry similar to that residue in isolation) (Drew et al. 2013; Mironov et al. 2019).

Commonly used QM simulation packages include the *General Atomic and Molecular Electronic Structure System* (GAMESS) (Alexeev et al. 2012), *NWChem* (Valiev et al. 2010), and *PSI4* (Parrish et al. 2017). All of these are either made available at no cost (GAMESS) or are free and open-source (NWChem, PSI4). Additionally, certain commercial drug discovery packages, such as the Jaguar and QSite applications in the Schrödinger software suite (Murphy et al. 2000; Bochevarov et al. 2013), provide QM methods as well as hybrid quantum mechanics/molecular mechanics (QM/MM) methods that can be used for local, high-accuracy QM calculations in the context of a many-atom MD simulation.

3.2.3.3 The Rosetta Software Suite

The Rosetta software suite is an extensive collection of algorithms and applications for heteropolymer modelling (Leaver-Fay et al. 2011). Originally developed for protein structure prediction, Rosetta has subsequently been expanded to allow protein design (Kuhlman et al. 2003; Koga et al. 2012), RNA and glycan structure prediction (Watkins et al. 2018; Roy Burman et al. 2020), heteropolymer docking (Chaudhury et al. 2011), small-molecule docking (Combs et al. 2013), symmetric structural modelling (DiMaio et al. 2011), protein therapeutic deimmunization (Yachnin et al. 2021), and many other applications. Rosetta stores structures in memory as *poses*, which contain all information about molecular geometry, kinematics, and energies. A given Rosetta protocol operates on a pose with a series of *movers*, which alter the pose in some way; *simple metrics*, which measure properties of the pose; and *filters*, which make decisions about the currently considered pose based on measured properties. Rosetta’s movers are based primarily around Monte Carlo and simulated annealing methods that progressively search the sequence and/or conformational space of a macromolecule or complex, often biased heavily by information mined from databases of macromolecules of known structure. The software suite’s protein structure prediction algorithms, for example, are largely based on “fragment insertion” moves, in which fragments of proteins of known structure are used to guide a Monte Carlo search of the conformation space of a fixed sequence of unknown structure (Simons et al. 1999). This reliance on known protein structures historically meant that Rosetta could not be applied readily to heteropolymer modelling problems involving more exotic chemical building blocks, for which few known structures exist. More recently, the development of more general sampling methods, like parametric design tools (Huang et al. 2014; Dang et al. 2017) and generalized kinematic closure methods (Bhardwaj et al. 2016) (described in detail in Sect. 3.3.3), have permitted more exotic modelling problems, involving synthetic heteropolymers built from non-natural chemical building blocks, to be tackled.

Like MD methods, Rosetta uses an energy function designed to be fast to compute and to update, at the expense of some accuracy. Historically, this combined

physics-based terms, such as a van der Waals potential with a functional form similar to that in AMBER or CHARMM, with statistical terms based on observed conformations of amino acids in Protein Data Bank (PDB) structures. Parameters in the energy function were tuned to reproduce experimentally solved protein structures (Leaver-Fay et al. 2011). The current version of the Rosetta energy function, REF2015, was re-parametrized using a combination of protein structure-based training and molecular dynamics simulations of fluids resembling amino acid side chains, with parameters adjusted to reproduce known bulk properties (freezing points, melting points, viscosities, etc.) of the various liquids (Park et al. 2016; Alford et al. 2017). Considerable effort also went into making the energy function better support mixtures of canonical and non-canonical amino acids, with particular focus on D-amino acids that are the mirror images of canonical L-amino acids (Alford et al. 2017). For these building-block types, the statistical potentials for the L-amino acids can be used with dihedral values mirrored (multiplied by -1). Accurately computing energies of more exotic non-canonical building blocks is more challenging. While physics-based parts of the REF2015 energy function (the van der Waals terms, electrostatic term, solvation terms, and hydrogen bond terms) can be used directly, statistical potentials such as the backbone Ramachandran potential, the side-chain Dunbrack potential, and the reference energies cannot. In cases in which a non-canonical amino acid closely resembles a canonical amino acid (e.g., halogenated derivatives of phenylalanine), the statistical potentials of the canonical amino acid can be “borrowed”. In cases in which the non-canonical side chain is more exotic, these potentials can be pre-computed using molecular mechanics or quantum mechanics calculations (see, e.g., the methods described in Mironov et al. (2019)). REF2015’s rama_prepro term allows additional pre-computed backbone potentials to be loaded (Mulligan et al. 2020), and its fa_dun term allows computed rotamer libraries with attached computed rotamer well depths and breadths to be loaded and used as a side-chain potential (Mulligan et al. 2021). Non-canonical rotamer libraries and potentials can be generated with Rosetta’s MakeRotLib application, which internally uses a general molecular mechanics force field (Renfrew et al. 2012).

The combination of these energy function improvements and the more general conformational sampling methods described above have allowed the design and computational validation of many well-folded peptide macrocycles. Beginning in 2016, these methods allowed the creation of peptides and polypeptides in the 15- to 50-amino acid size range, built from mixtures of canonical and non-canonical building blocks, that fold into rigid structures. These structures, which were confirmed experimentally by nuclear magnetic resonance (NMR) spectroscopy, were stabilized by disulphide bonds and/or by terminal amide bonds and by the intrinsic conformational preferences of D- and L-proline (Bhardwaj et al. 2016). In 2017, these methods were further advanced to allow the creation of small 7- to 14-amino acid macrocycles with rigid structures despite the lack of secondary structure (Hosseinzadeh et al. 2017), as well as a larger 60-amino acid macrocycle stabilized by a core linked by a three-way covalent hydrophobic cross-linker (Dang et al. 2017). By 2020, further refinement of these methods had permitted the design of

large panels of 8- to 24-amino acid macrocycles, built from mixtures of D- and L-amino acids, with exotic symmetries inaccessible to natural proteins. The structures of these were confirmed by X-ray crystallography. The largest of these was stabilized by a central, bound zinc ion and acted as a metal-dependent conformational switch, adopting an alternative rigid fold in the absence of the ion (Mulligan et al. 2020). In 2021, Rosetta's peptide macrocycle design methods were applied to the design of a macrocycle able to fit into and occlude the active site of the New Delhi metallo- β -lactamase 1 (NDM-1), an enzyme responsible for antibiotic resistance in certain pathogenic bacteria. The macrocycle was shown to fold into the designed structure and to bind to NDM-1 in the designed binding mode by X-ray crystallography (Mulligan et al. 2021). These methods have also permitted the creation of high-affinity inhibitors of histone deacetylases (HDACs) (Hosseinzadeh et al. 2021) and are being applied to other targets. Some of these macrocycles are shown in Fig. 3.2.

Rosetta's compiled executables and source code are made freely available for academic, not-for-profit, and governmental use and are licenced for a fee for commercial use. A community of several dozen academic laboratories maintains the software and works to expand its capabilities (Koehler Leman et al. 2020).

3.2.3.4 Other Macromolecular Modelling Software

Although this chapter will focus primarily on methods implemented in the Rosetta software suite, it is important to note that Rosetta is not the only software available for atomic-resolution modelling of peptide-protein complexes. Many tools exist for protein structure prediction and model refinement, including the *QUARK* (Xu and Zhang 2011, 2012), *I-TASSER* (Yang et al. 2015), and *RaptorX* (Källberg et al. 2012) tools. There are also other protein design tools, such as *Osprey* (Hallen et al. 2018), and these could conceivably be adapted for designing non-canonical peptides. Peptide design has also been carried out by iterated rounds of mutation and docking with docking software such as the *Molecular Operating Environment* (MOE) (Abe et al. 2007). However, due to the computational expense of docking and the vast combinatorial problem when designing sequences, this approach only works well for very short peptides with limited numbers of amino acid possibilities. Commercial drug discovery tools also exist, such as the Schrödinger software suite, which includes the macrocycles application for small-molecule macrocycle conformational sampling (Sindhikara et al. 2017), the Glide application for in silico small-molecule docking for drug screening (Halgren et al. 2004), and the MacroModel and Desmond tools for conformational sampling and scoring (Bhachoo and Beuming 2017). Recently, a combination of Schrödinger tools was used to develop a general protocol for designing macrocycles, including peptide macrocycles (Sindhikara et al. 2020), which may represent a commercial alternative to the publicly available Rosetta-based methods that will be the focus of this chapter.

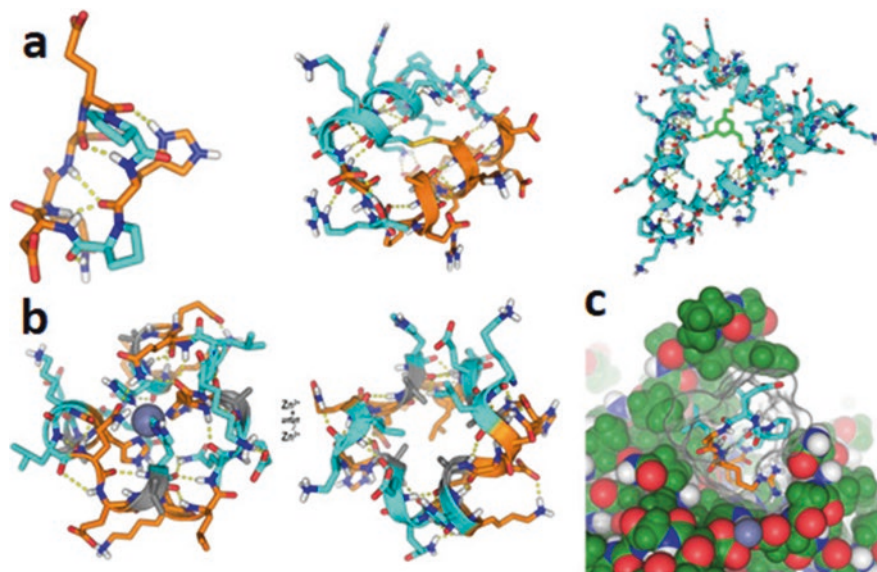


Fig. 3.2 Examples of computationally designed peptide macrocycles with experimentally validated structures, built from mixtures of L-amino acids (cyan), D-amino acids (orange), achiral non-canonical amino acids (grey), and exotic chemical cross-linkers (light green). Polar contacts including hydrogen bonds are shown as dashed yellow lines. **(a)** Peptide macrocycles with rigid structures spanning a broad range of sizes. From left to right, representative structures from the NMR ensemble for PDB structures 6BEW, 5KX0, and 5V2G, representing rigidly folded peptide macrocycles of 7, 26, and 60 amino acid residues, respectively. Smaller macrocycles tend to have structures resembling protein loops, stabilized primarily by the intrinsic conformational preferences of amino acids (particularly conformationally constrained amino acids like D- and L-proline). Midsized macrocycles can form secondary and tertiary structures, including exotic structures like the helices of opposite handedness that pack against one another in the 5KX0 structure. Larger macrocycles begin to have hydrophobic cores like those of proteins. **(b)** A designed peptide macrocycle with internal S4 symmetry, which binds a zinc ion as a central structural element (left). In the absence of zinc, the macrocycle inverts its conformation, exposing polar residues that had coordinated the zinc and burying hydrophobic residues to form a protein-like hydrophobic core (right). Such designed conformational dynamics could one day facilitate membrane permeability or other desired functions. PDB crystal structures 6UFA and 6UF9 are shown. **(c)** An 8-residue peptide macrocycle inhibitor of the New Delhi metallo- β -lactamase 1 (NDM-1). The macrocycle is shown as sticks and surfaces, bound to the NDM-1 target, which is shown as spheres. This peptide was designed both to make favourable contacts with the target and to fold rigidly into the binding-competent conformation (with high shape complementarity to the target binding pocket) in the absence of the target

3.3 Pipelines for Peptide Macrocycle Drug Design

To design peptide macrocycles intended to bind to a target of therapeutic interest, one must first obtain or generate a reliable model of the target and of a small-molecule “stub” bound to a functional site. This stub is then extended to produce a model of a macrocycle backbone in the functional site on the target, and diverse

conformations of the macrocycle are sampled. For each conformation, a sequence is designed in a guided process. Once a large pool of designs has been generated, each is subjected to low-cost filters to discard poor designs and then high-cost conformational sampling simulations to identify designs that stably fold into the binding-competent conformation in isolation. These high-cost simulations permit designs to be ranked to prioritize chemical syntheses and experimental evaluation. This pipeline, which proceeds from considering many possibilities at low computational cost per design to considering few possibilities at high computational cost per design, will be illustrated using the example of the previously published inhibitors of NDM-1 (Mulligan et al. 2021).

3.3.1 Target Modelling

Computational drug design approaches rely on having accurate structural and functional information about the target protein. Since the success of a drug development project depends not only on the methods used to design the drug but on the data that inform the design, care must be taken at the outset to obtain reliable starting data. Ideally, high-resolution (approximately 2 Å resolution or better) X-ray crystal structures are preferred as a starting point for computational design, and these must be analyzed to identify functional sites (such as enzyme active sites or binding interfaces) that can be modulated with a suitably designed binding molecule. The design of the first-generation HIV protease inhibitor saquinavir, for example, was based on high-resolution crystal structures of the HIV-1 protease (Ghosh and Gemma 2015). Peptide macrocycle inhibitors of the NDM-1 antibiotic resistance factor was designed rationally using the X-ray crystal structure of a small-molecule inhibitor bound to the NDM-1 active site (Mulligan et al. 2021). In both cases, the mechanism of action of these inhibitors is simple: They occlude the enzyme active site, blocking its function. In both cases, the active site was easily identified, though this is not the case for all functional sites in all potential protein targets. A binder can also modulate a protein's function through an allosteric mechanism, in which binding of a molecule to one site on a protein causes a cascade of small conformational changes that alter the function of a distant site – for example, enhancing or reducing catalysis at an enzyme active site or increasing or decreasing affinity for another molecule at a distal binding site. Long after their discovery, some commonly used drugs were found to exert their effects allosterically: The barbiturates, for example, modulate the effect of GABA_A receptors by increasing the duration of openings of an ion channel distinct from the barbiturate binding site (Leeb-Lundberg et al. 1980; Löscher and Rogawski 2012). However, rational design of allosteric modulators is challenging and fraught with risk: Errors in modelling the binding *or* in modelling the propagation of the effects of binding through the protein can result in a failure. The poorer the quality of the target structure, the more likely these failures grow.

Sometimes, X-ray crystal structures are of poorer resolution or fail to resolve key functional parts of a target protein, such as flexible loops. In other cases, a protein may fail to crystallize outright, necessitating other structural techniques such as nuclear magnetic resonance (NMR) spectroscopy or cryo-electron microscopy (cryo-EM). Although major advancements are being made in both techniques (Tan and Carragher 2020; Alderson and Kay 2021), rational drug design projects based on either can be riskier. NMR experiments yield information about interatomic distances that can be used to generate an ensemble of predicted structures, and although such ensembles may contain valuable information about dynamic motions of a protein, an NMR structure is rarely as precisely defined as a high-resolution X-ray crystal structure. Cryo-EM relies on thousands of separate images of protein molecules to construct an ensemble average. Historically, cryo-EM structures were severely limited in resolution, though recent advancements in microscope detector hardware and in image processing software have resulted in a “resolution revolution” that is beginning to allow this technique to compete with X-ray crystallography.

When experimental data are limited (e.g., of poor resolution) or incomplete (e.g., missing regions), computational structure prediction techniques can help to produce a high-quality starting model for drug design. MD-based structural refinement methods have been shown to improve models (Heo et al. 2019), as have physical restraints-based methods of refining X-ray crystal structures (Headd et al. 2012). In the Rosetta software suite, the CS-Rosetta method has been used successfully to generate high-quality models from limited NMR chemical shift data (Lange et al. 2012; Nerli and Sgourakis 2019), and scripted protocols have also been developed to permit model refinement guided by limited or low-resolution electron density from X-ray crystallography or cryo-EM experiments (Wang et al. 2016). Loop modelling tools, such as Rosetta Remodel (Huang et al. 2011), Rosetta LoopModel (Mandell et al. 2009), or Rosetta Next-Generation-KIC (Stein and Kortemme 2013), also permit predictions of regions of proteins that cannot be resolved experimentally. By combining data from many experiment types with computational prediction, one can improve models further (Leelananda and Lindert 2020). Importantly, all of these approaches rely on having some structural information from experiments, and it is most reliable to rely more heavily on experimental data and less on computational predictions.

In the absence of any experimental data about the target protein, known structures of homologous proteins can be used with homology modelling programs, such as Rosetta-CM (Song et al. 2013), I-TASSER (Zhang et al. 2018), or RaptorX (Källberg et al. 2012), to generate a homology model of the target protein. Although it is hoped that fully de novo structure prediction from amino acid sequence alone (i.e., with no structural data from experiments) could one day produce models of sufficient accuracy and precision to permit drug design, this approach is not yet commonly relied upon for drug development. Nevertheless, recent advancements in machine learning (discussed in Sect. 3.4.1, at the end of this chapter) are beginning to improve the prospects of drug design based on predicted structures.

A Target Protein Structure Can Come From

1. High-resolution x-ray crystal structures (preferred).
2. NMR or cryo-EM structures.
3. Low-resolution or incomplete experimental data bolstered by computational refinement.
4. Homology models.
5. Aspirationally, de novo computational structure prediction from amino acid sequence alone (though this is not yet reliable enough for common use).

3.3.2 Initial Stub Identification

As will be seen in the next section, current peptide macrocycle design approaches rely on extending a starting anchor or stub to build the macrocycle bound to the target. This means that a drug designer needs not only a reliable model of the target protein but of the target protein in complex with a starting stub. The starting stub is ideally a one- or two-residue peptide bound in a functional site of the target (such as the active site of an enzyme that one wishes to inhibit) but can also be a small molecule resembling an amino acid or other chemical building-block that can be incorporated into a peptide. Generally, this is a molecule that lacks the specificity or affinity to be a good drug on its own, but which could plausibly lose off-target interactions or gain target affinity when incorporated into a peptide that is able to make additional favourable interactions with the target.

In the best case, a high-resolution X-ray crystal structure (or other reliable experimentally determined structure) exists or may be obtained of a known small-molecule binder in complex with the target. This was the case for the peptide macrocycle inhibitors of NDM-1 that were designed previously: The X-ray crystal structure of L-captopril, a weak and nonspecific inhibitor, bound to NDM-1 had been solved (King et al. 2012). Fortuitously, this molecule resembles a dipeptide: By replacing a methyl group with an amine in the crystal structure of L-captopril bound to NDM-1, the molecule was converted to a D-cysteine-L-proline dipeptide that could be extended with additional amino acid residues. By extending the dipeptide to 8 residues and cyclizing it, it was possible to engineer additional favourable contacts with NDM-1 and to achieve higher-affinity binding (Mulligan et al. 2021). This approach is a general one that could convert a nonspecific small molecule that would not be usable as a drug due to off-target effects into a highly specific peptide therapeutic with few off-target effects.

In other cases, one may not be fortunate enough to have a small molecule as a starting point. Sometimes, however, there is a known protein interaction partner of one's target, particularly if one wishes to disrupt a protein-protein interaction surface. If a reliable experimentally determined structure exists of the two proteins in the complex, a drug designer may select one or two key interacting residues from the interaction partner as a starting stub for peptide macrocycle design.

If no experimentally determined structure of one's target in complex with another molecule exists, a drug designer may need to fall back on computational docking approaches. Tools like Rosetta's GALigandDock module can place single amino acid in a binding site on a target protein and can generate a pool of docked configurations as starting points for macrocycle design. Alternatively, MD simulations can also be used to generate pools of docked conformations of a single amino acid. While these approaches carry a higher risk of failure than relying on experimentally determined structures of a bound small molecule, predictions from docking trajectories or MD simulations using the known structure of a target protein in isolation are more reliable than, for example, full de novo predictions of a target protein structure – particularly when the goal is to identify many plausible weak binding modes for an amino acid that can subsequently be stabilized by other interactions between the peptide that presents the amino acid and the protein target.

An Initial Stub Can Come From

1. An experimentally determined structure of a small molecule or short peptide bound to a functional site on a target protein.
2. An experimentally determined structure of another protein interacting with a functional site on a target protein, from which one or two amino acids involved in the interaction may be cut out.
3. A docking algorithm or molecular dynamics simulation that places an amino acid or short peptide in a functional site on a target protein.

3.3.3 *Sampling Peptide Conformation*

Given a stub of one or two amino acids docked to a functional site on the target in a plausible docked binding orientation, the next steps are to extend the stub to produce a peptide of a desired length, to cyclize the resulting peptide, and to sample plausible conformations of the resulting macrocycle. Stub extension is best accomplished using an amino acid type with conformational preferences that represent the full set of building blocks that one wishes to consider during design, since sampling approaches are typically biased by the conformational preferences of this “stand-in” type (which will later be replaced by the types chosen when designing the sequence). If designing with only the 20 canonical amino acids (19 of which are L-amino acids), L-alanine tends to be a convenient choice since its accessible backbone conformations are similar to those of the other L-amino acids, and its side chain is minimal yet representative of the methyl or methylene group present in all L-amino acids. On the other hand, if designing with both L- and D-amino acids, glycine is a better choice for stub extension since it can access regions of backbone conformational space that are accessible to both L- and D-amino acids.

Once the stub has been extended on both its N- and C-termini with a length of alanine or glycine residues, the next challenge is to sample the conformations that

are “closed” – that is, those that are compatible with an amide bond connecting the termini (or another suitable covalent connection, such as a side-chain disulphide between terminal and near-terminal residues). Since the accessible conformation space grows exponentially with the length of the macrocycle ($O(C^N)$ scaling for an N residue macrocycle), fast methods are needed, and there will inevitably be an upper limit to the size of macrocycle that can be sampled regardless the speed of the method. Although this sampling could be accomplished with MD simulations, this is an approach that could require many core-hours of computing time per conformation sampled. Another approach would be to randomize the conformations of the extended chain, apply a harmonic potential to the termini (essentially, a “rubber band” or “elastic” drawing the termini together), and use gradient-descent minimization to “pull” the termini into an appropriate conformation to form a terminal amide bond. This would be faster but would still likely cost several core-seconds per sampled conformation and could introduce unknown bias into the sampled conformations. Since conformational sampling is an early step in the peptide design process, and one at which many hundreds of thousands or millions of conformations may be considered and most rejected, speed is of the essence at this step.

One of the fastest approaches is to use a kinematic closure algorithm borrowed from the robotics field to sample closed conformations in microseconds to milliseconds per conformation. Kinematic closure algorithms allow a user to set, perturb, or randomize most degrees of freedom in a chain and then to solve for the values of the remaining degrees of freedom to provide a desired rigid-body transformation from the start to the end of the chain (Coutsias et al. 2004; Mandell et al. 2009). When applied to peptide macrocycle conformational sampling, this means that the conformations of the amino acids added to the N- and C-termini of the stub can be randomized biased by the conformational preferences of their amino acid type, the interatomic distance and bond angles for the bond connecting the ends of the chains can be set to the ideal values for an amide bond (or other chemical bond type), and the values of six dihedral angles, flanking three pivot atoms, can be solved for analytically to produce a closed conformation, as shown in Fig. 3.3. In the Rosetta software suite, this is performed with a Rosetta mover called *GeneralizedKIC* (short for *generalized kinematic closure*) (Bhardwaj et al. 2016). When using *GeneralizedKIC*, a user must specify a selection of amino acid residues defining a closed chain of atoms; a set of rules for randomizing, perturbing, or setting degrees of freedom in the chain; and three pivot atoms. In addition, since each attempt can yield 0 to 16 solutions for the six torsion angles flanking the pivot atoms, the user may specify filtering rules to discard poor closure solutions, and must specify a selection rule to pick a “best” solution from those that remain. Filtration of solutions is typically done based on clashes with the target (i.e., rejecting those conformations in which backbone atoms overlap with one another or with the target) or number of internal hydrogen bonds within the peptide (i.e., rejecting those conformations with fewer than, e.g., two or three internal hydrogen bonds to stabilize the binding-competent conformation). In addition, a user may specify that some protocol be applied to each solution before a “best” solution is chosen based on computed energy. This permits a round of rapid fixed-backbone sequence design (see Sect. 3.3.4.1) to be

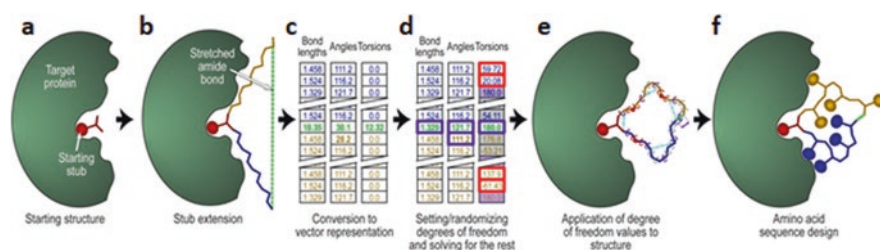


Fig. 3.3 Use of generalized kinematic closure (GeneralizedKIC) for efficient sampling of bound peptide conformations and docked orientations during peptide macrocycle design. (a) The starting structure. A structure of a one- or two-amino acid “stub” (red) bound in a binding pocket of the protein target (green) must be prepared. (b) Polyglycine chains are appended and prepended (brown and blue), and their termini are connected with an amide bond (dashed green line). The initial conformation of these chains results in a terminal amide bond with impossibly distorted geometry. (c) The portion of the peptide macrocycle that excludes the stub (brown, green, and blue) is converted to a vector representation, in which degrees of freedom (bond lengths, bond angles, and torsion angles) are listed in three vectors. At this point, the bond lengths, bond angles, and torsion angles of the terminal amide bond are all unrealistic (boldface). (d) Degrees of freedom are altered to close the loop. Specifically, the bond length, bond angles, and torsion angle for the terminal amide bond are set to ideal values (purple boxes), other amide torsion angles are set to 180° (purple shading), and all but six backbone ϕ and ψ torsion angles are randomized biased by the Ramachandran preferences of glycine (grey shading). Once this is done, a set of kinematic equations is solved analytically to produce the values of the six remaining backbone torsion angles (red boxes) in order to restore the rigid-body transform from the start of the loop (C-terminus of the stub) to the end of the loop (N-terminus of the stub). Each attempt with a given set of randomized values yields 0 to 16 solutions. (e) The vector representation is used to set degrees of freedom of the original structure, building from N-terminus to C-terminus. All backbone dihedral angles (purple, grey, and red curved arrows, representing the torsion values that were set, randomized, or solved for) are set, as well as bond angles and the bond length (purple) of the terminal amide bond (green). Solutions with fewer than a threshold number of backbone hydrogen bonds (dashed cyan lines) are discarded. (f) For each conformation remaining, an amino acid sequence is designed using the methods described in Sect. 3.3.4. The lowest-energy solution is returned as the result from a single kinematic closure attempt. Typically, the process is repeated thousands of times to produce thousands of designs

carried out for each solution, and the solution that presents amino acid side chains in the manner that maximizes favourable interactions with the target to be chosen.

Once an initial candidate conformation has been picked, this conformation can be further refined in an automated manner. One approach that has proven successful is to carry out a Monte Carlo search of the conformation space, in which each move involves perturbing the macrocycle slightly and re-closing with GeneralizedKIC, and possibly carrying out a fast round of sequence re-optimization before accepting or rejecting the move. Such an approach can yield backbone conformations that present amino acid side chains that are able to form very high-affinity interactions with the target. A protocol that implemented this approach for the NDM-1 target is examined in detail in Sect. 3.3.4.5.

3.3.4 Designing Peptide Sequence

Given a candidate conformation for a peptide macrocycle in a binding site on a target protein, the next step is to find a suitable amino acid sequence for that macrocycle. The sequence must be chosen to satisfy several criteria.

Criteria That a Good Peptide Sequence Must Satisfy

1. The sequence must create favourable interactions between target and peptide while avoiding unfavourable structural features that would destabilize the bound structure.
2. The sequence must be one that is amenable to chemical synthesis.
3. The sequence must have favourable physical properties (e.g., solubility) or pharmacological properties, allowing it to persist in the body long enough to find its target.
4. The sequence must stabilize the binding-competent conformation of the peptide.

The fourth criterion is needed since a high entropic cost is associated with ordering a flexible molecule on binding. As discussed in Sect. 3.1, this cost can shift the binding equilibrium heavily towards the unbound state (i.e., reduce the peptide's affinity for its target). This section discusses approaches for finding sequences that satisfy all of these criteria.

3.3.4.1 Approaches to the Fixed-Backbone Design Problem

The fixed-backbone design problem may be stated formally as follows: Given a candidate polypeptide backbone conformation with one or more chains, a selection of N amino acid positions (numbered 1 through N) at which side-chain identities and/or conformations are mutable, and a set of D_i discrete candidates for the side-chain identity and conformation (collectively called the *rotamer*) at the i^{th} position, one wishes to choose one rotamer per position such that some objective function E is minimized. In the particular case of designing a peptide macrocycle bound to a target protein, all of the residues in the macrocycle plus the residues in the target that are closest to the macrocycle would be part of the design problem. The rotamers considered for the macrocycle residues would include many discrete conformations for each of the possible amino acid building blocks that one had available for use in chemical synthesis. The rotamers considered for any given position in the target, on the other hand, would consist of discrete conformations of the particular amino acid type present at that position (i.e., only the side-chain conformation, and not the identity, of the residue would be permitted to change).

The objective function E scores any given selection of one rotamer at each of the N variable positions. Traditionally, this function was a molecular mechanics-style

energy function, with the property of *residue-level pairwise decomposability*. This meant that E could be expressed as a sum of one-body energy terms (based solely on a single amino acid residue) and two-body interaction energy terms (based solely on a pair of interacting amino acid residues), as shown in Eq. 3.10:

$$E(\vec{s}) = \sum_{i=1}^N \alpha_i(s_i) + \sum_{j=2}^N \sum_{k=1}^{j-1} \beta_{j,k}(s_j, s_k) \quad (3.10)$$

In the above, \vec{s} is the current rotamer selection vector listing the index of the selected rotamer at each of the N positions, $\alpha_i(s_i)$ is the one-body energy of the i th sequence position given the rotamer selection at this position, and $\beta_{j,k}(s_j, s_k)$ is the interaction energy between the j th and k th sequence positions given the rotamers selected these positions. Each of these in turn may be a sum of many energy terms representing particular physical interactions (e.g., van der Waals interactions, charge-charge attraction or repulsion, hydrogen bonds, torsional potentials, etc.). If the objective function is pairwise-decomposable, then it is possible to pre-compute tables of one- and two-body energies. This pre-computation can greatly accelerate the algorithms that solve for the lowest- E selection of rotamers. (Variations that include higher-order terms, such as three- or four-body energies, are also possible. However, the time and memory required to pre-compute and store n -body energies for N sequence positions and D rotamers per position scales with $O((ND)^n)$. For typically sized problems, it is computationally tractable to pre-compute and store a linearly scaling list of one-body energies and a quadratically scaling matrix of two-body energies, but higher-order terms rapidly become infeasible due to cubic, quartic, or higher-polynomial scaling.)

Choosing one rotamer at each of the N variable positions to produce selection vector \vec{s} such that $E(\vec{s})$ is minimized is a challenging problem. The search space for such a problem scales with $O(D^N)$, where D is the geometric average of the D_i values. As such, exhaustive enumeration is not feasible. Moreover, the problem has been proven to be NP -complete, meaning that there is currently no known polynomial-time algorithm for solving it (Pierce and Winfree 2002). Exact solvers such as Toulbar2 (Traoré et al. 2013) and the DEE/A* algorithm implemented in the Osprey protein design software (Hallen et al. 2018) use iterative branch-and-bound and dead-end elimination approaches to identify and eliminate rotamer selections that cannot yield the optimal solution and to converge on the unique selection that does. These methods have the advantage of provably converging to the global optimum. However, a peptide designer is often more interested in diverse near-optimal sequences than in the single sequence that optimizes an approximate objective function. Recently, a provably optimal algorithm for generating diverse near-optimal sequences was demonstrated that addresses this need (Ruffini et al. 2021), and this also has been made available in the Toulbar2 software.

While powerful, these methods scale non-linearly as the number of rotamers grows. A typical peptide binder design problem must include rotamers for all allowed amino acid types for every sequence position in the peptide, as well as rotamers for the fixed amino acid types in the target within some threshold distance

of the peptide. This means that even when designing small peptides, problems can involve thousands to tens of thousands of total rotamers. Moreover, the number of rotamers is increased by the inclusion of non-canonical amino acids as allowed types in the peptide. This can make it infeasible to use exact solvers for many of these problems. Heuristic approaches, such as Rosetta's *packer* module (Leaver-Fay et al. 2011), are an alternative that can be faster, allowing larger problems to be tackled, at the expense of ceasing to offer a guarantee of converging to the solution yielding the very lowest value of $E(\vec{s})$. The Rosetta *packer* solves the rotamer optimization problem by a simulated annealing-based search, in which moves consist of randomly swapping one rotamer for another at a randomly chosen position, and each move is accepted or rejected based on the Metropolis criterion (Eq. 3.6). Note that during the search, the updates can be computed very quickly because the value of $E(\vec{s}_{\text{new}}) - E(\vec{s}_{\text{old}})$ depends only on the change in one-body energy of the position that has been altered and the change in the two-body interaction energies involving that position, not on the energies of the rest of the structure.

Rosetta's *packer* is invoked through the Rosetta *fixbb* application or the *PackRotamersMover*. The former provides a command-line interface for the *packer* alone, while the latter provides an interface accessible to the PyRosetta (Chaudhury et al. 2010) and RosettaScripts (Fleishman et al. 2011) scripting languages, allowing rotamer optimization or sequence design steps to be interspersed with other structure manipulation or filtering steps to construct an overall design protocol. In addition, the *packer* is used by other protocols, such as the *FastRelax* (Khatib et al. 2011; Maguire et al. 2021) and *FastDesign* (Bhardwaj et al. 2016) protocols described in the next section.

While the Rosetta *packer* uses a very general simulated annealing-based heuristic, there has been research into more specialized heuristics tailored to the structure of a fixed-backbone rotamer optimization problem. One recent approach, called *Unified Decomposition-Guided Variable Neighbour Search* (UGDVNS), combines a heuristic global search with a deterministic local search, providing provably optimal solutions in the local neighbourhood of a starting solution. There is evidence that this method can consistently find lower-energy solutions than the Rosetta *packer*, albeit at greater computational cost (Ouali et al. 2020). This method is available in the Toulbar2 software but has not yet incorporated into peptide design pipelines.

Note that the fixed-backbone rotamer optimization methods described here are all completely general: nothing limits them to the design of proteins built only from the 20 canonical amino acids. In Rosetta, the default behaviour of the *packer* is to allow rotamers of the 20 canonical amino acids, though the set of building blocks used for design can be expanded by invoking a Rosetta module called a *packer palette*. For example, a custom *packer palette* could be used to specify that the *packer* should allow rotamers of the 20 canonical amino acids plus their mirror-image D-amino acid forms. The *packer* can thereby be applied to the design of peptides built from any combination of canonical or non-canonical building blocks, so long as all possible building blocks at a given position preserve the backbone geometry. By using the Rosetta REF2015 energy function or another similar energy function,

it is possible to maximize favourable interactions across the peptide-protein interface, the first criterion needed in a good sequence. In Sect. 3.3.4.3, we will see how modification of the energy function can help to find sequences that satisfy the other criteria.

3.3.4.2 Approaches for Flexible-Backbone Design

Frequently, the initially sampled backbone conformation is close to, but not identical to, one that is compatible with a particular amino acid sequence. Additionally, the ideal side-chain conformations are often close to, but not identical to, the discrete side-chain rotamers that are sampled. Composite protocols that perform alternating rounds of fixed-backbone rotamer optimization steps and flexible-backbone gradient-descent energy-minimization steps are often useful to allow backbones to move to accommodate the amino acid side chains. In Rosetta, gradient-descent energy minimization is performed using a module called the *minimizer* (Leaver-Fay et al. 2011). The minimizer computes the derivatives of an objective function E (usually, the Rosetta REF2015 energy function or some modification thereon) with respect to degrees of freedom of motion of the macromolecular system being modelled, and then performs downhill line searches in the reverse gradient direction, progressively incrementing the degrees of freedom until a local minimum in the one-dimensional search is reached. At this point, a new gradient calculation is performed, followed by a new line search in the new direction. When a local minimum in the multidimensional conformational space defined by the movable degrees of freedom is reached, the minimizer stops. (Note that this describes basic gradient-descent minimization. In practice, this is made more efficient with quasi-Newtonian methods that approximate second derivative information to alter the search direction and shorten the search.) Typically, allowed degrees of freedom of motion are the torsion angles of rotatable bonds and the rigid-body translational and rotational degrees of freedom relating separate polypeptide chains (called *jumps* in Rosetta), though the minimizer is also capable of performing Cartesian minimization by computing derivatives of E with respect to x , y , and z coordinates of atoms directly and allowing every atom to move independently, provided that E contains terms to constrain bond lengths and bond angles (Conway et al. 2014). The effect of the minimizer is to shift a conformation from a starting state to the nearest local energy minimum.

The simplest way to combine fixed-backbone design with flexible-backbone gradient-descent minimization would be to alternate the two. If one does so, however, the fixed-backbone rotamer optimization solutions that are closest to optimal (i.e., the ones which, if energy-minimized, would give the lowest-energy end points) are often discarded in the first rotamer optimization step due to small imperfections (often small overlaps between atoms that result in very high energies). For this reason, the energy function used for rotamer optimization is often modified in early stages to soften repulsive interactions, either by effectively spreading atomic density over a broader area or by simply reducing the weight on the repulsive van der

Waals term in the energy function. The Rosetta FastRelax (for optimizing conformations with fixed sequences) and FastDesign (for variable-sequence design and conformational optimization) protocols both perform alternating rounds of fixed-backbone rotamer optimization and flexible-backbone gradient-descent minimization, ramping the repulsive van der Waals term in the Rosetta REF2015 energy function from a small fraction of its normal value to full strength over the course of many rounds (Khatib et al. 2011; Bhardwaj et al. 2016; Maguire et al. 2021). This promotes densely packed rotamer selections in early rounds that can be “rescued” by a small amount of backbone and/or side-chain motion during gradient-descent minimization in later rounds.

An alternative approach is to use a “minimization aware” algorithm. Broadly, these algorithms implicitly consider the effects of side-chain and/or backbone minimization in the setup of the combinatorial optimization problem, instead of relying on post hoc minimization. One of the simplest approaches is to pre-compute rotamer pair energies, not for the fixed rotamers but for pairs of rotamers with side-chain conformations minimized for each pair in the context of the fixed molecular geometry (backbone and non-moving side chains) (Gaillard et al. 2016). While this allows the off-rotamer side-chain conformations to be considered to some degree, it ignores the fact that the minimized side-chain conformation found with different interaction partners may be very different. That is, a particular minimized conformation for a side chain may not be mutually compatible with all of the surrounding rotamers ultimately selected (Bouchiba et al. 2021). The iMinDEE algorithm, implemented in the Osprey software, improves on this idea by treating each rotamer not as a discrete conformation but as a continuum of conformations defined by small ranges of side-chain dihedral angles. Rotamer pair energies are likewise treated as a range with minimum and maximum bounds, calculated with energy minimization/maximization within each range for each pair. These bounds on the pairwise energies are used in the dead-end elimination (DEE) search of rotamer combinations, and the bounds on solution energies are narrowed as solutions are pruned until a provably optimal solution that considers the effect of side-chain minimization results. This serves to reconcile the different minimized conformations of pairs of rotamers in isolation (Gainza et al. 2012). An extension of this is the dead-end elimination with perturbations (DEEPer) approach, also available in Osprey, which includes small backbone motions in the ensemble implicitly considered for each rotamer (Hallen et al. 2013). Since arbitrary backbone motions propagate down the backbone and alter the interactions of all downstream amino acid residues, DEEPer permits only local backbone motions, focusing on the shear and backrub motions identified previously, which have local effects limited to three and four contiguous residues, respectively (Davis et al. 2006; Smith and Kortemme 2008).

Like the fixed-backbone rotamer optimization algorithm, gradient-descent minimization algorithms are general and usable with canonical or non-canonical building blocks. This means that FastDesign can be applied to peptide design problems, in which it frequently improves the docked conformation and finds sequences with more favourable intermolecular contacts and peptide-protein interaction energies as compared to simple fixed-backbone design algorithms alone. Minimization aware

approaches that act on side chains, such as iMinDEE, are also general enough to apply to canonical or non-canonical building blocks. However, those minimization aware approaches that allow backbone flexibility, such as DEEPER, rely on backbone motion types like backrub and shear, which are dependent upon assumptions about the covalent geometry of the backbone. Although these motion types are applicable to any chain of canonical or non-canonical α -amino acids, inclusion of building blocks with additional backbone atoms, such as β - or γ -amino acids, can prevent the use of these methods. Future research could produce fully general methods compatible with backbone minimization however.

3.3.4.3 Controlling the Design Process

The process of designing peptide therapeutics is as much an art as it is a science. Human experience and intuition must guide the design process. Generally, the design process is an iterative one: a designer creates an initial pool of designs, carries out various *in silico* validation tests to evaluate the pool, and then adjusts parameters guiding the design process to address the problems observed. In addition to iteratively improving the design process to maximize the success rate, a designer will typically also wish to impose prior knowledge about the desired chemical composition of the peptide macrocycles to be produced, based on what may be feasibly synthesized, what is likely to have physical properties compatible with the intended use or assay methods, etc. One of the most direct forms of control is over the set of amino acid building blocks allowed during design.

As mentioned in Sect. 3.3.4.1, basic fixed-backbone design algorithms (and their flexible-backbone extensions) can be used with any combination of side-chain rotamers. The user must specify the allowed side-chain identities and the degree of discretization of side-chain conformations to construct the allowed rotamer set and to define the side-chain optimization problem. In Rosetta, side-chain identities are controlled by packer palettes and by task operations, which offer global and local (position-specific) control, respectively.

Globally, the set of allowed chemical building blocks is defined by the packer palette. This is analogous to an artist's palette, which may define the set of colours with which the artist will paint. Rosetta's default packer palette allows design with the 20 canonical amino acids, but expanded palettes may also be specified by the user with command-line flags ("`-packing:packer_palette:extra_base_type_file <filename>`", where `<filename>`; is an ASCII file containing a whitespace-separated list of additional chemical building blocks to include by default) or by defining custom packer palettes in the PyRosetta (Chaudhury et al. 2010) or RosettaScripts (Fleishman et al. 2011) scripting languages.

Locally, the set of allowed chemical building blocks is pruned in a position-specific manner using task operations. This is analogous to an artist saying, "By default, my palette contains red, yellow, and blue paint, but in this corner of the painting, I will restrict myself to red and yellow by prohibiting blue." In addition to producing solutions that better align with prior knowledge, pruning allowed residue

types using task operations reduces the average number of rotamers per position. Since the solution space for a fixed-backbone design problem scales as $O(D^N)$, where D is the geometric average number of rotamers per position and N is the number of positions, pruning can greatly reduce the size of the search space and the length of Monte Carlo search needed to find reasonable solutions. Multiple task operations may be applied, but their action is always prohibitive: If the first prohibits tyrosine and tryptophan at positions 4 and 6 and the second prohibits lysine and arginine at position 5 and 6, the effect of applying both is to prohibit tyrosine and tryptophan at position 4, lysine and arginine at position 5, and all four types at position 6. The residues to which a given task operation applies in turn may be controlled by a *residue selector*, which defines rules for selecting amino acid positions in a structure. Particular residue selectors may select based on burial (e.g., the *Layer selector*), backbone dihedral values (e.g., the *Phi selector*), user-specified indices (e.g., the *Index selector*), or many other criteria. As an example, the RosettaScripts XML code shown in Listing 3.1 (in Appendix B) defines a fixed-backbone design job in which canonical amino acids and their D-amino acid counterparts are allowed by default, but D- and L-amino acids are each prohibited at positions at which backbone geometry disfavours one or the other, polar amino acids are prohibited in the peptide core, and hydrophobic amino acids are prohibited on the surface.

By finding sequences that minimize computed energy, the fixed- and flexible-backbone design algorithms reviewed in the previous section help to satisfy the first of the four criteria discussed in Section 3.3.4: by optimizing energy, one finds sequences that maximize favourable interactions between peptide and target and which minimize unfavourable interactions. However, the other three criteria (sequences must be amenable to chemical synthesis, sequences must possess favourable physical and pharmacological properties, and sequences must uniquely stabilize the binding-competent conformation of the peptide) are poorly satisfied by this approach. The energy function contains no information about the ease with which a molecule may be synthesized or about pharmacological properties. And although one would expect low-energy sequences to be stable in the binding-competent conformation, there is no guarantee that this conformation is *uniquely* stabilized: by optimizing energies of the bound complex without considering alternative conformations of the peptide, it is possible that one creates a peptide with many low-energy conformations or with a unique, lowest-energy conformation entirely different than the designed conformation. In addition to these problems, certain types of favourable or unfavourable interactions between peptide and target are poorly described by a pairwise-decomposable energy function, meaning that even the first criterion, which a sequence must maximize favourable interactions between peptide and target, may not be fully satisfied.

How may one find solutions to a design problem that satisfy multiple objectives? One approach is to augment the energy function used during the design process with additional, non-energetic terms that score the extent to which additional objectives are satisfied (Hosseinzadeh et al. 2017). Since the augmented scoring function is now a weighted sum of functions favouring different objectives (Eq. 3.11), this converts the multi-objective optimization problem into a single-objective

optimization problem that can be solved by existing fixed-backbone rotamer optimization algorithms.

$$E'(\bar{s}) = E(\bar{s}) + \lambda_1 f_1(\bar{s}) + \lambda_2 f_2(\bar{s}) + \lambda_3 f_3(\bar{s}) + \dots \quad (3.11)$$

In the above, $E'(\bar{s})$ is a combined objective function incorporating physical energies $E(\bar{s})$ as defined in Eq. 3.10) and non-physical, *design-centric guidance functions* $f_1(\bar{s})$, $f_2(\bar{s})$, $f_3(\bar{s})$, etc.). User-specified constants (or Lagrange multipliers) λ_1 , λ_2 , λ_3 , etc. set the relative weight or importance of each of these additional guidance terms. Provided the additional terms are finite-valued, the simulated annealing search of sequence space remains ergodic in the limit of a very long search; however, in practice, these terms can make it unlikely that non-productive regions of sequence space are ever sampled in a reasonable amount of time. Since the goal is to find the best sequence that one can find in reasonable time, and not to generate statistics about the entire ensemble of possible sequences, this practical nonergodicity is a benefit rather than a detriment.

In Rosetta, several design-centric guidance functions are available to the user. Some of these violate the assumption of pairwise decomposability applied to the energy function. Despite this, most are heavily optimized for minimal slowdown. Some of the most useful guidance functions are summarized here.

Frequently, one has some prior knowledge about the desired amino acid makeup of a peptide. By imposing certain rules, such as “my peptide must have at least two L- or D-proline residues”, one can considerably increase the probability of designing a peptide that uniquely favours the binding-competent conformation, satisfying the fourth criterion for a good sequence. Rosetta’s default REF2015 energy function has a “reference energy” term (Alford et al. 2017), which assigns a constant penalty or bonus for each instance of each type of amino acid, which is intended to correct for the fact that certain amino acid side chains (such as tryptophan) are bulkier and more able to engage in favourable interactions than others: by adding a penalty of +2.26 kcal/mol for each tryptophan residue in a structure, the Rosetta packer is discouraged from including tryptophan everywhere that it possibly can. Similarly, by adding a bonus of -2.72 kcal/mol for each glutamate residue, the packer is encouraged to include glutamate (which otherwise tends to be excluded from designs). Unfortunately, a constant per-residue-type instance penalty or bonus results in a linear penalty as a function of amino acid counts of each type (Fig. 3.4a), meaning that this term alone promotes solutions that are entirely composed of the amino acid type with the most negative bonus, and all compositional diversity comes from the influence of other terms. Because the other terms define a very complicated function, it is extremely difficult to anticipate exactly what amino acid composition is favoured by the combination of the reference energy term and the others. Designers may attempt to control amino acid composition by manually altering the bonuses and penalties placed on each residue type, but this is difficult for the same reason that controlling the position of a marble in a marble maze (a toy consisting of a board with a rugged or pitted surface) by tipping the maze board is

difficult: Given a complicated non-linear potential (in this analogy, the marble maze board with its pits and ridges), it is hard to anticipate or control where the global minimum will lie by adding a linear function (the tilt of the board). Even though the combination of the non-linear and linear functions can favour a point in the middle, the linear function that one is controlling always favours an edge or a corner. For a peptide designer attempting to control amino acid composition by altering reference energies, the process can be a very frustrating one of trial and error in which small “tilts” (small changes to the reference energies) have large and discontinuous impacts on the position of the global minimum and the resulting amino acid composition. Moreover, some desired amino acid compositions may be impossible to favour with any choice of reference energy values. The dimensionality of the space

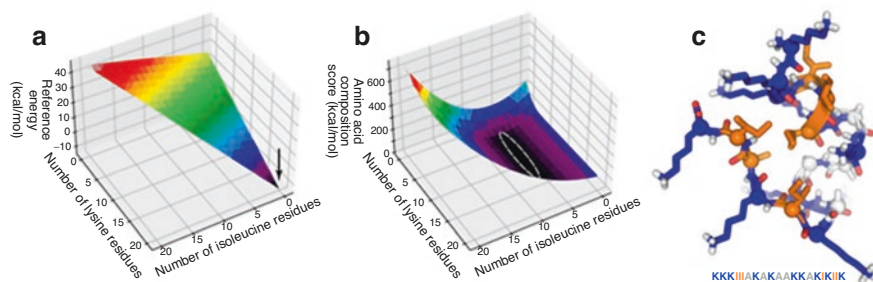


Fig. 3.4 Use of the `aa_composition` design-centric guidance scoring term to guide the design process to solutions with desired amino acid composition. The 20-residue trp cage mini-protein (PDB ID 1L2Y) is used as a sample scaffold for design in this example. **(a)** Effect of the reference energy term that is part of the default Rosetta energy function (REF2015). This term adds a constant value to the overall score for every instance of a particular amino acid type. Across all possible designed sequences, this results in a linear amino acid composition penalty as a function of the number of each allowed residue type, in which the lowest score is always achieved with 100% of the peptide being made of one residue type or another. In the example shown in Listing 3.3, in which allowed residue types are restricted to lysine, isoleucine, and alanine, the lowest reference energy score is with 100% lysine, 0% isoleucine, and 0% alanine (black arrow). Only the influence of other scoring terms (such as steric repulsion terms that might promote, e.g., the small side chain of alanine over the large side chains of lysine or isoleucine at certain tightly packed positions) can result in an intermediate amino acid composition, though the intermediate amino acid composition that is favoured is hard to anticipate and very hard to control with the linear reference energy term. **(b)** Effect of the `aa_composition` guidance term. Since `aa_composition` allows the user to specify quadratic penalties for deviation from a desired amino acid composition, lowest-scoring solutions can be at intermediate compositions. In the example shown in Listing 3.3, the amino acid composition constraints defined in Listing 3.2 define a quadratic penalty for greater than or fewer than six isoleucine residues and a quadratic penalty for fewer than four lysine residues. This means that solutions with exactly six isoleucine residues and at least four lysine residues are the most favourable (area surrounded by dashed white ellipse). When combined with other terms, the overall weight on the `aa_composition` term determines the extent to which deviation from the desired amino acid composition is tolerated. **(c)** Sample output from the script shown in Listing 3.3. Exactly six isoleucine residues (orange) are included though the influence of other scoring terms determines where they are best incorporated. The number of lysine residues (blue) is greater than four. Alanine residues are shown in white

compounds this: even given only the canonical amino acids, this is like a marble maze with a 19-dimensional board.

For this reason, the *aa_composition* design-centric guidance term was introduced. This term allows the user to define a desired amino acid composition and imposes a non-linearly ramping penalty, with a user-defined magnitude, to sequences that deviate from this desired composition. In addition to allowing direct control over the desired amino acid composition, this allows separate control over the strength of the constraint (i.e., the degree to which sequences that deviate slightly from the ideal composition are tolerated), which is not possible with a linear penalty. To return to the marble maze analogy, this is like pulling down on a particular point in the middle of a marble maze made of soft rubber to create a curved depression: If one pulls hard enough, it becomes trivial to get the marble to roll to this point, and the game is no longer hard at all.

As a simple example, let us imagine that we wished to design a peptide with only the amino acid types lysine, isoleucine, and alanine. Although this is only for illustrative purposes, it is worth noting that self-assembling peptides have been successfully designed with this minimal set of amino acid types (Boon et al. 2004; Frost et al. 2005), so this is not entirely absurd. We can prohibit all other residue types using task operations as before. However, we may also wish to impose the rule that final designs must contain at least 20% lysine and exactly six isoleucine residues. To this end, we may define a quadratically ramping penalty for fewer than 20% lysine residues, and for fewer or more than six isoleucine residues, using the composition constraints file shown in Listing 3.2 (in Appendix B). These constraints define the penalty function shown in Eq. 3.12. This function is graphed in Fig. 3.4b for an example case in which a small 20-amino acid mini-protein (the trp cage protein) is used as the backbone for which the sequence is to be designed.

$$E_{\text{aacomp,example}} = \begin{cases} \lambda_{\text{aacomp}} \left((4 - K)^2 + (6 - I)^2 \right), & K \leq 4 \\ \lambda_{\text{aacomp}} (6 - I)^2, & K > 4 \end{cases} \quad (3.12)$$

In Eq. 3.12, I and K represent the number of isoleucine and lysine residues, respectively, and λ_{aacomp} is the overall weight of the *aa_composition* scoring term. As a function of amino acid composition, the penalty in this example has a long, narrow minimum in which any solution with more than 20% (i.e., more than four) lysine residues is isoenergetic, but only solutions in which exactly six isoleucine residues are present are lowest scoring. Note that the particular functional form depends on the composition constraints applied and can vary considerably from example to example. In a RosettaScripts script, these constraints may be applied to a pose prior to a design run using a Rosetta mover that adds composition constraints, the *AddCompositionConstraintMover*, as shown in Listing 3.3 (in Appendix B).

When design is carried out using a scoring function that has the *aa_composition* term appended, an example of the output is shown in Figure 3.4c. Practically, for the design of peptide macrocycle therapeutics, it is often necessary to enforce the

presence of a minimum number of conformationally constrained amino acid types (L- or D-proline, 2-aminoisobutyric acid, etc.), a maximum of one of any residue type poorly amenable to chemical synthesis (L- or D-methionine, L- or D-arginine, etc.), and a minimum content of both polar amino acid types (for solubility) and hydrophobic types (for binding to a hydrophobic binding pocket). The weights on these constraints are generally tuned by an iterative guess-and-check process, in which a designer configures `aa_composition` constraints with his or her best estimate of the constraint weights needed, produces a pool of designs, evaluates the fraction of designs that pass visual inspection, low-cost automated filters (see Sect. 3.3.4.4), or high-cost validation (see Sect. 3.3.5), and forms hypotheses about the reasons for the failures. The `aa_composition` constraint weights are then adjusted based on these hypotheses. For example, a designer may upweight the proline constraint if he or she observes that too many designs are predicted not to fold and hypothesizes that this is due to too few proline residues appearing in designs. A new round of design is then performed. While more automated means of tuning constraint weights may be devised in the future, the complexity of peptide design and paucity of training data currently necessitates heavy involvement of a human capable of observing patterns and forming intelligent hypotheses about failure modes to guide the process.

Another common problem encountered during the design process is frequent occurrence of empty voids or holes in the peptide-protein interface. Ideally, a high-affinity binder should possess high shape complementarity to its target, binding in a conformation that leaves no empty spaces between the two. The presence of voids arises from the functional form of the Lennard-Jones term in many force fields, including the Rosetta REF2015 energy function. The Lennard-Jones term (divided into two terms, called `fa_rep` and `fa_atr` in Rosetta) approximates both the steric repulsion of atoms in close proximity and the effects of the London dispersion forces that attract widely spread atoms to one another. Typically, a 6–12 potential is used as shown in Eq. 3.13:

$$E_{LJ}(r_{i,j}) = w_{LJ} \left(\left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^{12} - 2 \left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^6 \right) \quad (3.13)$$

In the above, $r_{i,j}$ represents the distance between atoms i and j , $\sigma_{i,j}$ is a parameter that depends on the atom types of atoms i and j , and w_{LJ} is the overall weight of the term. A potential of this form asymptotically approaches zero as $r_{i,j}$ gets very large, dips to a small, negative value when the two atoms are separated by a contact distance equal to the sum of their van der Waals radii, and approaches infinity as $r_{i,j}$ goes to zero. Because the extreme scaling of the repulsive part of the potential (as atoms are pressed together) outweighs the mild cost of separating atoms beyond their contact distance, rotamer combinations that produce very small clashes (which could easily be resolved by a round of gradient-descent energy minimization) are often shunned by rotamer optimization algorithms in favour of rotamer combinations with large gaps or voids.

To correct for this, the *voids_penalty* design-centric guidance term was introduced. This term guides rotamer optimization algorithms to solutions with well-packed interfaces by penalizing solutions that are either under-packed (with empty voids) or overpacked (with clashes). The functional form of the *voids_penalty* term is shown in Eq. 3.14:

$$E_{\text{voids}} = w_{\text{voids}} \left(V_{\text{total}} - \sum_{i=1}^N v_i \right)^2 \quad (3.14)$$

Here, V_{total} is the volume to be filled by movable or designable side chains, $\sum_{i=1}^N v_i$ is the sum of the volumes of the currently selected rotamers at the N packable positions, and w_{voids} is the weight assigned to the *voids_penalty* term. To accelerate this calculation, both V_{total} and the volumes v_i of all rotamers are calculated in a pre-computation and cached for rapid lookup during the simulated annealing search of rotamer combinations as shown in Fig. 3.5. Since the difference is squared, the minimum is when $V_{\text{total}} = \sum_{i=1}^N v_i$ (i.e., when the currently selected rotamers have a total volume matching the volume to be filled). Such solutions can only have empty voids if they also possess atomic overlaps (clashes), but since these are heavily penalized by the Lennard-Jones term, the lowest-energy solutions are those that are well-packed without major clashes. Minor clashes can be subsequently relieved by gradient-descent energy minimization. Although frequently useful for interface design, the *voids_penalty* guidance term can also be used for conventional protein design to ensure that hydrophobic cores of designed proteins are well-packed.

Many other design-centric guidance functions exist that can help to eliminate observed pathologies in pools of designed peptide therapeutics and to improve overall design quality (Mulligan et al. 2021). Terms like the *buried_unsatisfied_penalty* penalize solutions in which hydrogen bond donors and acceptors buried in the interface are not satisfied (a common cause of a design failing in experiments) and promote fully satisfied solutions. The *hbnet* term, which can be useful for polar interfaces, promotes solutions with extensive hydrogen bond networks. The *netcharge* term promotes a desired net charge, which can be useful for, for example, limiting positive charges on a peptide drug to avoid cytotoxicity caused by membrane disruption. And the *aspartimide_penalty* term is useful for penalizing dipeptide subsequences that have a high likelihood of producing aspartimide by-products that can limit yield during chemical synthesis. The particular choice of which terms to use and how to configure them must be made on a case-by-case basis by a peptide macrocycle designer, but these tools provide many ways to fine-tune the design process to maximize the quality of designs produced. For a complete list of the design-centric guidance scoring terms available in Rosetta, see the online documentation at https://new.rosettacommons.org/docs/latest/rosetta_basics/scoring/design-guidance-terms.

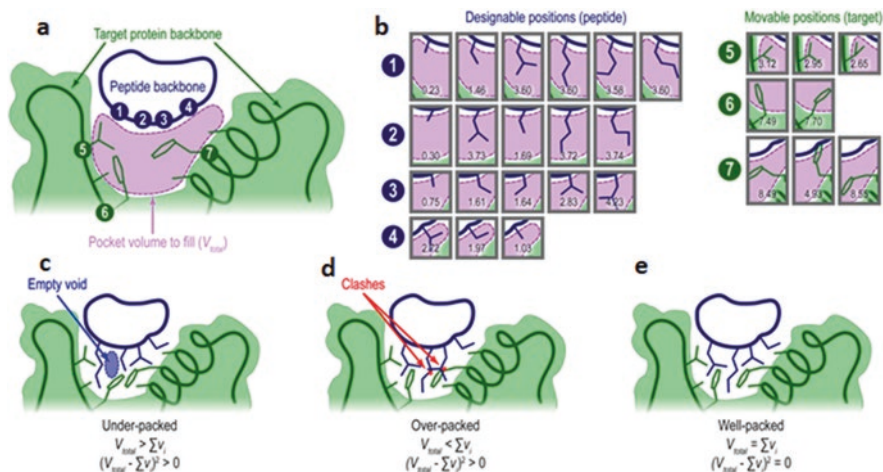


Fig. 3.5 Use of the `voids_penalty` design-centric guidance scoring term to produce designs with well-packed protein-peptide interfaces. In all panels, the protein target is shown in green, and the peptide macrocycle is shown in dark blue. (**a** and **b**) Pre-computation performed by the `voids_penalty` scoring term. For a given fixed-backbone conformation of the macrocycle, the `voids_penalty` term identifies the space, shown in pink, which is to be filled by designed peptide side chains or flexible protein side chains (panel **a**), and computes the volume, V_{total} , of this space. For every side-chain rotamer of the peptide (where both residue identity and conformation can vary) or target (where residue identity is fixed and only conformation varies), the volume of the side chain that lies within the volume to be filled is also pre-computed and stored (panel **b**). (**c** through **e**) Scoring of candidate solutions during the simulated annealing-based search of rotamer selections. At every step, the sum of the pre-computed volumes of currently selected rotamers ($\sum V_i$) is subtracted from the total volume (V_{total}), and the result squared. In cases in which there are empty voids (blue ellipse in panel **c**), the total volume is greater than the sum of the volumes of current rotamers. In cases in which there are clashes (red stars in panel **d**), the sum of the current rotamer volumes is greater than the total volume. Only when the solution is well-packed does V_{total} equal $\sum V_i$ so that the difference squared is a minimum (panel **e**). Since the packer optimizes the objective function, including the `voids_penalty` score guides the packer to solutions with well-packed peptide-protein interfaces

3.3.4.4 Filtering Out Poor Designs

Since peptide macrocycle design strategies are based on stochastic searches of both sequence and conformational space, and are scored with scoring functions that prioritize speed over accuracy, many design attempts will produce failures even with heavy guidance. For this reason, a designer produces a large pool of designs and then attempts to discard as many designs as possible to generate a short list of candidates that is highly enriched for successful designs to carry forward to experimental testing phases. In Rosetta, filters are used to examine poses and to accept or reject them. In addition to being useful for pruning designs from the final pool, filters are essential for aborting non-productive design trajectories at an early stage to avoid wasting computation on a design that is unlikely to be recoverable with more computation. Although some filters measure properties that are captured by the energy function or by a design-centric guidance function, filters are not intended to

be used for millions of evaluations during a rotamer optimization Monte Carlo trajectory; as such, they can use algorithms that are hundreds to thousands of times slower (costing core-seconds of computing time typically) but which can be considerably more accurate. This also means that certain properties can only be evaluated post hoc with filters for want of a means of evaluating them rapidly during the design process.

When designing peptides, common properties that one may wish to examine include the shape complementarity of the peptide to its binding pocket, the number of buried unsatisfied hydrogen bond donors and acceptors in the interface, and the number of hydrogen bond donors and acceptors that engage in non-physical numbers of hydrogen bonds. We will examine each of these in turn.

As discussed in Sect. 3.3.4.3, rotamer optimization algorithms often err to the side of promoting empty buried voids over small side-chain clashes. Even with design-centric guidance terms like the `voids_penalty` scoring term, some subset of designs may still have empty voids. When designing peptide-protein interfaces, it is equivalent to say that some subset of designs will have poor shape complementarity across the peptide-protein interface. Rosetta's *ShapeComplementarity* filter is useful for identifying and discarding these designs. It allows the user to define a minimum shape complementarity threshold (a number ranging from 0 to 1) and/or a minimum contact area (in square Ångstroms). The filter computes the molecular surface for the peptide and the target and then calculates a score based on the distance and the difference in normal vectors at the closest points across the interface as a means of scoring shape complementarity (Lawrence and Colman 1993). For each pair of points, the score is given by Eq. 3.15:

$$S(\vec{x}) = -(\vec{n} \cdot \vec{n}') e^{-w|\vec{x}-\vec{x}'|^2} \quad (3.15)$$

In the above, \vec{x} and \vec{x}' are a point on the molecular surface of one molecule and the nearest point on the molecular surface on the other molecule, respectively, \vec{n} and \vec{n}' are the normal vectors at these two points, and w is a fall-off weighting factor. Intuitively, if the normal vectors point in opposite directions and the points across the interface are close, then both the dot product and the exponential are nearly 1 and $S(\vec{x})$ approaches 1. Near perpendicular normals reduce the size of the dot product, and large distances between the surfaces reduce the value of the exponential, in either case resulting in a value of $S(\vec{x})$ near 0. The median value is computed for a sample of many points for each surface, and the two medians are averaged to produce a final score (Lawrence and Colman 1993). Because this slower but potentially more accurate approach is completely distinct from the method used by the `voids_penalty` term, it serves as a useful check on the output of design runs performed with that term. An example will be seen in Sect. 3.3.4.5.

Buried unsatisfied hydrogen bond donors and acceptors are another frequent pathology produced by rotamer optimization algorithms. In reality, buried unsatisfied hydrogen bond donors and acceptors are rarely seen in protein cores or

interfaces since they are highly destabilizing: a hydrogen-bonding group pays a price to be buried in the absence of a hydrogen-bonding partner if it could be satisfied if exposed to water. The presence of unsatisfied donors and acceptors in the output from rotamer-optimization algorithms results from the pairwise-decomposable nature of the energy function: hydrogen bonding terms typically award a negative score for the presence of a hydrogen bond and a score of zero for the absence of a hydrogen bond. Identifying the absence of a hydrogen bond in order to penalize it is fundamentally not a pairwise-decomposable problem: it is necessary to consider all possible hydrogen bonding partners in order to determine that none of them form a hydrogen bond with a particular group. Although non-pairwise design-centric scoring terms, such as the `buried_unsatisfied_penalty`, can help to discourage buried unsatisfied donors and acceptors, even with such terms, some designs will possess this undesirable feature. Rosetta's *BuriedUnsatHbonds* filter is useful for discarding designs with more than a user-specified threshold number of buried, unsatisfied hydrogen bond donors or acceptors.

A related problem is oversaturation of hydrogen bond acceptors. An oxygen atom can accept at most two hydrogen bonds. Unfortunately, since the pairwise-decomposable energy function awards a constant bonus for each additional hydrogen bond, rotamer optimization algorithms sometimes yield solutions with three or even four hydrogen bonds to a single oxygen atom. Rosetta's *OversaturatedHbondAcceptorFilter* can identify and discard designs with this pathology, and an example is shown in Sect. 3.3.4.5. An important caveat is that starting structures sometimes possess a small number of buried unsatisfied hydrogen bond donors and acceptors and occasionally one or two oversaturated acceptors. For this reason, the user must set the acceptable threshold for such filters on a case-by-case basis to catch only the new unsatisfied or oversaturated groups added by design steps. Again, both the *BuriedUnsatHbonds* filter and the *OversaturatedHbondAcceptorFilter* perform their calculations in a slower but more accurate manner distinct from that used by the `buried_unsatisfied_penalty`, potentially identifying cases missed by the faster method that is compatible with rotamer optimization.

Rosetta offers many additional filters that can be useful for peptide design. Designs with high overall scores or high `fa_rep` scores (indicating atomic clashes) are often discarded using the *ScoreType* filter for instance. For a full list of available filters, see the online documentation at https://new.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Filters/Filters-RosettaScripts.

3.3.4.5 Example of RosettaScripts Protocol

In this section, we will explore a full peptide macrocycle design protocol that was used previously to design eight-residue macrocyclic inhibitors of the NDM-1 antibiotic resistance factor (Mulligan et al. 2021). The full script is available from the Rosetta scripts repository, in the directory `Rosetta/main/rosetta_scripts_scripts/scripts/public/macrocycle_inhibitor_`

design/NDM1i-1_design_script/modernized/ in a standard installation of Rosetta. For instructive purposes, we will examine key pieces though the reader is encouraged to examine the full script to understand how the pieces fit together. Relevant excerpts from the NDM-1 inhibitor design script are shown in Listings 3.4 through 3.17 in Appendix B.

A RosettaScripts XML script is organized into sections. In earlier sections, instances of Rosetta modules of particular types (scoring functions, residue selectors, task operations, movers, etc.) are declared, assigned unique names, and configured. Once declared, these instances may be used in later steps, either as inputs to other modules (e.g., some movers can accept a mover that they apply at a particular stage of their own execution) or as steps applied in the PROTOCOLS section. The PROTOCOLS section, which occurs near the end of the script, lays out the overall sequence of steps defining the protocol. This basic layout is shown in Listing 3.4.

The input into the NDM-1 inhibitor design script is a model of NDM-1 with a dipeptide “stub”, consisting of D-cysteine-L-proline, bound in the active site. (This input is based on PDB ID 4EXS, which is the X-ray crystal structure of L-captopril, an NDM-1 inhibitor that resembles the D-cysteine-L-proline dipeptide, bound to the NDM-1 active site.) The output is a pool of designed structures, each consisting of an eight-residue peptide macrocycle with a designed sequence modelled in the NDM-1 active site. Let us first examine the PROTOCOLS section shown in Listing 3.5. Although located at the end of the script, this tends to be a good place to start since it defines the sequence of steps that will be executed. This section lists a series of 16 movers followed by two filtering steps. The movers, which are defined, configured, and assigned names in the earlier MOVERS section of the script, perform the steps of extending the dipeptide with six additional glycine residues (“extend”), setting up a new kinematic relationship for the resulting eight-residue peptide (“foldtree1”), initializing the backbone dihedral angles of the new chains (“initialize_tors”), and declaring a peptide bond between the N- and C-termini (“connect_termini”). Two special movers (“add_cutpoint_upper_pep” and “add_cutpoint_lower_pep”) are then used to annotate the starting and ending residues of the cutpoint in the macrocycle, which will allow subsequent energy-minimization steps following loop closure to preserve the peptide bond geometry. The next movers add global amino acid composition constraints to the peptide residues (“global_comp”), as well as local amino acid composition constraints to promote α -helix-favouring residue types at positions that are in the right- and left-handed helical regions of Ramachandran space (“L_alpha_comp” and “D_alpha_comp”, respectively), and local amino acid composition constraints to promote β -strand-favouring residue types at positions that are in the right- and left-handed strand regions of Ramachandran space (“L_beta_comp” and “D_beta_comp”, respectively). The next step is to close the macrocycle loop using the generalized kinematic closure algorithm (“genkic”), performing a fast-initial round of design on each of the closed solutions found before the best is selected. The best solution is further refined using a Monte Carlo protocol that perturbs and re-closes using generalized kinematic closure, refining the sequence at each step (“mc_search”). The best conformation and sequence from the Monte Carlo trajectory is redesigned a

final time, using the more expensive FastDesign protocol (“fdes”). The resulting structure is subjected to a final round of fixed-sequence conformational optimization and relaxation (“final_frlx”). The “connect termini” mover is applied a final time to correct the positions of the carbonyl oxygen and amide proton that depend on the peptide bond linking the termini, and the resulting structure is subjected to final filtering steps: a check that hydrogen bond acceptors are not accepting more than two hydrogen bonds (“oversat”) and a check that the shape complementarity between peptide and target is above a minimum cut-off (“shape3”). Those structures that pass these filters are written to disk.

Let us now examine the configuration of certain of the movers, in the `MOVERS` section of the script, prior to the `PROTOCOLS` section. Initial closure of the macrocycle is performed with an instance of the GeneralizedKIC mover that is assigned the name “genkic” as shown in Listing 3.6. The configuration of this mover closely follows the steps shown in Fig. 3.3. For the purposes of this example, residues 237 and 238 are the “stub”, which is not part of the loop to be closed. The loop to be closed, as defined by the `AddResidue` lines, runs forward from residues 239 through 241, passes through the peptide bond connecting residue 241 (the C-terminal residue) to residue 234 (the N-terminal residue), and then runs forward from residue 234 through residue 236. The six degrees of freedom that the kinematic closure algorithm will solve for analytically are the torsions flanking the three atoms defined in the `SetPivots` tag: the first is the C_{α} atom at the start of the loop (residue 239), the second is a C_{α} atom somewhere near the middle (residue 234), and the third is the C_{α} atom at the end of the loop (residue 236). Since the peptide bond joining residues 241 and 234 starts out in an impossibly stretched configuration, the `CloseBond` tag defines how its degrees of freedom must be updated to yield proper amide bond geometry, setting the bond length, bond angles, and dihedral (torsion) angle to ideal values. Finally, the perturbors defined in the `AddPerturber` block specify that all loop residues’ φ and ψ backbone dihedral angles should be randomized, biased by the Ramachandran map for glycine.

The settings in the `GeneralizedKIC` tag, at the top, specify that this mover will try up to 100 times to randomize backbone dihedrals and find a closed solution but will stop when at least one valid solution is found. If multiple solutions are found on a single attempt, the best solution will be chosen by energy (lowest_energy_selector). However, prior to choosing a solution, each candidate will undergo a series of steps laid out in a mover named “genkic_steps”. This mover is declared earlier in the `MOVERS` section of the script. As Listing 3.7 shows, the “genkic_steps” mover is a named instance of a “ParsedProtocol” mover, which wraps several previously defined mover and filter steps. These steps define a protocol that attempts to discard poor solutions using the minimum amount of computing power and which reserves computationally expensive design steps for those solutions that pass initial filters. The initial filters reject sampled backbone conformations with oversaturated hydrogen bond acceptors (“oversat”) and with fewer than a threshold number of backbone hydrogen bonds within the peptide (“total_hbonds”). They also reject solutions with a very high score of the Rosetta fa_rep scoring term – that is, those solutions in which part of the peptide backbone clashes egregiously with itself or with the

protein target. Listing 3.8 shows the earlier definition of some of these filters in the `FILTERS` section of the script. Next, the “genkic_steps” protocol carries out alternating rounds of design and minimization interspersed with additional filtering steps, beginning with relatively low-cost calls to the fixed-backbone `PackRotamersMover` (“softdesign” in Listing 3.7) and progressing to higher-cost calls to the flexible-backbone `FastDesign` protocol (“fdes” in Listing 3.7). As before, the “connect_termini” mover is reapplied periodically to correct placement of carbonyl oxygen and amide proton atoms that depend on the peptide bond connecting the termini. At the end of the protocol, oversaturation and shape complementarity are assessed, and designs failing these criteria are rejected. The pool of candidate closed solutions that the “genkic” mover chooses from when picking a best solution is the subset of candidates for which all filters listed in “genkic_steps” have passed.

We will not examine the declarations of all of the movers listed in the “genkic_steps” protocol, but the “softdesign”, “min1”, and “fdes” movers, defined previously in the `MOVERS` section, are worth examining in greater detail. First, the “softdesign” mover, the declaration of which is shown in Listing 3.9, is a named instance of a `PackRotamersMover`, which directly calls the Rosetta packer to perform optimization of discrete side-chain rotamers on a fixed backbone. The energy function used during this step is not the default Rosetta REF2015 energy function but a “softened” version that is more permissive of small clashes. This is a strategy frequently used during early rounds of design. Additionally, the “r15_soft” scoring function was configured in the `SCOREFXNS` section (not shown here) to include the `aspartimide_penalty` and `aa_composition` design-centric guidance terms (see Sect. 3.3.4.3).

The “softdesign” mover is configured with a packer palette (“design_palette”) that specifies that in addition to using the 19 canonical L-amino acids plus glycine, this design step will also use 18 of the 19 D-amino acids that are the mirror images of the canonical L-amino acids (omitting D-cysteine, which has the complication of forming disulphide bonds). The configuration of the packer palette is shown in Listing 3.10.

The allowed amino acids used by the “softdesign” mover are pruned in a position-specific manner using task operations. Task operations also provide additional configuration for the packer. In order, the task operations listed in Listing 3.9 specify that the conformation of the existing amino acid (if any) should be included as an allowed rotamer (“use_input_rotamer”), that target residue identities are immutable (“no_design_target”) and conformationally fixed (“no_repack_target”), that stub residue identities are immutable (“no_design_stub”), and that the D-cysteine part of the stub (which coordinates a metal ion) is conformationally fixed (“no_repack_stub_dcys”). The remaining task operations prune allowed residue types, specifying that cysteine and glycine are prohibited at all designable positions (“no_cys_gly”), that L- and D-amino acids are prohibited at non-buried backbone positions incompatible with either class of residue type (“D_design” and “L_design”, respectively), and that non-hydrophobic residue types should be prohibited at buried positions (“D_hydrophobic_design” and “L_hydrophobic_design”). The definitions of these task operations are shown in Listing 3.11.

Where the task operations shown in Listing 3.11 define a particular rule (e.g., “prohibit glycine” or “prohibit everything that is not hydrophobic”), the domain over which a given task operation acts is defined by a residue selector. Residue selectors are defined above the TASKOPERATIONS section, in the RESIDUE_SELECTORS section of the script, as shown in Listing 3.12. They can select residues by index (the various Index selectors), or by geometric properties, such as whether the ϕ dihedral angle is positive or negative (the Phi selector) – a key determinant of whether a position is compatible with D- or L-amino acids. Residue selectors can also be used to configure other residue selectors – particularly those that allow selections to be combined with Boolean logic (And, Or, and Not selectors).

In the “genkic_steps” protocol, the “softdesign” mover is used in conjunction with the “min1” mover, a named instance of a MinMover, which performs gradient-descent minimization to adjust the positions of the newly designed side chains of the peptides, relieving minor steric clashes. The definition of this mover is shown in Listing 3.13. This mover is configured with a *move map* (a Rosetta tool for specifying movable degrees of freedom within a pose) to alter neither backbone degrees of freedom nor rigid-body transforms (jumps) between chains, nor any side chain except for those of the peptide. Where the rotamer optimization step was performed with a softened energy function that was forgiving of small clashes, the minimization step is performed with the ordinary REF2015 energy function (“r15_cst”, defined earlier in the script in the SCOREFXNS section) so that any clashes introduced during packing are relieved.

The result of the “softdesign” and “min1” movers is an initial, crude sequence, designed at moderate computational cost, with no motion of target side chains or of the target or peptide backbone permitted. This marks a second point in the “genkic_steps” protocol (Listing 3.7) at which filters are applied to determine whether to accept or reject the candidate peptide conformation sampled by the generalized kinematic closure mover. If filters pass, more extensive flexible-backbone design is carried out with the FastDesign mover, shown in Listing 3.14. Since FastDesign alternates between packing steps and minimization steps (ramping the repulsive part of the van der Waals term in the energy function for successive steps), it is configured both with task operations (to configure packing or design steps) and with a move map (to configure gradient-descent minimization steps). In this case, the task operations are the same as those used to configure “softdesign” except that rather than prohibiting packing at any target position, only those positions far from the peptide-protein interface are prohibited from packing (“no_repack_target_far_from_interface”). This means that target residues close to the peptide can adopt new conformations that may better accommodate peptide residues. The move map permits motion of all side chains and of the peptide backbone but disables changes to interchain rigid-body transforms (jumps) or to the target backbone. Candidate peptide backbone conformations sampled by generalized kinematic closure can therefore move slightly to accommodate the peptide sequence prior to selection of a “best” sampled conformation for subsequent refinement. During this step, an energy function in which the backbone hydrogen bonding term has been upweighted (“r15_highhbond_aacomp_cst”) is used to discourage loss of any hydrogen bonds

formed during initial sampling as the backbone moves. As the name implies, the energy function also has design-centric guidance terms like `aa_composition` and the `aspartimide_penalty` added.

From those solutions that pass all filters in the “genkic_steps” protocol, the lowest-energy solution is selected. The main protocol, defined in the `PROTOCOLS` section of the script, continues with a Monte Carlo refinement step, “mc_search”. The “mc_search” mover, shown in Listing 3.15, is configured to carry out a 500-step Monte Carlo search. At each step, the “mc_steps” protocol, defined in Listing 3.16, is applied, and the result is accepted or rejected based on the Metropolis criterion (Eq. 3.6). At the end of the Monte Carlo trajectory, the lowest-energy state encountered is returned.

Each step in the Monte Carlo trajectory involves perturbation of the dihedral angles of the peptide, followed by re-closure using generalized kinematic closure. As shown in Listing 3.17, a second instance of the generalized kinematic closure mover, named “genkic_perturb”, is configured to perturb loop degrees of freedom rather than randomizing them fully, to carry out minimal filtering on the solutions found, and to return the solution closest to the previous peptide conformation (ensuring that the perturbation is minimal). Following re-closure of the perturbed peptide, rapid rounds of peptide design (“softdesign”) and minimization (“min1”) are performed, allowing the peptide sequence to adapt to the perturbed conformation. The Monte Carlo step concludes with a series of filtering steps that trigger immediate rejection of the move if any filter fails. Only if all filters pass is the Metropolis criterion applied. Over the course of the Monte Carlo trajectory, both conformation and sequence become better optimized for the target.

The protocol ends with another application of the expensive FastDesign mover (“fdes”) to allow additional flexible-backbone sequence design starting from the best conformation found during the Monte Carlo search. Final relaxation and filtering ensure that no new problems have been introduced by these additional rounds of design.

Although this protocol successfully produced inhibitors of the NDM-1 antibiotic resistance factor, it should be viewed as a starting point for future design efforts. A variant of this approach that was applied to create inhibitors of histone deacetylases (HDACs) also serves as a useful reference; see Hosseinzadeh et al. (2021) for details. Every design problem is different, and a drug designer must work to optimize a design protocol for a particular target, for the available computational resources, and for the experimental and biological constraints of the problem. It is advisable to maintain the general pattern of seeking to eliminate poor designs as early as possible and reserving expensive computations, such as flexible-backbone design algorithms or Monte Carlo searches of sequence and conformation space, for refining designs that pass initial filters. However, it remains to be seen whether the particular strategy of finding a first closed solution, carrying out Monte Carlo refinement, and then applying flexible-backbone design will be applicable to all problems. Designers should not fear to improve upon this protocol or to experiment with alternatives.

3.3.5 Computational Validation of Designs

Once a pool of peptide macrocycles has been designed in silico, a key challenge is selecting the best candidates for chemical synthesis. The design pipeline to this point may have consumed on the order of a core-hour of computing time per design, which, in an age of readily available parallel computing power, is modest. Chemical synthesis requires many person-hours of labour per design, however, as well as expensive reagents and valuable hours on instruments like peptide synthesizers and HPLC purification apparatus. Additionally, the in vitro test for the desired peptide function – binding to a particular target protein, inhibiting a particular enzyme’s activity, preventing a certain protein-protein interaction, etc. – will require a certain number of person-hours per design and some amount of wet-lab resources. While some experiments can be performed on a very large scale, other experiments are intrinsically low-throughput techniques. Since the cost per design is far higher in the synthetic and experimental phases than in the computational phase, one may find oneself in the position of having produced thousands of promising designs in silico while only being able to synthesize and assay dozens or so in vitro. An in silico means of ranking designs and prioritizing the best for synthesis is therefore needed. The goal of any such method is to select a small pool of designs from a larger pool while enriching the smaller pool for designs that will show success in experiments.

3.3.5.1 Ranking Designs by Computed Interaction Energy with the Target

Intuitively, one would expect the computed strength of interaction between a designed peptide and a target protein to correlate with the experimentally measured binding affinity of the peptide for its target. Indeed, if the Gibbs free energy change of a folded peptide transitioning from an unbound to a bound state, $\Delta G_{\text{binding}\#\text{folded}}$, could be computed exactly, one could predict the equilibrium constant $K_{\text{eq, binding}\#\text{folded}}$ for the binding by rearranging Eq. 3.3 to yield Eq. 3.16:

$$K_{\text{eq, binding}\#\text{folded}} = e^{\frac{-\Delta G_{\text{binding}\#\text{folded}}}{RT}} \quad (3.16)$$

In the above equation, R is the gas constant, and T is the absolute temperature. Interaction energies can be estimated in a computation examining a single conformation of a protein, using a molecular dynamics force field such as CHARMM or AMBER or the energy function of a molecular modelling package such as the Rosetta REF2015 energy function. Such a calculation is extremely rapid, typically consuming milliseconds on a single core of a modern CPU for a medium-sized protein with implicit solvent. To make valid comparisons between designs, however, all designed structures should be prepared in the same manner. In particular,

since energy landscapes are extremely rugged, the structure should be relaxed by gradient-descent energy minimization using the same force field or energy function that is to be used for ranking designs. Without this step, one could be comparing one designed sequence in a conformation that is at a local energy minimum with another that is in a high-energy conformation, but which could be very close to a very low-energy local energy minimum.

Interaction energy computations predominantly compute the enthalpy of forming favourable polar interactions with a target. For hydrophobic peptides binding to hydrophobic cavities, solvent entropy may be the dominant factor driving peptide binding however. Implicit-solvent energy functions such as Rosetta's REF2015 energy function contain solvation terms that approximate solvent entropy effects using a fictional potential, called *fa_sol*, that causes hydrophobic chemical groups to attract one another and hydrophobic and polar groups to repel one another. Since this is an approximation, this can be a source of error that limits the correlation between predicted interaction energy and measured interaction.

In practice, a point computation in implicit solvent using the same energy function that was optimized during the design process tends to yield an estimate of the energy of the system that correlates poorly with binding affinity in experiments (Fig. 3.6). There are three reasons for this. First, as discussed above, force field-based energy functions are approximate (particularly when modelling solvent entropy effects), and the computed energy has an intrinsic level of error. If one has already optimized a set of designs extensively to yield low computed energies, then the magnitude of the expected energy differences between designs is likely to be less than the magnitude of the error in the energy function. To address this, it is conceivable that energy minimization and rescoring with a higher-accuracy energy function (which may be more computationally expensive to evaluate) could yield better discriminating power, though this is an area of ongoing research. Second, a point calculation performed on a single conformation and binding orientation could yield an *accurate* estimate of the energy of an *inaccurate* microstate of the system. Put another way, one may not have sampled the true bound state, and this may mean that the states being compared are higher-energy states. Ongoing research is investigating whether increased sampling of conformations near the hypothesized binding conformation, using either Rosetta Monte Carlo sampling strategies or MD simulations, could yield better predicted energies with better correlations with experiment. However, even if all of these accuracy issues were addressed, a third factor would limit the utility of these predictions: Computations that consider only the binding-competent conformation of a peptide and the bound state cannot take into account the entropic cost of ordering a disordered molecule on binding: $\Delta G_{\text{binding} | \text{folded}}$ is not $\Delta G_{\text{binding}}$. The true change in Gibbs free energy associated with a peptide binding its target must include some consideration of the conformational entropy change. The next section discusses means of addressing this.

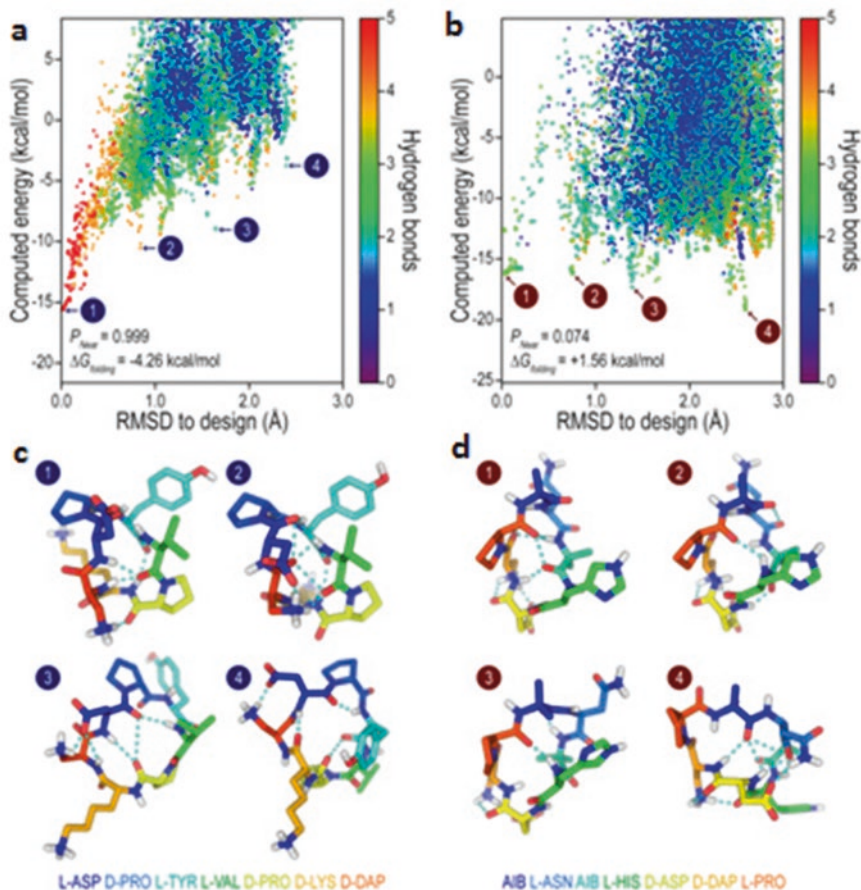


Fig. 3.7 Examples of peptide macrocycles with high and low fold propensities as evaluated with the Rosetta Simple_Cycpep_Predict application. (a) Computed conformational energy landscape of a peptide macrocycle with extremely high fold propensity ($P_{near} = 0.999$). Each point in this plot represents the energy (vertical axis) and RMSD to the designed conformation (horizontal axis) of a separate sampled conformation following rotamer optimization and energy minimization. Approximately 50,000 conformations were sampled in total. This peptide possesses a unique, low-energy state (point labelled 1) that closely matches the designed conformation. All other local minima in the energy landscape (e.g., points labelled 2, 3, and 4) are considerably higher in energy. Note that the designed state is stabilized by a large number of backbone hydrogen bonds (colours). (b) Computed conformational energy landscape for a peptide macrocycle with poor fold propensity ($P_{near} = 0.074$). The lowest-energy state near the designed conformation (point labelled 1) is nearly isoenergetic with states much further from the designed conformation (points labelled 2 and 3) and is considerably higher in energy than state 4. The overall shape of the landscape lacks the distinctive funnel shape shown in panel (a). (c and d) Sampled conformations of the peptides plotted in panels (a) and (b), respectively. Numbers correspond to labelled conformational states. Polar contacts including hydrogen bonds are shown as dashed cyan lines. Amino acid sequences are shown at the bottom. DAP = 2,3-diaminopropionic acid; AIB = 2-aminoisobutyric acid. Note that the energy difference between conformations 1 and 2 in panel (c) is large despite a very small difference in structure, indicating this peptide's rigid preference for its native state. Also note that the energy difference between conformations 1 and 4 in panel D is small, despite an enormous conformational rearrangement, indicating this peptide's flexibility

3.3.5.2 Ranking Designs by Computed Rigidity in the Binding-Competent Conformation

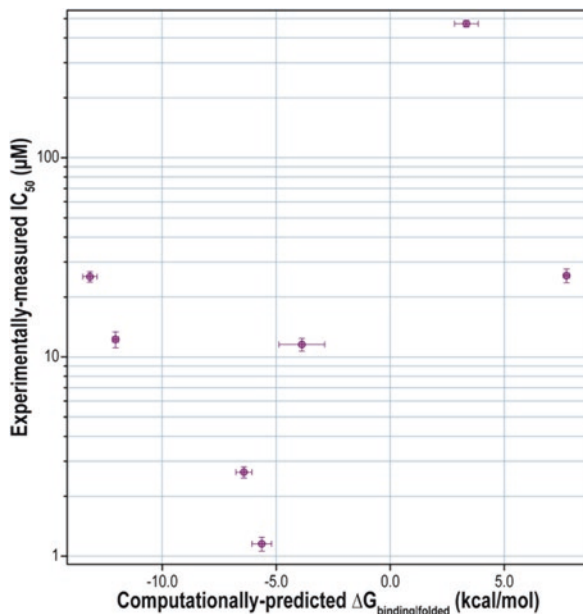
As shown in Eq. 3.4, the change in Gibbs free energy when a peptide macrocycle binds to a target has two components. The first is the free energy change of folding, $\Delta G_{\text{folding}}$, which is dominated by the entropic term $-T\Delta S_{\text{folding}}$. The second is the free energy change of the folded peptide binding to its target, $\Delta G_{\text{binding} \mid \text{folded}}$. We have already seen how design calculations attempt to minimize $\Delta G_{\text{binding} \mid \text{folded}}$. To estimate $\Delta G_{\text{folding}}$, which contains the contribution of conformational entropy to the overall value of $\Delta G_{\text{binding}}$, it is necessary to consider the ensemble of conformations that a peptide can access in the unbound state. This ensemble could be sampled using a single very long molecular dynamics simulation (perhaps representing milliseconds of simulated time) at enormous computational expense, but this is unattractive if one wishes to perform in silico screening of many candidate peptides. One could also conceivably use many short MD simulations, each starting from a different starting conformation, but this would require a means of generating a reasonably evenly distributed set of starting conformations. Such a means could itself be used as a means of sampling conformations, without the expense of MD simulations, which simulate the full molecular motions of transitioning from one metastable state to another at great computational cost.

The Rosetta software suite includes an application called `simple_cycpep_predict`, which uses generalized kinematic closure to rapidly sample conformations of an input peptide sequence in isolation (Bhardwaj et al. 2016; Hosseinzadeh et al. 2017). For every conformation sampled, the application performs side-chain optimization and energy minimization using the FastRelax protocol, and returns the relaxed structure and its energy. If the user provides a desired backbone conformation (e.g., the binding-competent conformation of a peptide designed to bind to a target), then the application can also compute the backbone root mean squared deviation (RMSD) between each sampled conformation and the desired conformation as well. After sampling tens or hundreds of thousands of conformations, the graph of energy versus RMSD provides insight into whether the peptide uniquely favours the binding-competent conformation or whether there exist alternative states that are equally or more favourable, as shown in Fig. 3.7.

While visual inspection can select high-quality designs, this is impractical given thousands of designs. To compare designs quantitatively and in an automated fashion, a metric called P_{near} was developed, which provides a rough estimate of the fraction of time that the peptide spends close to the desired conformation (Bhardwaj et al. 2016). P_{near} ranges from 0 to 1, with values near 0 meaning that the peptide spends very little of its time in native-like conformations and values near 1 meaning that the peptide spends the considerable majority of its time in native-like conformations. The definition of P_{near} is shown in Eq. 3.17.

$$P_{\text{near}} = \frac{\sum_{i=1}^N e^{-\left(\frac{r_i}{\lambda}\right)^2} e^{-\frac{E_i}{k_B T}}}{\sum_{j=1}^N e^{-\frac{E_j}{k_B T}}} \quad (3.17)$$

Fig. 3.6 Relationship between experimentally measured enzyme inhibition (vertical axis; lower is better) and computationally predicted binding free energy change (horizontal axis; more negative is better) for seven peptide macrocycle inhibitors of the NDM-1 antibiotic resistance factor. Since all designs were heavily optimized to maximize favourable interactions between target and peptide, there is no discernible correlation between the computed interaction energy and the experimentally measured inhibition, making interaction energy a poor metric for ranking designs. Figure redrawn from data published in Mulligan et al. (2021)



In the expression above, N is the number of samples, r_i is the RMSD of the i^{th} sample to the desired conformation, and E_i and E_j are the energies of the i^{th} and j^{th} samples, respectively. The expression also has two tunable parameters, λ and $k_B T$. The first, λ , is a parameter measured in Ångstroms that defines how close a structure must be to the native state in order for it to be considered “native-like”. For peptides in the 7- to 12-amino acid residue range, a value of 1.5 Å is typically used. The second, $k_B T$, is a parameter measured in energy units that determines the extent to which high-energy states are represented in the conformational ensemble. At higher values, the weighting grows more even, and as $k_B T$ approaches zero, the lowest-energy samples dominate. A value of 0.62 kcal/mol, which corresponds to physiological temperature, is typically used.

To understand the P_{near} metric, it can be simpler to express it as follows:

$$P_{\text{near}} = \frac{\sum_{i=1}^N f(r_i) e^{-\frac{E_i}{k_B T}}}{Z} \quad (3.18)$$

In the above, Z is the partition function (which in the limit of infinite sampling is equal to $\sum_{j=1}^N e^{-\frac{E_j}{k_B T}}$), and $f(r_i)$ is a weighting function that determines the extent to

which each sample is counted as native-like. Written as such, it is easier to see that P_{near} represents the Boltzmann probability of the peptide existing in conformations for which $f(r_i)$ is large. If $f(r_i)$ were a hat function, equal to 1 below some RMSD threshold and 0 above that threshold, then P_{near} would represent the probability of finding the peptide in a conformation within that threshold distance of the desired conformation. However, because sampling is discrete and stochastic, it is inconvenient to use a discrete threshold: small, random changes in the distribution of samples from run to run could result in the RMSD value of a few low-energy points drifting slightly from run to run, and if a point crossed the threshold, the value of P_{near} could change drastically. By replacing the hat function with a Gaussian function of RMSD, the definition of “native-like” becomes fuzzy: samples that exactly match the desired conformation count as perfectly “native-like”, samples very different from the desired conformation count as not at all “native-like”, but samples close to the native conformation can count as *somewhat* native-like (i.e., $f(r_i)$ can yield values intermediate between 0 and 1). This has a stabilizing effect on the numerical computation: small changes in the RMSD values from run to run have small effects on the value of P_{near} since there is no hard threshold that results in a large change in $f(r_i)$.

Since P_{near} can be interpreted as an estimate of the probability that a peptide is folded, it can be related to the folding equilibrium constant as follows:

$$K_{\text{eq.folding}} = \frac{[\text{folded}]}{[\text{unfolded}]} \approx \frac{P_{\text{near}}}{1 - P_{\text{near}}} \quad (3.19)$$

This means that $\Delta G_{\text{folding}}$ may be estimated by Eq. 3.20:

$$\Delta G_{\text{folding}} = -RT \ln(K_{\text{eq.folding}}) = -n_{\text{Av}} k_B T \ln(K_{\text{eq.folding}}) \approx -n_{\text{Av}} k_B T \ln\left(\frac{P_{\text{near}}}{1 - P_{\text{near}}}\right) \quad (3.20)$$

In the above, R is the gas constant, n_{Av} is Avogadro’s constant, $n_{\text{Av}} k_B T$ is the Boltzmann constant, and T is the absolute temperature. The value of $n_{\text{Av}} k_B T$ is typically 0.62 kcal/mol to represent a physiological temperature of 310 K.

When the `simple_cycpep_predict` application is run using the Message Passing Interface (MPI) to efficiently parallelize conformational sampling over many processes on many CPU cores, either on a single node or across many nodes in a computing cluster, the application will automatically compute P_{near} and $\Delta G_{\text{folding}}$ for the user from the ensemble of samples. (To conserve memory on each node, the application also permits thread-level parallelism.) Computational time for sampling scales with $O(N)$ (where N is the number of samples); however, the number of samples needed scales exponentially with the length of the peptide macrocycle since longer macrocycles can access far more conformations. This means that the application can only reliably sample exhaustively for macrocycles of lengths up to about 15 or 16 amino acids unless sampling is constrained in some way (e.g., using known conformational symmetry as described in Mulligan et al. (2020) or known

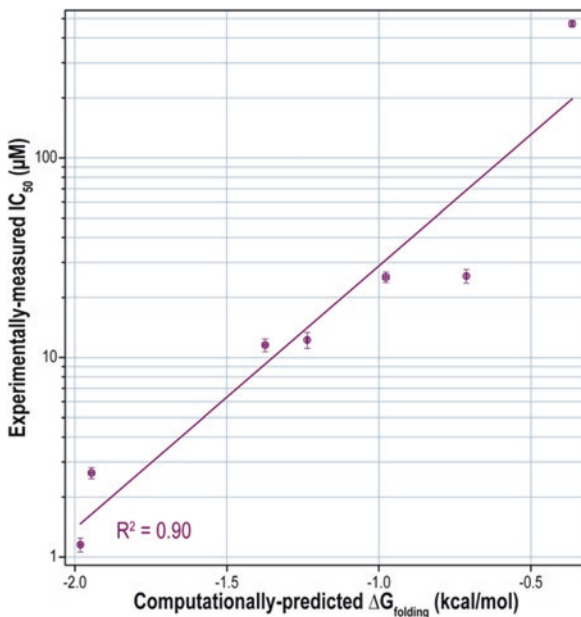
secondary structure as described in Dang et al. (2017)). For small macrocycles (7 to 12 amino acids), 20,000 to 100,000 samples are usually sufficient and will typically cost a few hundred core-hours to generate. This makes `simple_cycpep_predict` a somewhat costly computational method for final validation but still considerably less expensive than MD simulations.

Practically, it has been shown that when peptides are heavily optimized for $\Delta G_{\text{binding} \setminus \text{folded}}$ during the design process (the effect of using methods that optimize an energy function using point energy calculations in the bound configuration), $\Delta G_{\text{folding}}$ becomes an excellent predictor of binding affinity in wet-lab experiments, as shown in Fig. 3.8. This provides a powerful means of ranking designs to prioritize syntheses and wet-lab experiments.

3.3.6 Experimental Validation of Top Designs

Once a large pool of designs has been generated in silico, filtered to remove those designs with obvious problems or pathologies, and ranked using available validation methods, a short list of top designs can be chemically synthesized and

Fig. 3.8 Correlation between experimentally measured enzyme inhibition (vertical axis; lower is better) and computationally predicted fold propensity (horizontal axis; more negative is better) for seven peptide macrocycle inhibitors of the NDM-1 antibiotic resistance factor. Since all were heavily optimized during design for favourable interaction energies between peptide and target, fold propensity is a major predictor of binding and inhibition, even predicting the rank order of these seven peptides. This makes fold propensity a very useful metric for ranking designs. Figure redrawn from data published in Mulligan et al. (2021)



Validating Peptides by Computed Folding Free Energy

1. Conformational sampling simulations can estimate the Gibbs free energy of folding of small peptide macrocycles.
2. Propensity to favour the binding-competent conformation is an excellent predictor of success for peptides that have been heavily optimized for favourable binding interactions.
3. The computational expense of these simulations is somewhat high, but lower than MD-based validation.

evaluated experimentally. While full details of synthesis and experimental evaluation are outside of the scope of this chapter, experimental evaluation is often an iterative process. Well-constructed, quantitative, medium- to low-throughput experiments can yield information that high-throughput experiments yielding pass/fail results cannot, allowing identification of partial successes. Because peptides are mutable, binders of low to moderate affinity can be redesigned to promote tighter binding, yielding pools of computationally generated variants that can be screened in subsequent rounds of experimentation. Where structural studies are possible, experimentally determined structures of first-round designs in complex with the target are excellent starting points for design refinement.

3.4 Emerging Technologies

Computational methods for drug discovery advance rapidly. The rise of open-source software has permitted method developers to build iteratively upon one another's work, producing newer and more powerful workflows continuously. In addition, two powerful technologies are emerging as potential game changers in the drug design field: machine learning methods are enjoying a renaissance, and quantum computing hardware is starting to reach the point of being useful. This section reviews these exciting new technologies and what they mean for peptide macrocycle drug design.

3.4.1 Deep Neural Networks and Other Machine Learning Methods

The approaches reviewed in this chapter have all been based on traditional algorithms, developed rationally by human programmers. Machine learning (ML) represents a complementary approach, which, thanks to recent advancements in

computer hardware and software methods, has lately regained popularity. Given a computational task, a measure of how well a given algorithm performs that task, and a set of data that plausibly contain information relevant to executing that task, an ML algorithm is one that is able to improve its performance of the task, as measured by the performance measure, through exposure to the data (Goodfellow et al. 2016). Although a human typically constructs the general architecture of an ML model for a given type of task and manually tunes hyperparameters that govern how the ML model learns from the data, ML is distinct from traditional programming in that the ML model is then subsequently permitted to optimize its own internal parameters or other aspects of its inner workings without the direct intervention of a human. In this sense, the trained ML model is not wholly the product of rational human design and may end up operating by surprising, counter-intuitive, or even difficult-to-interpret internal processes. In the case of *supervised machine learning*, training data are provided in the form of pairs of algorithm inputs and known, desired outputs, the performance measure scores how well an algorithm produces outputs close to the known outputs, and the training process yields an algorithm better able to produce the expected outputs given the corresponding inputs. Although many classes of ML algorithms exist, recent advancements in using deep neural networks for image and language processing have generated enormous enthusiasm for *deep learning* or ML based on deep neural networks.

Machine Learning: Definitions

- Machine learning refers to the development of algorithms that improve their performance on a task, as measured by an objective function, on exposure to data.
- Supervised machine learning is a subset of machine learning in which algorithms are trained on training sets consisting of particular inputs and the corresponding, known desired outputs.
- Deep learning is a subset of machine learning that uses deep neural networks as a general type of architecture that can learn.

A deep neural network (DNN) is a generic computing structure with tunable complexity, which, given sufficient training data, can learn to approximate any function mapping inputs to outputs. Internally, it consists of a series of linear tensor operations interspersed with element-wise non-linear operations, with thousands to millions of tunable parameters that control the mapping of inputs to outputs (Goodfellow et al. 2016). Compared to other ML approaches, DNNs require large training datasets but can achieve impressive performance at complex tasks when sufficient data are available.

Recently, DNNs have begun to feature heavily in the biennial Critical Assessment of Protein Structure Prediction (CASP) competitions in which competitors attempt to predict protein structures from amino acid sequence alone or from sequence and limited other data. Since the Protein Data Bank now contains over 160,000

experimentally determined protein structures (Burley et al. 2019), and since there are now over 197 million protein sequences in the NCBI Reference Sequence Database (O’Leary et al. 2016), there are considerable pools of training data for building and training fold prediction DNNs. This has allowed methods such as RaptorX-Contact (Wang et al. 2018), trRosetta (Yang et al. 2020), and AlphaFold (Senior et al. 2019, 2020) to achieve remarkable accuracies when predicting structures of proteins built from canonical amino acids. These DNNs operate on multiple sequence alignment (MSA) input data and infer long-range contacts between pairs of amino acids based on co-evolutionary relationships discernible from the MSAs, producing as output a contact matrix, estimated distance probability distribution matrix, or estimated residue-residue orientation probability distributions. Although these methods rely on more traditional simulation-based algorithms to convert this contact information into a full three-dimensional protein structure, more recently, the RoseTTAFold DNN has produced full backbone 3D coordinates as output from an input MSA, with accuracy sufficient to allow X-ray crystal structures to be solved by molecular replacement (MR) phasing (Baek et al. 2021). The AlphaFold2 DNN takes this one step further, providing both backbone and side-chain geometry as output, though it still relies on energy minimization using a molecular mechanics force field for final all-atom refinement. To date, AlphaFold2 produces the most accurate structure predictions of any ML or simulation-based method, frequently achieving sub-Ångstrom precision, but still has difficulty with protein complexes dominated by interchain contacts or MSAs with fewer than 30 sequences (Jumper et al. 2021). Ongoing challenges include building DNNs that can produce predictions using smaller amounts of memory or computing time: AlphaFold2 and RoseTTAFold require high-memory GPUs and GPU-minutes to GPU-hours per prediction. Beyond the prediction of protein structure, DNNs have been employed to predict protein function from structure and sequence (Gligorijević et al. 2021) and to compare protein surface features (Simonovsky and Meyers 2020). For a more detailed review of deep learning in macromolecular modelling, see Mulligan (2021).

Although they show impressive ability to interpolate based on patterns in large training datasets, DNNs often have more difficulty extrapolating beyond the domain of the training data. A particular challenge when modelling peptides is the fact that peptides can be built from non-canonical amino acids, possessing physical and conformational preferences distinct from those of the canonical amino acids. Unfortunately, where canonical amino acid training data are abundant, non-canonical amino acid training data are sparse. This hinders attempts to develop neural networks based on experimental data. However, an alternative is to train neural networks based on the output of expensive but general physics-based simulations, which can accurately model non-canonical amino acids, in order to produce a fast means of approximating the information from a simulation without performing the simulation. This is a topic of ongoing research. In coming years, deep learning may begin to complement expensive simulations (such as peptide conformational sampling simulations for validation) in order to reduce the computational cost of designing and validating peptide therapeutics.

3.4.2 Quantum Computing

All but the simplest quantum mechanical systems are computationally intractable even with simplifying methods such as Hartree-Fock theory, DFT, or FMO methods. Part of the reason for this computational intractability is that quantum mechanical systems can exist in a superposition of quantum states, where the possible number of superimposed states for a many-component system scales with the product of the number of states accessible to each component. This means that computing the contribution of all states as the system evolves requires operations and memory that scale with $O(C^N)$, where C is the number of states per component and N is the number of components. (For a multi-electron spin system, e.g., each electron can exist in one of two spin states, and the composite N -electron system can exist in 2^N states. A system with just 270 electrons can exist in a superposition of more states than there are atoms in the universe!)

In the early 1980s, Richard Feynman proposed that where this represents a perhaps intractable obstacle for conventional or classical computing, it may represent an opportunity for developing new types of computers (Feynman 1982). Since quantum mechanical systems evolve in accordance with the Schrödinger equation, the simulation of which requires time and memory that scales with $O(C^N)$, the observation of the outcome of a series of real quantum mechanical operations applied to a real quantum mechanical system provides the solution to an $O(C^N)$ problem. If a hard problem of interest can be mapped to the form of the equations describing the quantum mechanical system, it too could be solved in this way. That is, computers that use quantum mechanical systems could conceivably provide a means of solving $O(C^N)$ problems efficiently, perhaps in linear or polynomial time. While one obvious application of quantum computers would be to QM simulations of molecules, many other difficult combinatorial problems that arise in peptide design and heteropolymer modelling could also conceivably be solved efficiently on quantum computers. The problem of designing peptide sequences, for example, is one with a solution space that scales with $O(C^N)$, where C is the number of amino acid residue types allowed at each position and N is the length of the peptide. This scaling necessitates Monte Carlo methods or other heuristics that search small subsets of the possible sequences on classical computers, but a quantum computer could potentially model *all possible sequences* as a superposition of quantum states. A suitable quantum algorithm could optimize over all possible sequences, returning the one that yields the best score for some objective function. Another possibility would be the conformational problem: Given C discrete backbone conformations for a single amino acid, exhaustively exploring the conformations of a peptide or polypeptide with N amino acids is an $O(C^N)$ operation. Again, where classical approaches involve simulated annealing-based searches of the conformation space that consider one possible conformation at a time, a quantum algorithm could conceivably represent *all possible conformations* and rapidly sample from the lowest-energy conformations.

From the 1990s onward, considerable progress has been made on the theory of quantum computing, particularly with two-state components (termed *qubits*, in analogy with the binary bits of classical computers). Current-generation quantum computers are limited by small size (few qubits that can be entangled to produce a composite quantum mechanical system) and by sensitivity to noise. The former limits the information content of problems that can be modelled, and the latter limits the number of operations that can be performed in the course of a quantum algorithm before the system deteriorates into a random state. For these reasons, we are far from the point at which these computers will be useful for QM calculations or for true conformational sampling. However, various groups have mapped 2D HP lattice toy models, a simplified version of the protein conformational sampling problem, to current-generation quantum computers (Perdomo-Ortiz et al. 2012; Babej et al. 2018). Additionally, it has been demonstrated that it is possible to design peptide sequences on a quantum computer (Mulligan et al. 2019). In both of these cases, the quantum computer used was the D-Wave quantum annealer, which is a special-purpose quantum computer capable of solving optimization problems with a very specific functional form to the objective function. The restrictions on the form of the objective function have allowed the construction of larger quantum computers of this type (up to 5,000 qubits at the time of this writing, albeit with very limited inter-qubit connectivity (McGeoch and Farré 2020)) but narrow the range of problems to which these systems will be applicable. For instance, when used for peptide design, the D-Wave quantum annealer could only optimize rotamer selections using energy functions that were residue level pairwise-decomposable, preventing use of non-pairwise design-centric guidance terms (Mulligan et al. 2021). Ongoing research aims to map conformational sampling and design algorithms to general-purpose, gate-based quantum computers, though current general-purpose quantum computing hardware is extremely limited. (The largest IBM gate-based quantum computer at the time of this writing has 65 qubits and is limited to tens of operations before losing quantum coherence (Cho 2020), but unlike the D-Wave system, any qubit can be entangled with any other.) This means that although algorithms can be planned out on paper, and even tested for the smallest problems using a classical supercomputer simulating a small quantum computer, it is very hard to test these algorithms on real general-purpose quantum computing hardware at present. In the long run, as the engineering challenges of building large and robust general-purpose, gate-based quantum computers are surmounted, quantum computing may prove to be a powerful technology for moving past the intrinsic limits of classical computing to tackle the hard, exponentially scaling problems that frequently arise in peptide and protein modelling.

Toward the Future: Quantum Computers for Peptide and Protein Modelling

- Quantum computers are computers that use quantum mechanical principles, such as superpositions of particle states, in their operation. Early, primitive quantum computers have been built, and the hardware is rapidly advancing.
- Quantum computers may be specialized or general-purpose. Adiabatic quantum annealers are specialized quantum computers that solve a particular class of optimization problem, with a particular functional form. Gate-based quantum computers are general-purpose devices able to implement a much broader range of algorithm types but which are currently at an earlier stage of development.
- Quantum algorithms are computational approaches that take advantage of the massive parallelism that comes from the fact that a collection of N particles with C states each can exist in a superposition of C^N states at once, all of which may be operated upon simultaneously. However, on measurement, only one state is observed. This makes quantum algorithms most useful for search problems in which astronomical numbers of possibilities must be considered, but a single “best” possibility is of interest.
- Conformational sampling, sequence design, and QM energy calculations are all problems that could potentially benefit from robust quantum computers.

3.5 Conclusions

This chapter has covered the theory, major algorithms, and current pipelines for computationally designing peptide macrocycles that are able to bind to targets of therapeutic targets. The exploration of a current protocol for NDM-1 inhibitor drug design presented here is intended to introduce drug developers to current best practices, though those wishing to learn more are encouraged to examine the relevant scripts in full detail and to alter or adapt these methods as needed for the unique challenges presented by their own target proteins. At the time of writing, the most robust software for this purpose is the Rosetta software suite, which offers a full end-to-end pipeline for building and sampling peptide backbones, designing peptide sequences, and computationally validating and ranking pools of designs to prioritize syntheses. Nevertheless, computational methods for drug design are an area of active research, and new tools are emerging at a rapid pace. The current renaissance in machine learning methods is likely to be transformative in the coming years, particularly as new means of using traditional algorithms and simulation-based methods in conjunction with deep neural networks are devised. And on the horizon, new computing paradigms such as quantum computers may make possible certain design tasks that are currently intractable, such as exhaustively exploring the conformation or sequence space of a peptide or even a protein. The rapid pace of the technology presents challenges for a drug designer who must stay abreast it but also creates exciting opportunities for developing powerful, next-generation peptide macrocycle therapeutics at much lower cost in time, effort, and resources.

Appendixes

Appendix A: Running Examples

Obtaining and Compiling the Rosetta Software

Although it is not open-source software, Rosetta is free for academic, government, or non-profit use and is licenced for commercial use. Free and commercial licences may be obtained from the University of Washington's CoMotion office (<https://els2.comotion.uw.edu/product/rosetta>). For licenced users, the software's source code can be downloaded from <https://www.rosettacommons.org/software>.

Once downloaded and unpacked, Rosetta must be *compiled* before use. (This means converting the human-readable C++ code into machine instructions that can be executed.) To compile Rosetta for default single-process, single-threaded execution, open a Linux or MacOS terminal and navigate to Rosetta's source directory. For example, if Rosetta had been installed in the directory `/home/myusername/rosetta/`, one would type `> cd /home/myusername/rosetta/Rosetta/main/source.`

Now, one can use the `scons` command to compile all Rosetta binaries. The command takes a flag, `-j`, used to specify the number of parallel processes to launch in order to compile more quickly. Typically, one should set this to the number of cores available on one's system. For example, if one had an eight-core system, one would type `> ./scons.py -j8 mode=release bin`

Compilation will take on the order of an hour on eight cores. If compilation is interrupted (either deliberately by pressing `Ctrl+C` or accidentally by a crash or a power outage), the `scons` compilation system is intelligent enough to continue from where it left off when the command is run again. It should be unnecessary to delete the existing subset of compiled binaries.

Certain features require special builds of Rosetta, specified with `extras=<desired feature>` in the `scons` command. For multithreading, one should compile with `extras=cxx11thread`. For process-based parallelism, include `extras=mpi,serialization`. These can be combined, as shown below:

```
> ./scons.py -j8 mode=release extras=cxx11thread,mpi,serialization bin
```

This will produce a separate build in a different subdirectory in the `Rosetta/main/source/build/` directory. Symlinks to executables are placed in the `Rosetta/main/source/bin/` directory and are given names with the pattern `<application_name>.<extras>.<OS-compiler-buildmode>`. For example, on a Linux system using the GCC compiler, a build with no extras will create applications with names like `rosetta_scripts.default.linuxgccrelease` or `simple_cycpep_predict.default.linuxgccrelease`, while on MacOS

with the clang compiler, a build with `extras=mpi,serialization` will create applications with names like `rosetta_scripts.mpiserialization.macosclangrelease` or `simple_cycpep_predict.mpiserialization.macosclangrelease`.

Obtaining Example XML Scripts and Inputs

The New Delhi metallo- β -lactamase 1 inhibitor design scripts described in Sect. 3.3.4.5 are distributed with recent releases of the Rosetta software and may be found in the directory `Rosetta/main/rosetta_scripts_scripts/scripts/public/macrocycle_inhibitor_design/NDMli-1_design_script/modernized/`. Alternatively, they may be downloaded from the Github repository `vmullig/ndml_design_scripts`. To do this, first create a directory for the repository and navigate to it:

```
> mkdir NDM1_scripts
> cd NDM1_scripts
```

Next, clone the git repository:

```
> git clone git@github.com:vmullig/ndml_design_scripts.git.
```

Running Example XML Scripts

To run the example, first navigate to the `modernized/` directory. This directory contains a subdirectory called `inputs/` with a `rosetta.flags` file containing all flags needed to run the example. This is done by invoking the `rosetta_scripts` application and passing the flags file. If Rosetta is installed in the `/home/username/rosetta` directory, one would type:

```
> <rosetta_path>/Rosetta/main/source/bin/rosetta_scripts.default.linuxgccrelease @inputs/rosetta.flags
```

In the command above, `<rosetta_path>` should be replaced with the absolute path to the Rosetta installation directory, and the suffix `.default.linuxgccrelease` should be updated as needed for one's compilation extras, operating system, compiler, and compilation mode. (For example, on MacOS with the clang compiler and multithreading support, this would be `rosetta_scripts.cxx11thread.macosclangrelease`.)

When running MPI builds of Rosetta to launch many parallel processes, one should consult the documentation for the version of MPI installed on one's system.

Often, the `mpirun` command must be invoked. For example, to launch four parallel processes for the `simple_cycpep_predict` application, the command would be:

```
> mpirun -np 4 <rosetta_path>/Rosetta/main/source/bin/simple_cycpep_
predict.mpiserialization.linuxgccrelease @inputs/rosetta.flags
```

Additional documentation for the Rosetta software suite is available online from <https://new.rosettacommons.org/docs/latest/Home>.

Appendix B: Code Listings

The example RosettaScripts XML code and input files referred to in the body of the chapter are listed here, as are excerpts from the New Delhi metallo- β -lactamase 1 inhibitor design script described in Sect. 3.3.4.5.

Listing 3.1 Example RosettaScripts XML script demonstrating the use of packer palettes to enable design with non-canonical amino acids and task operations to prune amino acid types in a position-specific manner. Long lines are wrapped with a backslash to indicate the wrapping

```
<ROSETTASCRIPTS>
<SCOREFXNS>
<!--
  This section of the script defines scoring functions. The default
  Rosetta REF2015 energy function will be used for this example:
  -->
<ScoreFunction name="r15" weights="ref2015" />
</SCOREFXNS>
<RESIDUE_SELECTORS>
<!--
  Residue selectors define rules for selecting regions of a pose.
  The first two here select residues with positive and negative
  values of the phi backbone torsion angle (a key determinant of
  whether a position is compatible with D- or L-amino acids):
  -->
<Phi name="select_positive_phi" select_positive_phi="true" />
<Not name="select_negative_phi" selector="select_positive_phi" />
<!--
  The next two selectors select buried positions and exposed
  positions, respectively:
  -->
```

```

<Layer name="select_core" select_core="true" select_
boundary="false" select_surface="false" />
<Layer name="select_surface" select_core="false" select_
boundary="false" select_surface="true" />
</RESIDUE_SELECTORS>
<PACKER_PALETTES>
<!--
  Packer palettes define the total set of amino acids that will be
  used for design. Here, we enable the canonical amino acids plus
  their D-amino acid counterparts:
  -->
<CustomBaseTypePackerPalette name="design_palette"
additional_residue_types="DALA,DASP,DCYS,DPHE,DGLU,DHIS,DILE, \
DLYS,DLEU,DMET,DASN,DPRO,DGLN,DARG,DSER,DTHR,DVAL,DTRP,DTYR"
/>
</PACKER_PALETTES>
<TASKOPERATIONS>
<!--
  Task operations take residue selectors as input, and prune the
  allowed amino acid types in a position-specific manner. The first
  two prohibit L- and D-amino acids at positions with positive and
  negative values of the phi backbone dihedral angle,
  respectively:
  -->
<ProhibitResidueProperties name="D_design" properties="L_AA"
selector="select_positive_phi"
/>
<ProhibitResidueProperties name="L_design" properties="D_AA"
selector="select_negative_phi"
/>
<!--
  The next two prohibit polar and hydrophobic residues in the core
  and on the surface, respectively:
  -->
<ProhibitResidueProperties name="hydrophobic_core"
properties="POLAR"
selector="select_core"
/>
<ProhibitResidueProperties name="polar_surface"
properties="HYDROPHOBIC"
selector="select_surface"
/>
</TASKOPERATIONS>

```

```

<MOVERS>
<!--
  Movers operate on a pose and modify it in some way. Here we
  configure a PackRotamersMover, using the four task operations
  defined above to allow only L-amino acids at negative-phi
  positions, only D-amino acids at positive-phi positions, only
  polar residues on the surface, and only hydrophobic residues in
  the core.
-->
<PackRotamersMover name="pack" scorefxn="r15"
  task_operations="D_design,L_design,hydrophobic_core,polar_surface"
  />
</MOVERS>
<PROTOCOLS>
<!--
  This section lists previously-defined movers to construct an
  overall protocol. Here we list only one, the "pack" mover
  defined above.
-->
<Add mover="pack" />
</PROTOCOLS>
<OUTPUT scorefxn="r15" />
</ROSETTASCRIPTS>

```

Listing 3.2 Penalty definition file `desired.comp` defining the desired amino acid composition for a sample design problem: exactly six isoleucine residues and no fewer than 20% charged (lysine) residues

```

# Define a penalty for having fewer than or more than 6 isoleucine
residues:
PENALTY_DEFINITION
TYPE ILE
DELTA_START -1
DELTA_END 1
PENALTIES 3 0 3
ABSOLUTE 6
BEFORE_FUNCTION QUADRATIC
AFTER_FUNCTION QUADRATIC
END_PENALTY_DEFINITION

```

```
# Define a penalty for having fewer than 20% charged (lysine)
residues:
PENALTY_DEFINITION
PROPERTIES CHARGED
DELTA_START -1
DELTA_END 1
PENALTIES 10 0 0
FRACTION 0.2
BEFORE_FUNCTION QUADRATIC
AFTER_FUNCTION CONSTANT
END_PENALTY_DEFINITION
```

Listing 3.3 Example RosettaScripts XML script that defines a protocol for designing a peptide. The design process is restricted to consider only the amino acids lysine, isoleucine, and alanine using a task operation and is further constrained using amino acid composition constraints defined in the file “design.comp” (Listing 3.2)

```
<ROSETTASCRIPTS>
<SCOREFXNS>
<!--
  In this section, we define a default scoring function, used for
  final output, as well as a custom scoring function, which has the
  aa_composition term appended to the default Rosetta ref2015
  energy function.
-->
<ScoreFunction name="default_scorefunction" weights="ref
2015.wts" />
<ScoreFunction name="custom_scorefunction"
weights="ref2015.wts" >
<Reweight scoretype="aa_composition" weight="1.0" />
</ScoreFunction>
</SCOREFXNS>
```

```
<TASKOPERATIONS>
<!--
  In this section, we define task operations, which control the
  behaviour of the Rosetta packer. We will define a single task
  operation for this example, which restricts design to use only
  lysine, isoleucine, and alanine.
-->
<RestrictToSpecifiedBaseResidueTypes
  name="only_use_KIA"
  base_types="LYS,ILE,ALA"
/>
</TASKOPERATIONS>
<MOVERS>
<!--
  Movers alter a pose in some way. Here, we define two. The first
  adds amino acid composition constraints, read from a file, to the
  pose. The second performs design, using a scoring function that
  respects the constraints because it includes the aa_composition
  scoring term.
-->
<AddCompositionConstraintMover
  name="add_composition_constraints"
  filename="desired.comp"
/>
<PackRotamersMover
  name="perform_design"
  scorefxn="custom_scorefunction"
  task_operations="only_use_KIA"
/>
</MOVERS>
<PROTOCOLS>
<!--
  In this section, we list the linear series of previously-defined
  operations (movers) to apply. This defines the overall protocol.
-->
<Add mover="add_composition_constraints" />
<Add mover="perform_design" />
</PROTOCOLS>
```

```
<!--  
  Prior to final output, the pose will be rescored using the  
  default REF2015 scoring function, and the scoring information  
  will be included in the PDB file that is written.  
  -->  
<OUTPUT scorefxn="default_scorefunction" />  
</ROSETTASCRIPTS>
```

Listing 3.4 Basic layout of a RosettaScripts XML script

```
<ROSETTASCRIPTS>  
<SCOREFXNS>  
<!--  
  In this section, scoring functions are declared that will be  
  used later in the protocol.  
  -->  
</SCOREFXNS>  
<RESIDUE_SELECTORS>  
<!--  
  This section declares residue selectors, which allow rule-based  
  selections of subsets of poses. They serve as inputs into  
  movers, filters, and task operations, allowing these to operate  
  on subsets of the pose instead of on the whole pose.  
  -->  
</RESIDUE_SELECTORS>  
<PACKER_PALETTES>  
<!--  
  This section declares packer palettes, which define the set of  
  amino acid building blocks used for a particular design step.  
  -->  
</PACKER_PALETTES>  
<TASKOPERATIONS>  
<!--  
  This section declares task operations, which allow position-  
  specific restrictions to be placed on the allowed residue types  
  for a particular design step. Task operations can also alter  
  other design parameters, such as the degree of discretization  
  of rotamers.  
  -->  
</TASKOPERATIONS>
```



```
<FILTERS>
<!--
  This section declares filters, which measure properties of a pose
  and make decisions about whether to continue the protocol or to
  abandon it in favour of a new attempt.
-->
</FILTERS>
<MOVERS>
<!--
  Movers are the workhorses of the protocol. Each mover modifies a
  pose in some way. They are declared and configured here.
-->
</MOVERS>
<PROTOCOLS>
<!--
  In this section, a sequence of previously-declared movers and
  filters is listed to define a protocol. Each will be executed in
  sequence.
-->
</PROTOCOLS>
<OUTPUTS>
<!--
  This section allows options to be set related to the final output
  from the protocol.
-->
</OUTPUTS>
</ROSETTASCRIPTS>
```

Listing 3.5 Series of steps for NDM-1 inhibitor design as laid out in the PROTOCOLS section of the RosettaScripts XML script

```
<PROTOCOLS>
<Add mover="extend" />
<Add mover="foldtree1" />
<Add mover="initialize_tors" />
<Add mover="connect_termini" />
<Add mover="add_cutpoint_upper_pep" />
<Add mover="add_cutpoint_lower_pep" />
<Add mover="global_comp" />
```

```

<Add mover="L_alpha_comp" />
<Add mover="D_alpha_comp" />
<Add mover="L_beta_comp" />
<Add mover="D_beta_comp" />
<Add mover="genkic" />
<Add mover="mc_search" />
<Add mover="fdes" />
<Add mover="final_frlx" />
<Add mover="connect_termini" />
<Add filter="oversat" />
<Add filter="shape3" />
</PROTOCOLS>

```

Listing 3.6 Configuration of the “genkic” mover used for initial closure of the peptide macrocycle

```

<GeneralizedKIC name="genkic" closure_attempts="100"
pre_selection_mover="genkic_steps"
stop_when_n_solutions_found="1"
  selector="lowest_energy_selector"
selector_scorefunction="r15_highhbond_cst"
>
<AddResidueres_index="239" />
<AddResidueres_index="240" />
<AddResidueres_index="241" />
<AddResidueres_index="234" />
<AddResidueres_index="235" />
<AddResidueres_index="236" />
<SetPivots
  res1="239" res2="234" res3="236"
  atom1="CA" atom2="CA" atom3="CA"
/>
<CloseBond atom1="C" res1="241" atom2="N" res2="234"
torsion="180"
bondlength="1.328685" angle1="116.2" angle2="121.7"
/>
<AddPerturber effect="randomize_alpha_backbone_by_rama" >
<AddResidue index="234" />
<AddResidue index="235" />
<AddResidue index="236" />
<AddResidue index="239" />
<AddResidue index="240" />
<AddResidue index="241" />
</AddPerturber>
</GeneralizedKIC>

```

Listing 3.7 Definition of the “genkic_steps” mover, a pre-selection mover applied to each candidate solution produced by the generalized kinematic closure algorithm in the NDM-1 design script

```
<ParsedProtocol name="genkic_steps">
<Add filter="oversat" />
<Add filter="total_hbonds" />
<Add filter="low_stringency_clash" />
<Add mover="softdesign" />
<Add mover="min1" />
<Add mover="connect_termini" />
<Add filter="oversat" />
<Add mover="min2" />
<Add mover="connect_termini" />
<Add filter="oversat" />
<Add filter="total_hbonds_2" />
<Add filter="shape1" />
<Add mover="fdes" />
<Add mover="connect_termini" />
<Add filter="oversat" />
<Add filter="shape2" />
</ParsedProtocol>
```

Listing 3.8 Some of the filters defined in the FILTERS section of the script and used in the “genkic_steps” protocol to remove initial closure solutions with oversaturated hydrogen bond acceptors (“oversat”), fewer than three internal hydrogen bonds in the peptide (“total_hbonds” and “total_hbonds2”), egregious clashes between the peptide backbone and the target (“low_stringency_clash”), or poor shape complementarity following fast initial or slower refinement design steps (“shape1” and “shape2”)

```
<FILTERS>
<OversaturatedHbondAcceptorFilter name="oversat" scorefxn="r15"
max_allowed_oversaturated="0" consider_mainchain_only="false"
/>
<PeptideInternalHbondsFilter name="total_hbonds" hbond_cutoff="3"
exclusion_distance="1" residue_selector="select_peptide"
/>
```

```

<PeptideInternalHbondsFilter
name="total_hbonds_2" hbond_cutoff="3"
exclusion_distance="1" residue_selector="select_peptide"
/>
<ScoreType name="low_stringency_clash" scorefxn="r15"
score_type="fa_rep" threshold="400"
/>
<ShapeComplementarity name="shape1" min_sc="0.63"
min_interface="150" jump="3"
/>
<ShapeComplementarity name="shape2" min_sc="0.66"
min_interface="150" jump="3"
/>
</FILTERS>

```

Listing 3.9 Declaration of the “softdesign” mover, a named instance of a PackRotamersMover used for initial peptide sequence design at relatively low cost in the NDM-1 design script. Backslashes are used to indicate long lines that have been wrapped

```

<PackRotamersMover name="softdesign"
scorefxn="r15_soft" packer_palette="design_palette"
task_operations="use_input_rotamer,no_design_target, \
no_repack_target,no_design_stub,no_repack_stub_dcys, \
no_cys_gly,D_design,L_design,D_hydrophobic_design, \
L_hydrophobic_design"
/>

```

Listing 3.10 The PACKER_PALETTES section of the script used to define the set of amino acid building blocks that will be used for design steps in the NDM-1 design script. Long lines are marked with a backslash to indicate that they have been wrapped

```

<PACKER_PALETTES>
<CustomBaseTypePackerPalette name="design_palette"
additional_residue_types="DALA,DASP,DGLU,DPHE,DHIS,DILE,DLYS,
DLEU,DMET,DASN,DPRO,DGLN,DARG,DSER,DTHR,DVAL,DTRP,DTYR"
/>
</PACKER_PALETTES>

```

Listing 3.11 A subset of the task operations defined in the `TASKOPERATIONS` section of the NDM-1 design script. Task operations are used to define rules for restricting allowed residue types during rotamer optimization steps and can also provide additional rules for configuring the packer (such as directing inclusion of the current side-chain conformation at a position as an allowed rotamer for that position). In turn, task operations may be configured using residue selectors, which define subsets of amino acid residues on which a given task operation will act.

```
<TASKOPERATIONS>
<IncludeCurrent name="use_input_rotamer" />
<OperateOnResidueSubset name="no_repack_target"
  selector="select_target"
>
<PreventRepackingRLT />
</OperateOnResidueSubset>
<OperateOnResidueSubset name="no_design_target"
  selector="select_target"
>
<RestrictToRepackingRLT />
</OperateOnResidueSubset>
<OperateOnResidueSubset name="no_repack_stub_dcys"
  selector="select_stub_dcys"
>
<PreventRepackingRLT />
</OperateOnResidueSubset>
<OperateOnResidueSubset name="no_design_stub"
  selector="select_stub"
>
<RestrictToRepackingRLT />
</OperateOnResidueSubset>
<ProhibitSpecifiedBaseResidueTypes name="no_cys_gly"
base_types="CYS,DCYS,GLY"
  selector="select_design_positions"
/>
<ProhibitResidueProperties name="D_design" properties="L_AA"
  selector="select_D_nonhydrophobic_positions"
/>
<RestrictToSpecifiedBaseResidueTypes name="D_hydrophobic_design"
base_types="DPHE,DILE,DLEU,DMET,DPRO,DVAL,DTRP,DTYR"
  selector="select_D_hydrophobic_positions"
/>
```

```

<ProhibitResidueProperties name="L_design" properties="D_AA"
  selector="select_L_nonhydrophobic_positions"
  />
<RestrictToSpecifiedBaseResidueTypes name="L_hydrophobic_design"
base_types="PHE, ILE, LEU, MET, PRO, VAL, TRP, TYR"
  selector="select_L_hydrophobic_positions"
  />
<OperateOnResidueSubset
name="no_repack_target_far_from_interface"
  selector="select_target_far_from_interface"
>
<PreventRepackingRLT />
</OperateOnResidueSubset>
</TASKOPERATIONS>

```

Listing 3.12 A selection from the RESIDUE_SELECTORS section of the NDM-1 design script used to define rules for selecting residues to configure task operations, movers, filters, or other residue selectors. Long lines are wrapped with a backslash to indicate the wrapping.

```

<RESIDUE_SELECTORS>
<Index name="select_peptide" resnums="234-241" />
<Index name="select_stub" resnums="237-238" />
<Index name="select_stub_dcys" resnums="237" />
<Index name="select_pep_start" resnums="234" />
<Index name="select_pep_end" resnums="241" />
<Not name="select_target" selector="select_peptide" />
<Phi name="select_neg_phi" select_positive_phi="false" />
<Not name="select_pos_phi" selector="select_neg_phi" />
<Index name="select_design_positions" res-
nums="234-236,239-241" />
<And name="select_D_positions"
  selectors="select_design_positions,select_pos_phi"
  />
<And name="select_L_positions"
  selectors="select_design_positions,select_neg_phi"
  />
<Neighborhood name="select_interface" resnums="234-241"
  distance="8.0"
  />
<Not name="not_interface" selector="select_interface" />
<And name="select_target_far_from_interface"
  selectors="select_not_interface,select_target"
  />

```

```

<Layer name="select_core" select_core="true" select_boundary="false"
select_surface="false"
/>
<And name="select_hydrophobic_positions"
selectors="select_design_positions,select_core"
/>
<Not name="select_nonhydrophobic_positions"
selector="select_hydrophobic_positions"
/>
<And name="select_nonhydrophobic_design_positions"
selectors="select_nonhydrophobic_positions,select_design_positions"
/>
<And name="select_L_hydrophobic_positions"
selectors="select_hydrophobic_positions,select_L_positions"
/>
<And name="select_D_hydrophobic_positions"
selectors="select_hydrophobic_positions,select_D_positions"
/>
<And name="select_L_nonhydrophobic_positions"
selectors="select_nonhydrophobic_design_positions, \
select_L_positions"
/>
<And name="select_D_nonhydrophobic_positions"
selectors="select_nonhydrophobic_design_positions, \
select_D_positions"
/>
</RESIDUE_SELECTORS>

```

Listing 3.13 Definition of the “min1” mover, a MinMover used for gradient-descent conformational optimization of peptide side chains in the NDM-1 design script after initial design with the “softdesign” PackRotamersMover

```

<MinMover name="min1" scorefxn="r15_cst" type="dfpmin"
tolerance="0.001" bb="false" chi="false"
>
<MoveMap name="min1_mm" >
<Jump number="1" setting="false" />
<Jump number="2" setting="false" />
<Jump number="3" setting="false" />
<Jump number="4" setting="false" />
<Jump number="5" setting="false" />
<Jump number="6" setting="false" />
<Jump number="7" setting="false" />
<Jump number="8" setting="false" />

```

```

<Jump number="9" setting="false" />
<Span begin="1" end="999" chi="false" bb="false" />
<Span begin="234" end="241" chi="1" bb="false" />
</MoveMap>
</MinMover>

```

Listing 3.14 Definition of FastDesign mover “fdes” used for refinement of initial sequence design in the NDM-1 design script. Long lines are marked with a backslash to indicate that they have been wrapped

```

<FastDesign
name="fdes" repeats="3" scorefxn="r15_highhbond_aacomp_cst"
min_type="dfpmin" packer_palette="design_palette"
task_operations="use_input_rotamer, \
no_repack_target_far_from_interface,no_design_target, \
no_design_stub,no_repack_stub_dcys,no_cys_gly,D_design, \
L_design,D_hydrophobic_design,L_hydrophobic_design"
>
<MoveMap name="fdes_mm" >
<Jump number="1" setting="0" />
<Jump number="2" setting="0" />
<Jump number="3" setting="0" />
<Jump number="4" setting="0" />
<Jump number="5" setting="0" />
<Jump number="6" setting="0" />
<Jump number="7" setting="0" />
<Jump number="8" setting="0" />
<Jump number="9" setting="0" />
<Span begin="1" end="999" chi="1" bb="0" />
<Span begin="234" end="241" chi="1" bb="1" />
</MoveMap>
</FastDesign>

```


Listing 3.15 Definition of the “mc_search” mover, which performs Monte Carlo refinement of the initial NDM-1 inhibitor design, in the MOVERS section of the script

```
<GenericMonteCarlo name="mc_search"  
mover_name="mc_steps" filter_name="mc_filter"  
  trials="500" temperature="0.5"  
/>
```

Listing 3.16 Definition of the “mc_steps” protocol performed at each step in the Monte Carlo search of peptide conformation and sequence space used to refine the initial design

```
<ParsedProtocol name="mc_steps">  
<Add mover="perturb_tors" />  
<Add mover="genkic_perturb" />  
<Add mover="softdesign" />  
<Add mover="min1" />  
<Add mover="connect_termini" />  
<Add filter="oversat" />  
<Add mover="min2" />  
<Add mover="connect_termini" />  
<Add filter="oversat" />  
<Add filter="total_hbonds_2" />  
</ParsedProtocol>
```

Listing 3.17 Definition of the “genkic_perturb” mover used to perturb the conformation of the peptide during the Monte Carlo refinement of the initial design, and the “genkic_perturb_steps” protocol, applied to every solution produced by the “genkic_perturb” mover

```
<ParsedProtocol name="genkic_perturb_steps">  
<Add mover="connect_termini" />  
<Add filter="oversat" />  
<Add filter="total_hbonds_2" />  
</ParsedProtocol>
```

```
<GeneralizedKIC name="genkic_perturb" closure_attempts="5"
pre_selection_mover="genkic_perturb_steps"
stop_when_n_solutions_found="1"
  selector="lowest_rmsd_selector"
selector_scorefunction="r15_highhbond_cst"
>
<AddResidueres_index="239" />
<AddResidueres_index="240" />
<AddResidueres_index="241" />
<AddResidueres_index="234" />
<AddResidueres_index="235" />
<AddResidueres_index="236" />
<SetPivots res1="239" res2="234" res3="236"
  atom1="CA" atom2="CA" atom3="CA"
  />
<CloseBond atom1="C" res1="241" atom2="N" res2="234"
torsion="180"
bondlength="1.328685" angle1="116.2" angle2="121.7" />
<AddPerturber effect="perturb_dihedral" >
<AddAtoms res1="239" atom1="N" res2="239" atom2="CA" />
<AddAtoms res1="239" atom1="CA" res2="239" atom2="C" />
<AddAtoms res1="240" atom1="N" res2="240" atom2="CA" />
<AddAtoms res1="240" atom1="CA" res2="240" atom2="C" />
<AddAtoms res1="241" atom1="N" res2="241" atom2="CA" />
<AddAtoms res1="241" atom1="CA" res2="241" atom2="C" />
<AddAtoms res1="234" atom1="N" res2="234" atom2="CA" />
<AddAtoms res1="234" atom1="CA" res2="234" atom2="C" />
<AddAtoms res1="235" atom1="N" res2="235" atom2="CA" />
<AddAtoms res1="235" atom1="CA" res2="235" atom2="C" />
<AddAtoms res1="236" atom1="N" res2="236" atom2="CA" />
<AddAtoms res1="236" atom1="CA" res2="236" atom2="C" />
<AddValue value="2.5" />
</AddPerturber>
</GeneralizedKIC>
```

References

- Abe K, Kobayashi N, Sode K, Ikebukuro K. Peptide ligand screening of α -synuclein aggregation modulators by in silico panning. *BMC Bioinform.* 2007;8:451. <https://doi.org/10.1186/1471-2105-8-451>.
- Alderson TR, Kay LE. NMR spectroscopy captures the essential role of dynamics in regulating biomolecular function. *Cell.* 2021;184:577–95. <https://doi.org/10.1016/j.cell.2020.12.034>.
- Alexeev Y, Mazanetz MP, Ichihara O, Fedorov DG. GAMESS as a free quantum-mechanical platform for drug research. *Curr Top Med Chem.* 2012;12:2013–33. <https://doi.org/10.2174/156802612804910269>.
- Alford RF, Leaver-Fay A, Jeliakzov JR, et al. The Rosetta all-atom energy function for macromolecular modeling and design. *J Chem Theory Comput.* 2017;13:3031–48. <https://doi.org/10.1021/acs.jctc.7b00125>.
- Arslan PE, Mulligan VK, Ho S, Chakrabartty A. Conversion of Abeta42 into a folded soluble native-like protein using a semi-random library of amphipathic helices. *J Mol Biol.* 2010;396:1284–94. <https://doi.org/10.1016/j.jmb.2009.12.019>.
- Babek T, Ing C, Fingerhuth M. Coarse-grained lattice protein folding on a quantum annealer. *arXiv181100713.* 2018; <https://doi.org/10.48550/arXiv.1811.00713>.
- Baek M, DiMaio F, Anishchenko I, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science.* 2021;373(6557):871–6. <https://doi.org/10.1126/science.abj8754>.
- Bernardi RC, Melo MCR, Schulten K. Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochim Biophys Acta BBA Gen Subj.* 2015;1850:872–7. <https://doi.org/10.1016/j.bbagen.2014.10.019>.
- Bertsimas D, Tsitsiklis J. Simulated annealing. *Stat Sci.* 1993;8:10–5. <https://doi.org/10.1214/ss/1177011077>.
- Bhachoo J, Beuming T. Investigating protein–peptide interactions using the Schrödinger computational suite. In: Schueler-Furman O, London N, editors. *Modeling peptide–protein interactions: methods and protocols.* New York, NY: Springer; 2017. p. 235–54.
- Bhardwaj G, Mulligan VK, Bahl CD, et al. Accurate de novo design of hyperstable constrained peptides. *Nature.* 2016;538:329–35.
- Bochevarov AD, Harder E, Hughes TF, et al. Jaguar: a high-performance quantum chemistry software program with strengths in life and materials sciences. *Int J Quantum Chem.* 2013;113:2110–42. <https://doi.org/10.1002/qua.24481>.
- Bogetti AT, Piston HE, Leung JMG, et al. A twist in the road less traveled: the AMBER ff15ipq-m force field for protein mimetics. *J Chem Phys.* 2020;153:064101. <https://doi.org/10.1063/5.0019054>.
- Boon CL, Frost D, Chakrabartty A. Identification of stable helical bundles from a combinatorial library of amphipathic peptides. *Biopolymers.* 2004;76:244–57. <https://doi.org/10.1002/bip.20074>.
- Bouchiba Y, Cortés J, Schiex T, Barbe S. Molecular flexibility in computational protein design: an algorithmic perspective. *Protein Eng Des Sel.* 2021;34:gzab011. <https://doi.org/10.1093/protein/gzab011>.
- Bovet DP, Crescenzi P. *Introduction to the theory of complexity.* New York: Prentice Hall; 1994.
- Burley SK, Berman HM, Bhikadiya C, et al. RCSB protein data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res.* 2019;47:D464–74. <https://doi.org/10.1093/nar/gky1004>.
- Case DA, Cheatham TE, Darden T, et al. The Amber biomolecular simulation programs. *J Comput Chem.* 2005;26:1668–88. <https://doi.org/10.1002/jcc.20290>.
- Černý V. Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm. *J Optim Theory Appl.* 1985;45:41–51. <https://doi.org/10.1007/BF00940812>.

- Chaudhury S, Lyskov S, Gray JJ. PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinforma Oxf Engl*. 2010;26:689–91. <https://doi.org/10.1093/bioinformatics/btq007>.
- Chaudhury S, Berrondo M, Weitzner BD, et al. Benchmarking and analysis of protein docking performance in Rosetta v3.2. *PLoS One*. 2011;6:e22477. <https://doi.org/10.1371/journal.pone.0022477>.
- Cho A. IBM promises 1000-qubit quantum computer—a milestone—by 2023. *Science*. 2020; <https://doi.org/10.1126/science.abe8122>.
- Combs SA, Deluca SL, Deluca SH, et al. Small-molecule ligand docking into comparative models with Rosetta. *Nat Protoc*. 2013;8:1277–98. <https://doi.org/10.1038/nprot.2013.074>.
- Conway P, Tyka MD, DiMaio F, et al. Relaxation of backbone bond geometry improves protein energy landscape modeling: relaxation of backbone bond geometry. *Protein Sci*. 2014;23:47–55. <https://doi.org/10.1002/pro.2389>.
- Coutsias EA, Seok C, Jacobson MP, Dill KA. A kinematic view of loop closure. *J Comput Chem*. 2004;25:510–28. <https://doi.org/10.1002/jcc.10416>.
- Dang B, Wu H, Mulligan VK, et al. De novo design of covalently constrained mesosize protein scaffolds with unique tertiary structures. *Proc Natl Acad Sci*. 2017;114:10852–7. <https://doi.org/10.1073/pnas.1710695114>.
- Davis IW, Arendall WB, Richardson DC, Richardson JS. The backrub motion: how protein backbone shrugs when a sidechain dances. *Structure*. 2006;14:265–74. <https://doi.org/10.1016/j.str.2005.10.007>.
- Dill KA, Bromberg S, Yue K, et al. Principles of protein folding – a perspective from simple exact models. *Protein Sci*. 1995;4:561–602. <https://doi.org/10.1002/pro.5560040401>.
- DiMaio F, Leaver-Fay A, Bradley P, et al. Modeling symmetric macromolecular structures in Rosetta3. *PLoS One*. 2011;6:e20450. <https://doi.org/10.1371/journal.pone.0020450>.
- Dintzis HM, Symer DE, Dintzis RZ, et al. A comparison of the immunogenicity of a pair of enantiomeric proteins. *Proteins Struct Funct Bioinforma*. 1993;16:306–8. <https://doi.org/10.1002/prot.340160309>.
- Drew K, Renfrew PD, Craven TW, et al. Adding diverse noncanonical backbones to Rosetta: enabling Peptidomimetic design. *PLoS One*. 2013;8:e67051. <https://doi.org/10.1371/journal.pone.0067051>.
- Eleftheriou M, Rayshubski A, Pitera JW, et al. Parallel implementation of the replica exchange molecular dynamics algorithm on blue gene/L. In: *Proceedings 20th IEEE international parallel & distributed processing symposium*. Rhodes Island, Greece: IEEE; 2006. p. 8.
- Fedorov DG. The fragment molecular orbital method: theoretical development, implementation in GAMESS, and applications. *WIREs Comput Mol Sci*. 2017;7:e1322. <https://doi.org/10.1002/wcms.1322>.
- Feynman RP. Simulating physics with computers. *Int J Theor Phys*. 1982;21:467–88. <https://doi.org/10.1007/BF02650179>.
- Fleishman SJ, Leaver-Fay A, Corn JE, et al. RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoS One*. 2011;6:e20161. <https://doi.org/10.1371/journal.pone.0020161>.
- Fock V. Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Z Für Phys*. 1930;61:126–48. <https://doi.org/10.1007/BF01340294>.
- Frappier V, Jenson JM, Zhou J, et al. Tertiary structural motif sequence statistics enable facile prediction and design of peptides that bind anti-apoptotic Bfl-1 and Mcl-1. *Structure*. 2019;27:606–617.e5. <https://doi.org/10.1016/j.str.2019.01.008>.
- Frost DWH, Yip CM, Chakrabartty A. Reversible assembly of helical filaments by de novo designed minimalist peptides. *Biopolymers*. 2005;80:26–33. <https://doi.org/10.1002/bip.20188>.
- Gaillard T, Panel N, Simonson T. Protein side chain conformation predictions with an MMGBSA energy function. *Proteins Struct Funct Bioinforma*. 2016;84:803–19. <https://doi.org/10.1002/prot.25030>.

- Gainza P, Roberts KE, Donald BR. Protein design using continuous Rotamers. *PLoS Comput Biol*. 2012;8:e1002335. <https://doi.org/10.1371/journal.pcbi.1002335>.
- Ghosh AK, Gemma S. HIV-1 protease inhibitors for the treatment of HIV infection and AIDS: design of Saquinavir, Indinavir, and Darunavir. In: *Structure-based design of drugs and other bioactive molecules*. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA; 2015. p. 237–70.
- Gligorijević V, Renfrew PD, Kosciolatek T, et al. Structure-based protein function prediction using graph convolutional networks. *Nat Commun*. 2021, 2021;12(3168) <https://doi.org/10.1038/s41467-021-23303-9>.
- Goldreich O. P, NP, and NP-completeness: the basics of computational complexity. New York: Cambridge University Press; 2010.
- Goodfellow I, Bengio Y, Courville A. *Deep learning*. Cambridge, MA: The MIT Press; 2016.
- Grigoryan G, Reinke AW, Keating AE. Design of protein-interaction specificity gives selective bZIP-binding peptides. *Nature*. 2009;458:859–64. <https://doi.org/10.1038/nature07885>.
- Halgren TA, Murphy RB, Friesner RA, et al. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem*. 2004;47:1750–9. <https://doi.org/10.1021/jm030644s>.
- Hallen MA, Keedy DA, Donald BR. Dead-end elimination with perturbations (DEEPer): a provable protein design algorithm with continuous sidechain and backbone flexibility. *Proteins Struct Funct Bioinforma*. 2013;81:18–39. <https://doi.org/10.1002/prot.24150>.
- Hallen MA, Martin JW, Ojewole A, et al. OSPREY 3.0: open-source protein redesign for you, with powerful new features. *J Comput Chem*. 2018;39:2494–507. <https://doi.org/10.1002/jcc.25522>.
- Hartree DR. The wave mechanics of an atom with a non-coulomb central field. Part I. theory and methods. *Math Proc Camb Philos Soc*. 1928;24:89–110. <https://doi.org/10.1017/S0305004100011919>.
- Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*. 1970;57:97–109. <https://doi.org/10.1093/biomet/57.1.97>.
- Headd JJ, Echols N, Afonine PV, et al. Use of knowledge-based restraints in phenix.Refine to improve macromolecular refinement at low resolution. *Acta Crystallogr D Biol Crystallogr*. 2012;68:381–90. <https://doi.org/10.1107/S0907444911047834>.
- Heo L, Arbour CF, Feig M. Driven to near-experimental accuracy by refinement via molecular dynamics simulations. *Proteins Struct Funct Bioinforma*. 2019;87:1263–75. <https://doi.org/10.1002/prot.25759>.
- Honeyman MC, Brusica V, Stone NL, Harrison LC. Neural network-based prediction of candidate T-cell epitopes. *Nat Biotechnol*. 1998;16:966–9. <https://doi.org/10.1038/nbt1098-966>.
- Hossein-zadeh P, Bhardwaj G, Mulligan VK, et al. Comprehensive computational design of ordered peptide macrocycles. *Science*. 2017;358:1461–6. <https://doi.org/10.1126/science.aap7577>.
- Hossein-zadeh P, Watson PR, Craven TW, et al. Anchor extension: a structure-guided approach to design cyclic peptides targeting enzyme active sites. *Nat Commun*. 2021;12:3384. <https://doi.org/10.1038/s41467-021-23609-8>.
- Huang P-S, Ban Y-EA, Richter F, et al. RosettaRemodel: a generalized framework for flexible backbone protein design. *PLoS One*. 2011;6:e24109. <https://doi.org/10.1371/journal.pone.0024109>.
- Huang P-S, Oberdorfer G, Xu C, et al. High thermodynamic stability of parametrically designed helical bundles. *Science*. 2014;346:481–5. <https://doi.org/10.1126/science.1257481>.
- Huang J, Rauscher S, Nawrocki G, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods*. 2017;14:71–3. <https://doi.org/10.1038/nmeth.4067>.
- Hwang TJ, Carpenter D, Lauffenburger JC, et al. Failure of investigational drugs in late-stage clinical development and publication of trial results. *JAMA Intern Med*. 2016;176:1826–33. <https://doi.org/10.1001/jamainternmed.2016.6008>.

- Joshi S, Chen L, Winter MB, et al. The rational Design of Therapeutic Peptides for aminopeptidase N using a substrate-based approach. *Sci Rep.* 2017;7:1424. <https://doi.org/10.1038/s41598-017-01542-5>.
- Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021;596:583–9. <https://doi.org/10.1038/s41586-021-03819-2>.
- Källberg M, Wang H, Wang S, et al. Template-based protein structure modeling using the RaptorX web server. *Nat Protoc.* 2012;7:1511–22. <https://doi.org/10.1038/nprot.2012.085>.
- Kästner J. Umbrella sampling: Umbrella sampling. *Wiley Interdiscip Rev Comput Mol Sci.* 2011;1:932–42. <https://doi.org/10.1002/wcms.66>.
- Khatib F, Cooper S, Tyka MD, et al. Algorithm discovery by protein folding game players. *Proc Natl Acad Sci U S A.* 2011;108:18949–53. <https://doi.org/10.1073/pnas.1115898108>.
- King DT, Worrall LJ, Gruninger R, Strynadka NCJ. New Delhi metallo- β -lactamase: structural insights into β -lactam recognition and inhibition. *J Am Chem Soc.* 2012;134:11362–5. <https://doi.org/10.1021/ja303579d>.
- Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by simulated annealing. *Science.* 1983;220:671–80. <https://doi.org/10.1126/science.220.4598.671>.
- Kitaura K, Ikeo E, Asada T, et al. Fragment molecular orbital method: an approximate computational method for large molecules. *Chem Phys Lett.* 1999;313:701–6. [https://doi.org/10.1016/S0009-2614\(99\)00874-X](https://doi.org/10.1016/S0009-2614(99)00874-X).
- Koehler Leman J, Weitzner BD, Renfrew PD, et al. Better together: elements of successful scientific software development in a distributed collaborative community. *PLoS Comput Biol.* 2020;16:e1007507. <https://doi.org/10.1371/journal.pcbi.1007507>.
- Koga N, Tatsumi-Koga R, Liu G, et al. Principles for designing ideal protein structures. *Nature.* 2012;491:222–7. <https://doi.org/10.1038/nature11600>.
- Kohn W, Sham LJ. Self-consistent equations including exchange and correlation effects. *Phys Rev.* 1965;140:A1133–8. <https://doi.org/10.1103/PhysRev.140.A1133>.
- Kuhlman B, Dantas G, Ireton GC, et al. Design of a novel globular protein fold with atomic-level accuracy. *Science.* 2003;302:1364–8. <https://doi.org/10.1126/science.1089427>.
- Lange OF, Rossi P, Sgourakis NG, et al. Determination of solution structures of proteins up to 40 kDa using CS-Rosetta with sparse NMR data from deuterated samples. *Proc Natl Acad Sci.* 2012;109:10873–8. <https://doi.org/10.1073/pnas.1203013109>.
- Lawrence MC, Colman PM. Shape complementarity at protein/protein interfaces. *J Mol Biol.* 1993;234:946–50. <https://doi.org/10.1006/jmbi.1993.1648>.
- Leaver-Fay A, Tyka M, Lewis SM, et al. Rosetta3. In: *Methods in enzymology*. In: Elsevier; 2011. p. 545–74.
- Leeb-Lundberg F, Snowman A, Olsen RW. Barbiturate receptor sites are coupled to benzodiazepine receptors. *Proc Natl Acad Sci U S A.* 1980;77:7468–72. <https://doi.org/10.1073/pnas.77.12.7468>.
- Leelananda SP, Lindert S. Using NMR chemical shifts and Cryo-EM density restraints in iterative Rosetta-MD protein structure refinement. *J Chem Inf Model.* 2020;60:2522–32. <https://doi.org/10.1021/acs.jcim.9b00932>.
- Limongelli V, Bonomi M, Parrinello M. Funnel metadynamics as accurate binding free-energy method. *Proc Natl Acad Sci.* 2013;110:6358–63. <https://doi.org/10.1073/pnas.1303186110>.
- Lombardino JG, Lowe JA. The role of the medicinal chemist in drug discovery -- then and now. *Nat Rev Drug Discov.* 2004;3:853–62. <https://doi.org/10.1038/nrd1523>.
- Löscher W, Rogawski MA. How theories evolved concerning the mechanism of action of barbiturates. *Epilepsia.* 2012;53:12–25. <https://doi.org/10.1111/epi.12025>.
- Maguire JB, Haddox HK, Strickland D, et al. Perturbing the energy landscape for improved packing during computational protein design. *Proteins Struct Funct Bioinforma.* 2021;89:436–49. <https://doi.org/10.1002/prot.26030>.
- Mandell DJ, Coutsias EA, Kortemte T. Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nat Methods.* 2009;6:551–2. <https://doi.org/10.1038/nmeth0809-551>.

- Manousiouthakis VI, Deem MW. Strict detailed balance is unnecessary in Monte Carlo simulation. *J Chem Phys.* 1999;110:2753–6. <https://doi.org/10.1063/1.477973>.
- McGeoch C, Farré P. The D-wave advantage system: an overview. D-Wave Systems Inc.; 2020.
- Metropolis N, Rosenbluth AW, Rosenbluth MN, et al. Equation of state calculations by fast computing machines. *J Chem Phys.* 1953;21:1087–92. <https://doi.org/10.1063/1.1699114>.
- Mironov V, Alexeev Y, Mulligan VK, Fedorov DG. A systematic study of minima in alanine dipeptide. *J Comput Chem.* 2019;40:297–309. <https://doi.org/10.1002/jcc.25589>.
- Møller C, Plesset MS. Note on an approximation treatment for many-electron systems. *Phys Rev.* 1934;46:618–22. <https://doi.org/10.1103/PhysRev.46.618>.
- Mulligan VK. The emerging role of computational design in peptide macrocycle drug discovery. *Expert Opin Drug Discov.* 2020;15:833–52. <https://doi.org/10.1080/17460441.2020.1751117>.
- Mulligan VK. Current directions in combining simulation-based macromolecular modeling approaches with deep learning. *Expert Opin Drug Discov.* 2021;16(9):1025–44. <https://doi.org/10.1080/17460441.2021.1918097>.
- Mulligan VK, Melo H, Merritt HI, et al. Designing peptides on a quantum computer. *bioRxiv.* 2019:752485. <https://doi.org/10.1101/752485>.
- Mulligan VK, Kang CS, Sawaya MR, et al. Computational design of mixed chirality peptide macrocycles with internal symmetry. *Protein Sci.* 2020;29:2433–45. <https://doi.org/10.1002/pro.3974>.
- Mulligan VK, Workman S, Sun T, et al. Computationally designed peptide macrocycle inhibitors of New Delhi metallo- β -lactamase 1. *Proc Natl Acad Sci.* 2021;118:e2012800118. <https://doi.org/10.1073/pnas.2012800118>.
- Murphy RB, Philipp DM, Friesner RA. A mixed quantum mechanics/molecular mechanics (QM/MM) method for large-scale modeling of chemistry in protein environments. *J Comput Chem.* 2000;21:1442–57. [https://doi.org/10.1002/1096-987X\(200012\)21:16<1442::AID-JCC3>3.0.CO;2-O](https://doi.org/10.1002/1096-987X(200012)21:16<1442::AID-JCC3>3.0.CO;2-O).
- Nerli S, Sgourakis NG. CS-ROSETTA. *Methods Enzymol.* 2019;614:321–62. <https://doi.org/10.1016/bs.mie.2018.07.005>.
- O’Leary NA, Wright MW, Brister JR, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 2016;44:D733–45. <https://doi.org/10.1093/nar/gkv1189>.
- Ouali A, Allouche D, de Givry S, et al. Variable neighborhood search for graphical model energy minimization. *Artif Intell.* 2020;278:103194. <https://doi.org/10.1016/j.artint.2019.103194>.
- Park H, Bradley P, Greisen P, et al. Simultaneous optimization of biomolecular energy functions on features from small molecules and macromolecules. *J Chem Theory Comput.* 2016;12:6201–12. <https://doi.org/10.1021/acs.jctc.6b00819>.
- Parrish RM, Burns LA, Smith DGA, et al. Psi4 1.1: an open-source electronic structure program emphasizing automation, advanced libraries, and interoperability. *J Chem Theory Comput.* 2017;13:3185–97. <https://doi.org/10.1021/acs.jctc.7b00174>.
- Patel S, Brooks CL. CHARMM fluctuating charge force field for proteins: I parameterization and application to bulk organic liquid simulations. *J Comput Chem.* 2004;25:1–16. <https://doi.org/10.1002/jcc.10355>.
- Perdomo-Ortiz A, Dickson N, Drew-Brook M, et al. Finding low-energy conformations of lattice protein models by quantum annealing. *Sci Rep.* 2012;2:571. <https://doi.org/10.1038/srep00571>.
- Phillips JC, Braun R, Wang W, et al. Scalable molecular dynamics with NAMD. *J Comput Chem.* 2005;26:1781–802. <https://doi.org/10.1002/jcc.20289>.
- Pierce NA, Winfree E. Protein design is NP-hard. *Protein Eng.* 2002;15:779–82. <https://doi.org/10.1093/protein/15.10.779>.
- Pronk S, Páll S, Schulz R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics.* 2013;29:845–54. <https://doi.org/10.1093/bioinformatics/btt055>.

- Renfrew PD, Choi EJ, Bonneau R, Kuhlman B. Incorporation of noncanonical amino acids into Rosetta and use in computational protein-peptide Interface design. *PLoS One*. 2012;7:e32637. <https://doi.org/10.1371/journal.pone.0032637>.
- Roy Burman SS, Nance ML, Jeliazkov JR, et al. Novel sampling strategies and a coarse-grained score function for docking homomers, flexible heteromers, and oligosaccharides using Rosetta in CAPRI rounds 37–45. *Proteins*. 2020;88:973–85. <https://doi.org/10.1002/prot.25855>.
- Ruffini M, Vucinic J, de Givry S, et al. Guaranteed diversity and optimality in cost function network based computational protein design methods. *Algorithms*. 2021;14:168. <https://doi.org/10.3390/a14060168>.
- Senior AW, Evans R, Jumper J, et al. Protein structure prediction using multiple deep neural networks in the 13th critical assessment of protein structure prediction (CASP13). *Proteins Struct Funct Bioinforma*. 2019;87:1141–8. <https://doi.org/10.1002/prot.25834>.
- Senior AW, Evans R, Jumper J, et al. Improved protein structure prediction using potentials from deep learning. *Nature*. 2020;577:706–10. <https://doi.org/10.1038/s41586-019-1923-7>.
- Simonovsky M, Meyers J. DeeplyTough: learning structural comparison of protein binding sites. *J Chem Inf Model*. 2020;60:2356–66. <https://doi.org/10.1021/acs.jcim.9b00554>.
- Simons KT, Bonneau R, Ruczinski I, Baker D. Ab initio protein structure prediction of CASP III targets using ROSETTA. *Proteins Suppl*. 1999;3:171–6. [https://doi.org/10.1002/\(sici\)1097-0134\(1999\)37:3+<171::aid-prot21>3.3.co;2-q](https://doi.org/10.1002/(sici)1097-0134(1999)37:3+<171::aid-prot21>3.3.co;2-q).
- Sindhikara D, Spronk SA, Day T, et al. Improving accuracy, diversity, and speed with prime macrocycle conformational sampling. *J Chem Inf Model*. 2017;57:1881–94. <https://doi.org/10.1021/acs.jcim.7b00052>.
- Sindhikara D, Wagner M, Gkeka P, et al. Automated design of macrocycles for therapeutic applications: From small molecules to peptides and proteins. *J Med Chem*. 2020; <https://doi.org/10.1021/acs.jmedchem.0c01500>.
- Slater JC. Note on Hartree's method. *Phys Rev*. 1930;35:210–1. <https://doi.org/10.1103/PhysRev.35.210.2>.
- Smith CA, Kortemme T. Backrub-like backbone simulation recapitulates natural protein conformational variability and improves mutant side-chain prediction. *J Mol Biol*. 2008;380:742–56. <https://doi.org/10.1016/j.jmb.2008.05.023>.
- Song Y, DiMaio F, Wang RY-R, et al. High-resolution comparative modeling with RosettaCM. *Structure*. 2013;21:1735–42. <https://doi.org/10.1016/j.str.2013.08.005>.
- Stein A, Kortemme T. Improvements to robotics-inspired conformational sampling in rosetta. *PLoS ONE*. 2013;8. <https://doi.org/10.1371/journal.pone.0063090>.
- Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett*. 1999;314:141–51. [https://doi.org/10.1016/S0009-2614\(99\)01123-9](https://doi.org/10.1016/S0009-2614(99)01123-9).
- Tan YZ, Carragher B. Seeing atoms: single-particle cryo-EM breaks the atomic barrier. *Mol Cell*. 2020;80:938–9. <https://doi.org/10.1016/j.molcel.2020.11.043>.
- Torrie GM, Valleau JP. Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling. *J Comput Phys*. 1977;23:187–99. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).
- Traoré S, Allouche D, André I, et al. A new framework for computational protein design through cost function network optimization. *Bioinforma Oxf Engl*. 2013;29:2129–36. <https://doi.org/10.1093/bioinformatics/btt374>.
- Valiev M, Bylaska EJ, Govind N, et al. NWChem: a comprehensive and scalable open-source solution for large scale molecular simulations. *Comput Phys Commun*. 2010;181:1477–89. <https://doi.org/10.1016/j.cpc.2010.04.018>.
- Vanhee P, van der Sloot AM, Verschuereen E, et al. Computational design of peptide ligands. *Trends Biotechnol*. 2011;29:231–9. <https://doi.org/10.1016/j.tibtech.2011.01.004>.
- Wang CK, Swedberg JE, Northfield SE, Craik DJ. Effects of cyclization on peptide backbone dynamics. *J Phys Chem B*. 2015;119:15821–30. <https://doi.org/10.1021/acs.jpcc.5b11085>.

- Wang RY-R, Song Y, Barad BA, et al. Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta. *eLife*. 2016;5:e17219. <https://doi.org/10.7554/eLife.17219>.
- Wang S, Sun S, Xu J. Analysis of deep learning methods for blind protein contact prediction in CASP12. *Proteins*. 2018;86:67–77. <https://doi.org/10.1002/prot.25377>.
- Watkins AM, Geniesse C, Kladwang W, et al. Blind prediction of noncanonical RNA structure at atomic accuracy. *Sci Adv*. 2018;4:eaar5316. <https://doi.org/10.1126/sciadv.aar5316>.
- Weinstock MT, Francis JN, Redman JS, Kay MS. Protease-resistant peptide design-empowering nature's fragile warriors against HIV. *Biopolymers*. 2012;98:431–42. <https://doi.org/10.1002/bip.22073>.
- Xu D, Zhang Y. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys J*. 2011;101:2525–34. <https://doi.org/10.1016/j.bpj.2011.10.024>.
- Xu D, Zhang Y. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins*. 2012;80:1715–35. <https://doi.org/10.1002/prot.24065>.
- Yachnin BJ, Mulligan VK, Khare SD, Bailey-Kellogg C. MHCEpitopeEnergy, a flexible rosetta-based biotherapeutic deimmunization platform. *J Chem Inf Model*. 2021;61(5):2368–82. <https://doi.org/10.1021/acs.jcim.1c00056>.
- Yang J, Yan R, Roy A, et al. The I-TASSER suite: protein structure and function prediction. *Nat Methods*. 2015;12:7–8. <https://doi.org/10.1038/nmeth.3213>.
- Yang J, Anishchenko I, Park H, et al. Improved protein structure prediction using predicted interresidue orientations. *Proc Natl Acad Sci*. 2020;117:1496–503. <https://doi.org/10.1073/pnas.1914677117>.
- Zhang C, Mortuza SM, He B, et al. Template-based and free modeling of I-TASSER and QUARK pipelines using predicted contact maps in CASP12. *Proteins Struct Funct Bioinforma*. 2018;86:136–51. <https://doi.org/10.1002/prot.25414>.
- Zhou J, Panaitiu AE, Grigoryan G. A general-purpose protein design framework based on mining sequence–structure relationships in known protein structures. *Proc Natl Acad Sci*. 2020;117:1059–68. <https://doi.org/10.1073/pnas.1908723117>.

Chapter 4

Strategies to Optimize Peptide Stability and Prolong Half-Life



Sophia M. Shi and Li Di

Contents

4.1	Introduction.....	165
4.2	Strategies to Improve Proteolytic Peptide Stability and Prolong Half-Life.....	167
4.3	Replacing L-Amino Acids with D-Amino Acids or Unnatural Amino Acids.....	168
4.4	Other Backbone Modification Strategies.....	171
4.5	Protection of N- and C-Termini.....	171
4.6	Cyclization.....	172
4.7	Increase Molecular Weight.....	173
4.8	Noncovalent π - π Interactions.....	175
4.9	Enzyme Inhibition.....	175
4.10	Flip-Flop Kinetics.....	175
4.11	Methods to Assess Peptide Stability.....	176
4.12	Blood, Serum, and Plasma Stability Assays.....	176
4.13	Hepatic Stability Assays.....	176
4.14	Kidney and Tissue Stability Assays.....	177
4.15	Gastric Intestinal Stability Assays.....	177
4.16	Conclusions.....	178
	References.....	178

Abstract Peptide drugs represent 5% of the global pharmaceutical market, but growing twice as fast as the rest of the drug market. The development of peptide therapeutics is challenging due to their low stability, short half-life, and poor oral bioavailability. However, peptides typically have exquisite potency, selectivity, and low toxicity, making them particularly attractive for certain disease targets. Significant technological innovations have enabled the rapid advancement of

S. M. Shi
Department of Chemistry, Stanford University, Stanford, CA, USA
e-mail: smshi@stanford.edu

L. Di (✉)
Pharmacokinetics, Dynamics and Metabolism, Pfizer Worldwide Research and Development,
Groton, CT, USA
e-mail: Li.Di@Pfizer.com

peptide therapeutics to the clinic. Here, strategies to improve peptide proteolytic stability and prolong half-life are discussed. Structural modifications are highly effective for enhancing peptide stability, including replacing the natural L-amino acids with D- or unnatural amino acids, peptide backbone modifications, protecting N- and C-termini, and cyclization. In vitro and in vivo assays are available to assess peptide stability and develop structure-stability relationships to enable the design of more stable peptides. Peptides will continue to play an important role in filling the gaps between small molecule drugs and protein therapeutics.

Keywords Peptide stability · Proteolysis · Half-life · Peptidase · Cyclic peptides · Stability assays

Abbreviations

ADA	antidrug antibody
BBMV	brush border membrane vesicles
CaV	voltage-gated calcium channel
CNS	central nervous system
CYP	cytochrome P450
DPP-4	dipeptidyl peptidase 4
EDTA	ethylenediaminetetraacetic acid
F _a	fraction absorbed
Fc	fragment crystallizable
GI	gastrointestinal
GIP	glucose-dependent insulinotropic polypeptide
GLP-1	glucagon-like-peptide-1
HIV	human immunodeficiency virus
HPLC	High-performance liquid chromatography
LC-MS	liquid chromatography–mass spectrometry
MW	molecular weight
PK/PD	pharmacokinetics/pharmacodynamics
PMSF	phenylmethylsulfonyl fluoride
SAR	structure-activity relationship
SC	subcutaneously
SGF	simulated gastric fluid
SIF	simulated intestinal fluid
T2DM	type 2 diabetes mellitus
TMDD	target-mediated drug disposition
UGT	uridine 5'-diphospho-glucuronosyltransferase

4.1 Introduction

Since the discovery of insulin as the first peptide drug in the early 1920s, peptide therapeutics have established a unique niche in the pharmaceutical industry. A variety of diseases are being treated using peptide drugs, such as diabetes, cancer, osteoporosis, multiple sclerosis, HIV infection, and chronic pain (Muttenthaler et al. 2021; Di 2015). Cancer and diabetes are currently the two major disease areas of focus for peptide drug discovery. At the present moment, over 80 peptide drugs have been approved worldwide, more than 150 peptides are in clinical development, and 400–600 peptides are in preclinical development (Muttenthaler et al. 2021). Blockbuster status has been achieved for several peptide drugs, including dulaglutide, liraglutide, and semaglutide, with annual sales exceeding \$1 billion (Muttenthaler et al. 2021). The current peptide drug market is valued at ~\$28 billion and is expected to reach ~\$51 billion in 2026 (<https://www.globenewswire.com/news-release/2021/03/01/2184060/0/en/The-Peptide-Therapeutics-Market-is-valued-at-approximately-USD-28-510-60-million-in-2020-and-is-expected-to-witness-a-revenue-of-USD-51-360-30-million-in-2026-with-a-CAGR-of-9-66-o.html>) ([https://www.researchandmarkets.com/reports/5265155/peptide-therapeutics-market-growth-trends?utm_source=BW&utm_medium=PressRelease&utm_code=nbjvs&utm_campaign=1507648%20PLUS_SPI%20-PLUS_%20SPI%20Global%20PLUS_SPI%20Peptide%20PLUS_SPI%20Therapeutics%20PLUS_SPI%20Market%20PLUS_SPI%20\(2021%20PLUS_SPI%20to%20PLUS_SPI%202026\)%20PLUS_SPI%20-PLUS_%20SPI%20Growth%2c%20PLUS_SPI%20Trends%2c%20PLUS_SPI%20COVID-19%20PLUS_SPI%20Impact%2c%20PLUS_SPI%20and%20PLUS_SPI%20Forecasts&utm_exec=jamu273prd%E2%80%99](https://www.researchandmarkets.com/reports/5265155/peptide-therapeutics-market-growth-trends?utm_source=BW&utm_medium=PressRelease&utm_code=nbjvs&utm_campaign=1507648%20PLUS_SPI%20-PLUS_%20SPI%20Global%20PLUS_SPI%20Peptide%20PLUS_SPI%20Therapeutics%20PLUS_SPI%20Market%20PLUS_SPI%20(2021%20PLUS_SPI%20to%20PLUS_SPI%202026)%20PLUS_SPI%20-PLUS_%20SPI%20Growth%2c%20PLUS_SPI%20Trends%2c%20PLUS_SPI%20COVID-19%20PLUS_SPI%20Impact%2c%20PLUS_SPI%20and%20PLUS_SPI%20Forecasts&utm_exec=jamu273prd%E2%80%99)).

Although peptide drugs represent only 5% of the global pharmaceutical market, they are growing twice as fast as the rest of the drug market and are predicted to soon occupy a larger niche (Muttenthaler et al. 2021). The growth of peptide therapeutics has been greatly accelerated by technological advances, including solid-state peptide synthesis, recombinant peptide production, genetic engineering, display technology, HPLC peptide purification, and peptide delivery systems. Peptide therapeutics will continue to be an integral part of drug discovery and development.

Peptides have the advantages of exquisite potency, excellent target selectivity, low toxicity, and relatively low molecular weight compared to therapeutic proteins (Table 4.1). Peptides are defined as compounds that consist of 2 to 40 amino acids, whereas proteins are defined to contain 50 or more amino acids (Booka et al. 2021; Craik et al. 2013; <https://www.fda.gov/media/135421/download>). The unique physicochemical and pharmacological properties of peptides (Table 4.1) make them attractive candidates for certain challenging disease targets compared to small molecules or therapeutic proteins. One successful example of a peptide drug is ziconotide, a peptidic inhibitor of voltage-gated calcium (CaV) 2.2 channels derived

Table 4.1 Comparisons of small molecules, peptides, and biologics (Muttenthaler et al. 2021; Di 2015; Makurvet 2021; Mocsai et al. 2014)

Properties	Small molecules	Peptides	Biologics
Percentage in global pharmaceutical market	75%	5%	20%
MW	Low	Medium	High
Structure	Well-defined and completely characterizable	Well-defined and completely characterizable	Complex, heterogeneous, and not entirely characterizable
Membrane permeability	Mostly high	Low	Low
Stability	Mostly high	Low	Low
Oral bioavailability	Mostly high	Low	Low
Delivery routes	Mostly oral	Mostly parental	Mostly parental
Elimination Mechanisms	Mostly metabolism by CYP/UGT/ others	Mostly proteolysis by peptidases and renal clearance	TMDD, lysosomal degradation, Fc, and ADA-mediated clearance, by nucleases, peptidases, proteinases, and hydrolases
Half-life	Moderate	Very short to long	Long
Distribution	Via blood circulation, easily distributed	Mostly via blood circulation, moderately distributed	Via blood circulation and lymphatics, limited distribution
Disease targets	Broad coverage	Mostly limited to extracellular targets	Mostly limited to extracellular targets
Target binding affinity	Moderate	High	High
Target specificity	Moderate	High	High
Mechanism of action	May not known	Well understood	Well understood
Off-target toxicity	Moderate	Low	Low
Immunotoxicity	Low	Moderate	High
Drug-drug interactions	Moderate-high risk	Low risk	Low risk
Impact of generics	High	Moderate	Moderate-low
Production	Chemical synthesis, relatively easier to make	Chemical synthesis or isolation from living systems, relatively more challenging to make	Living cell culture, challenging to make
Process dependency	Mostly process-independent	Mostly process-independent	Strongly process-dependent
Manufacturing cost	Low	Moderate	High

TMDD target-mediated drug disposition, Fc fragment crystallizable, ADA antidrug antibody

from the venom of a predatory marine cone snail (Muttenthaler et al. 2021). Ziconotide is 1000-fold more potent than morphine and does not possess the addictive properties associated with opioid pain medications. The major limitations of peptides as therapeutic drugs are poor oral bioavailability, susceptibility to proteolysis, short half-life, and low cell membrane permeability (Table 4.1). As a result, 90% of peptide drugs are delivered by injection rather than via oral administration. In addition, due to low cell membrane permeability and poor central nervous system (CNS) penetration, peptide therapeutics typically focus on extracellular and peripheral disease targets. Although the lack of oral bioavailability remains highly challenging for peptide drug development, there are some peptides that are orally bioavailable. The most well-known example is cyclosporin A, which has been shown to have a fraction absorbed (F_a) greater than 86% (Wu et al. 1995; Chiu et al. 2003). Permeation enhancers can also help to achieve oral absorption of peptides (Maher et al. 2016). A recent example is the development of orally available semaglutide, a GLP-1 (glucagon-like-peptide-1) agonist for the treatment of type 2 diabetes mellitus (T2DM). By using permeation enhancer SNAC (sodium *N*-[8-(2-hydroxybenzoyl) aminocaprylate]), transcellular permeability of semaglutide in the stomach increased leading to ~4% oral bioavailability (Knudsen and Lau 2019; Buckley et al. 2018). In addition, cyclic and conformationally constrained peptides are promising candidates to achieve oral bioavailability (Nielsen et al. 2017). Reducing hydrogen-bond donors and flexibility generally favors oral absorption of peptides.

Many great reviews are available on the design and development of peptide therapeutics (Muttenthaler et al. 2021; Di 2015; Yao et al. 2018; Zaman et al. 2019; Lee et al. 2019; Davenport et al. 2020; Craik et al. 2013; Jing and Jin 2020). Improving peptide proteolytic stability is one of the major goals for peptide drug discovery and development. This chapter focuses on the discussion of structural modification strategies to optimize peptide stability and prolong half-life. Approaches on formulations and novel peptide delivery systems will be covered in other chapters of the book.

4.2 Strategies to Improve Proteolytic Peptide Stability and Prolong Half-Life

One of the greatest challenges for peptide therapeutics is their high elimination rate and short half-life in systemic circulation. Many native peptides have half-lives of only a few minutes in the blood, which greatly limits their use as therapeutic agents. Several mechanisms are involved in clearing peptides from the blood including renal clearance (glomerular filtration and active secretion), proteolysis in blood and tissues (e.g., kidney, liver, brain), and endocytosis and degradation by the proteasome. Strategies to reduce peptide renal clearance have been well-documented (Di 2015; Zaman et al. 2019; Wu and Huang 2018). Here, we focus the discussion on

peptide clearance by proteolysis. Peptides are susceptible to proteolysis by peptidases due to the characteristic peptide bonds in their structures. Over 1,500 different peptidases exist in the human body that can cleave peptides, such as aminopeptidases, dipeptidases, dipeptidyl-peptidases and tripeptidyl-peptidases, peptidyl-dipeptidases, serine-type carboxypeptidases, metallo-carboxy peptidases, cysteine-type carboxypeptidases, and serine endopeptidases (Yao et al. 2018; Werle and Bernkop-Schnuerch 2006). Peptidases are expressed all over the body including in tissues (e.g., intestine, liver, kidney, brain, lung, pancreas, skeletal muscle, heart, spleen, skin, nasal epithelial cells, placenta, ovary, uterus, testis, prostate, thymus, and adrenal glands) and body fluids (plasma, blood, gastrointestinal mucosa, pancreatic juice) as membrane-bound or soluble enzymes. Peptides can be cleaved by multiple peptidases in complex steps to yield smaller peptides and amino acids as final products, which can be used as building blocks for synthesis of various proteins. Endopeptidases (e.g., pepsin and elastase) cleave peptide bonds near the middle of the peptide, and exopeptidases (e.g., aminopeptidases and carboxypeptidases) hydrolyze terminal peptide bonds. Several software programs are available to predict cleavage sites of peptides (Di 2015). In practice, the peptidases involved in catalyzing hydrolysis of a peptide are not always fully characterized. Instead, the hydrolytic products are typically identified using liquid chromatography–mass spectrometry (LC-MS) from *in vitro* incubation of various matrices or *in vivo* samples (Liao et al. 2015; Sharma et al. 2013; Jyrkas and Tolonen 2021). This information can help pinpoint the labile sites and guide structural modifications of peptides to minimize enzymatic degradation. A number of strategies have been developed to improve peptide stability via a variety of modifications using medicinal chemistry approaches. It is important to maintain or improve potency and circumvent toxicity while improving peptide stability through structural modification. Some of the commonly applied strategies to improve peptide proteolytic stability and prolong half-life are discussed here.

4.3 Replacing L-Amino Acids with D-Amino Acids or Unnatural Amino Acids

Many peptides containing natural L-amino acids are susceptible to enzymatic degradation. Some amino acids are more labile than others, e.g., aspartic acid is sensitive to isomerization, asparagine is prone to deamination, and methionine is susceptible to oxidation (Boschi-Muller et al. 2008; Wakankar and Borchardt 2006; Jing and Jin 2020). One effective strategy to increase proteolytic peptide stability is to replace the natural L-amino acids with the unnatural D- or modified amino acids (Fig 4.1 and Table 4.2), as the peptidases may not be able to recognize the unnatural amino acids for cleavage. However, a simple substitution of all the L-amino acids with D-amino acids is generally ineffective, because the resulting peptide conformation and side chain orientation can prevent the correct binding geometry and

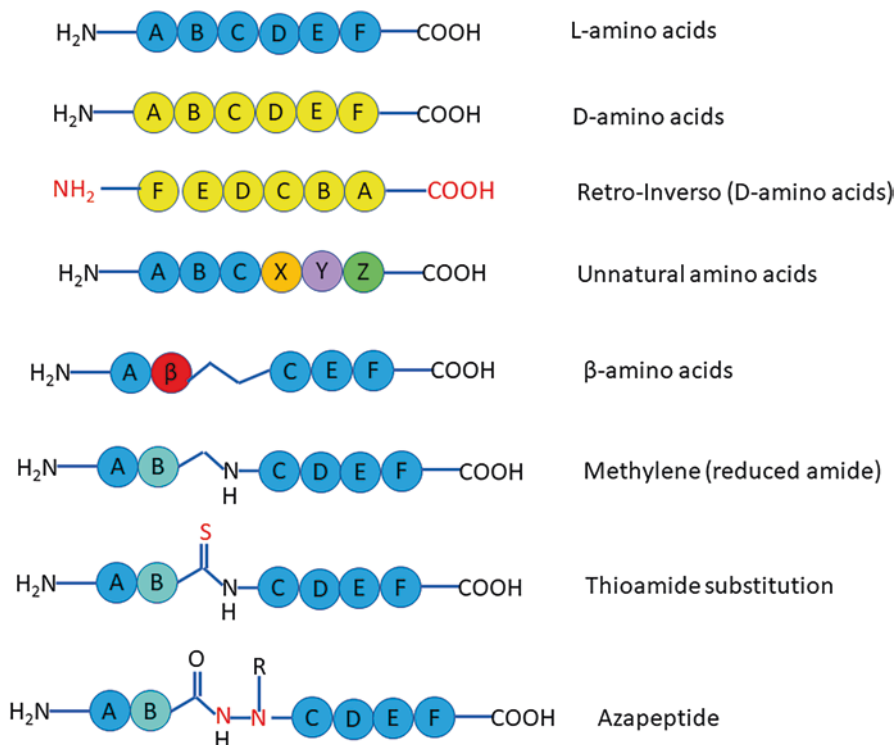
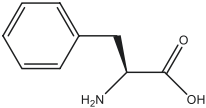
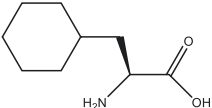
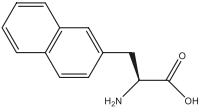
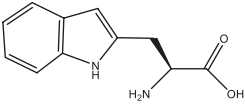
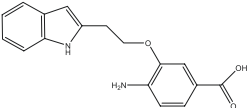
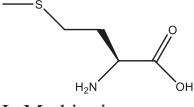
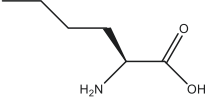
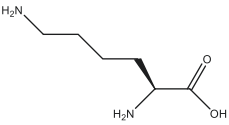
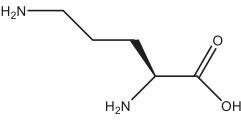
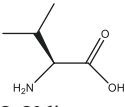
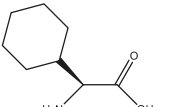
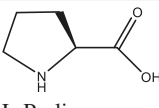
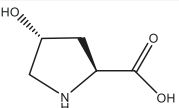


Fig. 4.1 Selected backbone modification strategies to stabilize peptides. (Booka et al. 2021; Gentilucci et al. 2010)

have poor potency towards the target (Evans et al. 2020). Retro-inversion (all the L-amino acids are replaced with D-amino acids, and the sequence is reversed as well (Fig. 4.1)) is a strategy developed to help restore activity by revising the sequence of D-peptides (Li et al. 2015; Ben-Yedidia et al. 2002).

There are many examples of peptide drugs that have successfully been developed by replacing L-amino acids with D-amino acids. For example, octreotide is a drug for the treatment of gastrointestinal tumors. It is a synthetic peptide based on the amino acid sequence of the endogenous hormone somatostatin, which has a very short half-life of only a few minutes. Octreotide was developed by shortening the amino acid sequence of somatostatin from 14 to 8 amino acids and replacing all L-amino acids with D-amino acids. These modifications subsequently led to significant enhancement of enzymatic stability and extended plasma half-life to 1.5 h (Harris 1994; Werle and Bernkop-Schnuerch 2006). Another example is the development of desmopressin (synthetic antidiuretic hormone) from vasopressin (native hormone), which normally has a short half-life of 10–35 min in humans. Replacing the L-Arg in vasopressin with D-Arg led to the discovery of desmopressin with an improved half-life of 3.7 h. Furthermore, it was reported that the replacement of

Table 4.2 Examples of mimics for natural amino acids (Jing and Jin 2020)

Natural amino acids	Unnatural amino acid mimics	
 L-Phenylalanine	 L-Cyclohexylalanine	 L-Naphthylalanine
 L-Tryptophan	 3-(2-(1 <i>H</i> -indol-2-yl)ethoxy)-4-aminobenzoic acid	
 L-Methionine	 L-Norleucine	
 L-Lysine	 L-Ornithine	
 L-Valine	 L-Cyclohexylglycine	
 L-Proline	 L-4-hydroxyproline	

L-histidine and L-cysteine residues in the $\alpha 5\beta 1$ integrin inhibitor peptide (anticancer) with the respective D-amino acids yielded a peptide with significantly improved potency and systemic stability (Veine et al. 2014). There are some cases where replacement of an L-amino acid with the D-amino acid may not improve half-life or yield an even shorter half-life (Werle and Bernkop-Schnuerch 2006; Darlak et al. 1988; Rafferty et al. 1988). Stability studies must be conducted to verify if the changes improve peptide stability. Another advantage of using D-amino acids for peptides is to reduce the toxicity of PEGylated peptides. PEGylated peptides can generate anti-PEG antibodies upon repeat dosing leading to toxicity, which has been linked with accelerated blood clearance mediated by an anti-PEG antibody response, resulting in rapid clearance of PEGylated carriers, complement activation, and anaphylactic reaction (Sylvestre et al. 2021). However, by substituting the L-amino acids with their D-amino acid enantiomers in PEGylated peptides, anti-PEG

antibody generation and toxicity were significantly reduced (Sylvestre et al. 2021). Unlike the L-amino acid PEGylated peptides, the D-amino acid PEGylated peptides exhibited good tolerability and did not elicit generation of anti-PEG antibodies.

4.4 Other Backbone Modification Strategies

There are several other peptide backbone modification strategies that have been successfully employed to improve the proteolytic stability of peptide drugs. In the thioamide substitution strategy, the oxygen atom in the peptide backbone is replaced by sulfur (Fig. 4.1). It has been shown that thioamide substitution near the scissile bond of GLP-1 and GIP (glucose-dependent insulintropic polypeptide) significantly enhances proteolytic stability without compromising their cellular activity (Chen et al. 2017). Peptides containing a thioamide were shown to be up to 750-fold more stable than the corresponding oxo peptides against cleavage by dipeptidyl peptidase 4 (DPP-4) (Chen et al. 2017). Thioamide substitution has the potential as a simple and effective approach to increase proteolytic stability while maintaining target potency. Substitution of one or more amide groups in the peptide backbone with sulfonamides can also increase proteolytic stability towards peptidases compared to their unmodified analogues, while maintaining biological activity (Evans et al. 2020). Another useful strategy to improve peptide stability is the replacement of peptide bonds with reduced amides (i.e., -CONH- to -CH₂NH-). This strategy has been used in the development of various opioid peptides (peptides that bind to opioid receptors) that demonstrate high stability against chemical and enzymatic degradation while maintaining potency and selectivity (Schiller et al. 2000; Gentilucci et al. 2010). Additionally, replacement of the α -CH group of the peptide backbone with a nitrogen yielding an azapeptide can improve stability (Gentilucci et al. 2010). Atazanavir is a highly active azapeptide for the treatment of HIV (von Hentig 2008) and is taken orally once a day due to its moderate half-life (~7 h) (https://www.accessdata.fda.gov/drugsatfda_docs/label/2011/021567s026lbl.pdf). Incorporation of β -amino acids into peptides is another strategy to improve stability by decreasing recognition by peptidases (Evans et al. 2020).

4.5 Protection of N- and C-Termini

Peptide half-life in plasma generally correlates with the N-termini residue. Peptides with N-termini residues of Met, Ser, Ala, Thr, Val, or Gly typically have longer half-lives, whereas Phe, Leu, Asp, Lys, or Arg at N-termini tend to yield shorter half-lives (Werle and Bernkop-Schnuerch 2006; Goldberg and Gomez-Orellana 2003). There are several peptidases in the blood and tissues that break down peptides from the N- or C-termini, including exopeptidases, aminopeptidases, and

carboxypeptidases. Therefore, modifications of either or both N- and C-termini can significantly improve peptide stability and increase peptide resistance to proteolysis. The most common approaches to protect the N- and C-termini are N-terminal acetylation and C-terminal amidation. Many examples have been reported in the literature demonstrating the success of using these strategies. For example, N-acetylated somatostatin analogs are much more stable than the native peptide (Adessi and Soto 2002). C-terminal amidation and N-terminal acetylation significantly improve plasma stability of the immunogenic peptide MART-I₂₇₋₃₅ (Brinckerhoff et al. 1999). Tesamorelin, which has a hexanoyl group attached to the N-terminus tyrosine residue, has a longer half-life (1 h) than the natural growth hormone-releasing hormone (GHRH, 6.8 min) (Ferdinandi et al. 2007). N-terminal pyroglutamylation leads to DPP-4 stable analogues of GLP-1 7-36 while maintaining potent antihyperglycemic activity (Green et al. 2004). The N-acetylated GLP-1 7-34 was reported to be much more stable than the unprotected peptides (John et al. 2008). N-terminal acetylation of a somatostatin analogue (RC-160) with various long chain fatty acids has greater resistance towards trypsin and serum degradation compared to the unmodified peptide and maintains antiproliferative activity in human breast adenocarcinoma cells (Dasgupta et al. 2002). Palmitate-derived analogues of N-terminal pyroglutamyl GIP are not only resistant to DPP-4 degradation but also have improved potency (Irwin et al. 2005). The dual modifications of N-terminal palmitoylation and C-terminal PEGylation lead to a full agonist of comparable potency to native GIP and stable to DPP-4 cleavage (Salhanick et al. 2005). Head-to-tail cyclization by forming an amide bond between the N- and C-termini can help to stabilize peptides by minimizing cleavage from exopeptidases. The head-to-tail cyclic hexapeptide analogues of peptide T maintain significant biological activity and have high resistance to enzymatic degradation in plasma and brain tissue compared to the linear peptide (Marastoni et al. 1994). In addition to improving peptide stability against exopeptidases, N-acetylation and C-amidation, in combination with amino acid substitutions, also improved resistance against endopeptidases for the EFK17 peptide (EFKRIVQRIKDFLRNLV) (Stroemstedt et al. 2009).

4.6 Cyclization

Cyclic peptides are considered a gold mine for discovering therapeutic drugs (Jing and Jin 2020). Over 40 cyclic peptides have been approved for clinical use, and the vast majority of which are derived from natural products (Jing and Jin 2020). Some well-known cyclic peptide drugs are the antibiotics daptomycin and vancomycin; the hormone analogs octreotide, oxytocin, and vasopressin; and the immunosuppressant cyclosporine (Craik et al. 2013; Driggers et al. 2008; Jing and Jin 2020). Cyclization minimizes peptide degradation in the gastrointestinal tract, blood, and tissues by removing cleavable N- and C-termini (Nielsen et al. 2017). Cyclization also shields the labile peptide bonds from enzymes to reduce degradation.

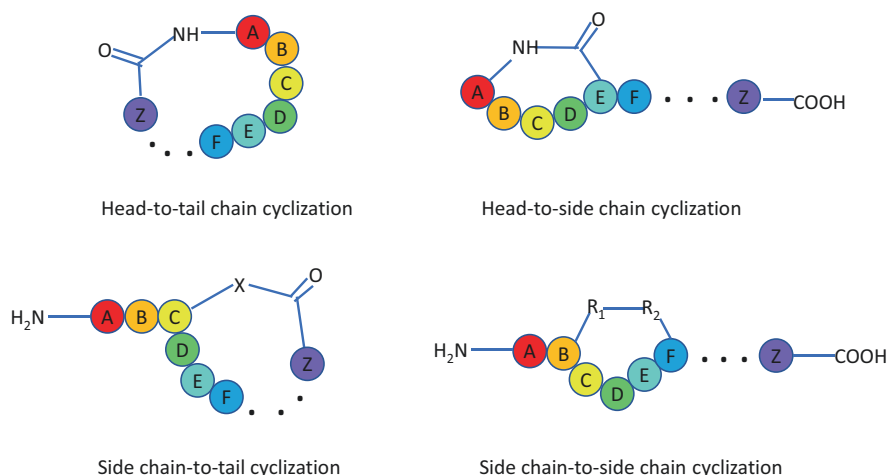


Fig. 4.2 Peptide cyclization patterns. (Modified from reference Jing and Jin 2020)

Cyclization not only improves peptide stability but also enhances membrane permeability of peptides. Depending on the functional groups, peptides can be cyclized in four different manners (Fig 4.2): head-to-tail, head-to-side chain, side chain-to-tail, and side chain-to-side chain (Jing and Jin 2020). Head-to-tail cyclization is the most common approach to generate cyclopeptides. Several cyclization strategies have been developed, including utilization of disulfide bridges, ether bridges, biaryl bridges, stapled peptides, macro-lactones/lactams, and head-to-tail cyclization (Jing and Jin 2020; Gentilucci et al. 2010; Goodwin et al. 2012) (Fig 4.3). As an example, ziconotide, a cyclic peptide for the treatment of pain, contains three disulfide bonds (Aridoss et al. 2021). Another example reported was cyclization of α -conotoxin TxIB that improved its stability and half-life in serum, while maintaining activity (Li et al. 2020). Additionally, compared to a linear peptide, a cyclic epitope peptide derived from herpes simplex virus glycoprotein was completely stable in 50% human serum (Tugyi et al. 2005). Furthermore, cyclic RGD (Arg-Gly-Asp) systems have been shown to be more potent, specific, and resistant to proteolysis than their linear analogues (Schaffner and Dard 2003). Finally, N- to C-terminal cyclization was also used in the development of pasireotide, which led to its long half-life of ~12 h compared to somatostatin (3 min) (Muttenthaler et al. 2021).

4.7 Increase Molecular Weight

Sometimes protection of enzyme-labile sites alone may not be enough to prolong half-life. Increasing MW (molecular weight) through lipidation, fusion to long-lived proteins, or conjugation to polymers not only reduces renal clearance but also

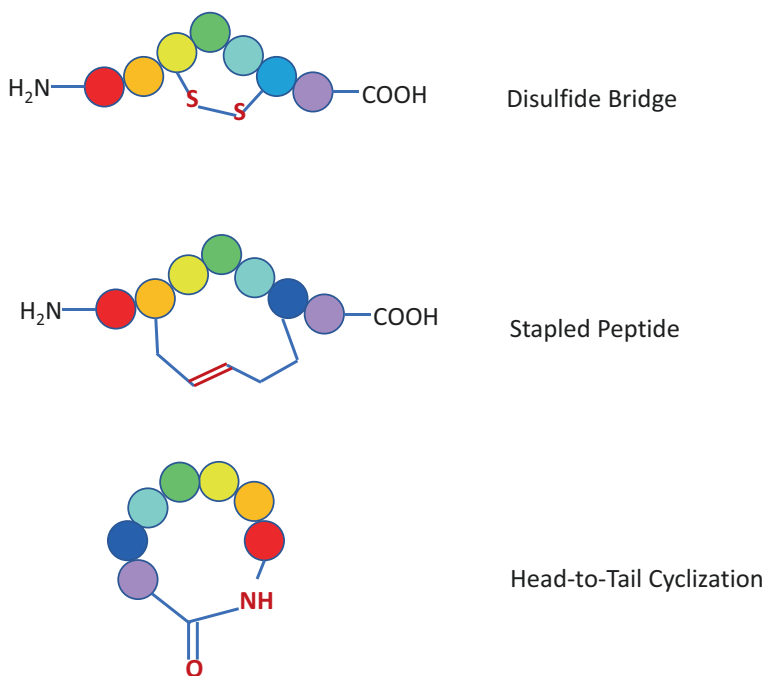


Fig. 4.3 Examples of cyclization strategies to stabilize peptides. (Modified from reference Booka et al. 2021)

protects against proteolytic degradation due to steric hindrance (Pollaro and Heinis 2010; Zaman et al. 2019). As an example, a PEG 2,40K conjugate of INF- α 2b exhibited a 330-fold longer plasma half-life compared to the native protein (Ramon et al. 2005). Another example is the GLP-1 agonist family of peptides for the treatment of T2DM. The native GLP-1 peptide is known to undergo rapid proteolysis with a half-life of a few minutes. Modification of the labile amino acid at the DPP-4 cleavage site led to the discovery of exenatide with improved serum stability (half-life 2–3 h subcutaneously (SC)) (Lau et al. 2015). Covalent attachment of C16 fatty acid to GLP-1 peptide resulted in liraglutide with a long half-life (13–15 h SC) to enable once daily dosing compared to twice a day dosing for exenatide (Lau et al. 2015). Further improvement of stability and renal clearance led to the discovery of semaglutide. Semaglutide contains a C₁₈ fatty diacid covalently bound to GLP-1 peptide. Semaglutide is highly plasma protein bound resulting in decreased renal clearance and protection from metabolic degradation. Semaglutide has a half-life of 1 week and is administered once weekly SC (Kapitza et al. 2015; Agerso et al. 2002; Elbrond et al. 2002). To increase MW, albiglutide (conjugated to albumin, withdrawn from the worldwide market in 2018) and dulaglutide (conjugated to modified IgG4 Fc domain) were developed to prolong half-life of GLP-1 peptide and enable weekly dosing.

4.8 Noncovalent π - π Interactions

A strategy was developed to use noncovalent π - π interactions between aromatic amino acid residues in peptides and synthetic electron-deficient aromatics. These interactions improve peptide stability through the introduction of steric hindrance and the enhancement of peptide α -helicity (Chen et al. 2017). These organic molecule-peptide hybrids display exceptionally high resistance to proteolysis. This strategy is useful because one need not manipulate the sequence of natural peptide extensively or use any unnatural amino acids for modification of the peptide and hence should not alter the potency and safety profiles of a natural peptide developed.

4.9 Enzyme Inhibition

Coadministration of enzyme inhibitors can prolong peptide half-life. It has been shown that co-dosing of a specific DPP-4 inhibitor (ile-thiazolidide) with GLP-1₁₇₋₃₆ improved circulation half-life (Marastoni et al. 1994). Coadministration of NVP DPP728, a DPP-4 inhibitor, prolonged the half-life of native GLP-1 peptide and can be used to treat type 2 diabetes (Ahren et al. 2002). Upon secretion, GLP-1 and GIP are rapidly degraded and inactivated by DPP-4. DPP-4 inhibitors developed have been known to prevent degradation of endogenously released GLP-1 and GIP, such as alogliptin, anagliptin, gemigliptin, linagliptin, saxagliptin, sitagliptin, teneligliptin, and vildagliptin (Chen et al. 2015). Inhibition of GLP-1 and GIP increases plasma levels of active incretins in circulation, prolonging the actions of the incretins, consequently leading to increased insulin levels.

4.10 Flip-Flop Kinetics

Flip-flop kinetics is when the rate of absorption of a compound is significantly slower than its rate of elimination from the body, and it can be explored as a strategy to prolong peptide half-life. Following SC injection, some peptides can depot at the site of injection and slowly release into the systemic circulation leading to long half-life, although clearance remains unchanged (Davenport et al. 2020). For example, degarelix, a synthetic peptide based on GnRH (gonadotropin-releasing hormone), has a half-life of 42–72 days due to flip-flop kinetics (Steinberg 2009; Anderson 2009). Lanreotide, a somatostatin agonist containing unnatural amino acids, has a half-life of 22 days as a result of slowly releasing from the depot of the injection site (Kyriakakis et al. 2014).

4.11 Methods to Assess Peptide Stability

The major sites for peptide degradation *in vivo* are the blood, liver, kidney, and intestine. *In silico* methods have been developed to predict half-life of peptides in mammalian blood (Mathur et al. 2018) and in intestine-like environment (Sharma et al. 2014). Assays can also be designed to evaluate peptide stability in various biomatrices. Quantitative measurement of kinetic parameters, such as half-life and intrinsic clearance, can be accomplished by monitoring parent depletion over time. Identification of degradation products also provides valuable information on the cleavage sites. This information is useful to enable the design of more stable peptides. Depending on the instability issues of compounds in various matrices, specific assays can be selected for screening to guide SAR and improve peptide stability. The stability information in different bioassay matrices can be used to predict clearance and exposure of peptides in human, as well as develop PK/PD relationship and estimate dose.

4.12 Blood, Serum, and Plasma Stability Assays

The blood is a major site of peptide degradation. In the incubation of peptides, fresh blood, serum (obtained as supernatants after complete blood coagulation), and plasma (centrifugation of blood supplemented with anticoagulants, such as EDTA or heparin) have all been used to examine the stability of peptides in systemic circulation (Bottger et al. 2017). However, the stability of a given peptide can change depending on the chosen reagents. For instance, enzyme activity can change during the coagulation process in serum preparation. In addition, anticoagulants added for plasma preparation can inhibit enzyme activity. For example, EDTA (ethylenediaminetetraacetic acid) inhibits metallo- and Ca^{2+} -dependent proteases, and heparin inhibits thrombin and factor Xa (Bottger et al. 2017). Red blood cells contain additional enzymes that are not present in plasma or serum (Cossum 1988). Thus, it is also important to evaluate peptide stability in blood. Barr et al. reported using serum and brain tissue homogenate to study the stability of peptides targeting soluble β -amyloid oligomers (Barr et al. 2016). Bottger et al. compared the stability of peptides from three different families in mouse serum, plasma, and blood (Bottger et al. 2017). They found that generally peptides had faster degradation rates in serum than in plasma and surprisingly were most stable in fresh blood (Bottger et al. 2017). Therefore, serum or plasma stability results may be misleading and need to be verified *in vivo*.

4.13 Hepatic Stability Assays

The liver is a primary site of degradation for some peptides, such as insulin (Duckworth et al. 1998). Liver microsomes are commonly used to evaluate the stability of membrane-bound enzymes of the liver, and liver cytosols are often used for soluble enzymes. The liver S9 fraction contains both membrane-bound and soluble

enzymes and can be used to evaluate peptide hepatic stability. Cryopreserved hepatocytes are also commonly applied as they contain a complete complement of enzymes and cofactors. However, low hepatocyte cell membrane permeability of peptides may limit their metabolic rate leading to artificially longer half-lives and underestimated *in vivo* clearance (Keefer et al. 2020). Therefore, liver S9 is usually a preferred system over hepatocytes to study peptide stability in order to eliminate the potential impact of a membrane barrier. A comparative study was conducted to evaluate peptide stability using human liver S9 and hepatocytes (Jyrkas and Tolonen 2021). The results showed that both systems performed similarly for peptide hydrolysis, but liver S9 was more effective for CYP metabolism and metabolite generation of some peptides with low cell membrane permeability (Jyrkas and Tolonen 2021). It is important to note that reagents need to be free of PMSF (phenylmethylsulfonyl fluoride) and EDTA as they can inhibit peptidase activity and lead to falsely more stable results. In a study of small peptide hormones as potential doping agents, human liver microsomes, liver and kidney S9 fractions, serum, and α -chymotrypsin were used to identify metabolites that might be used for the detection of substance abuse (Esposito et al. 2015). Zvereva et al. found that kidney microsomes and liver S9 were the most efficient *in vitro* systems for generating diverse metabolite profiles of peptides (Zvereva et al. 2016). Furthermore, Yang et al. reported the use of plasma, liver and brain homogenates with and without paraoxon (esterase inhibitor) to examine peptide stability (Yang et al. 2002).

4.14 Kidney and Tissue Stability Assays

The kidney is a major site of peptide degradation that is not very common for small molecules. For example, significant degradation of exenatide and insulin occurs in the kidneys (Liao et al. 2015; Duckworth et al. 1998). Kidney brush border membrane vesicles (BBMV) (Biber et al. 2007), microsomes, or homogenates can be used to assess peptide degradation in the kidneys (Liao et al. 2015; Duckworth et al. 1998; Copley et al. 2006; Zvereva et al. 2018). In a doping control study to determine abuse of GnRH (gonadotropin-releasing hormone) and its synthetic analogues, human kidney microsomes were used to identify the metabolites that may help detect substance abuse (Zvereva et al. 2018). Bendre et al. reported the use of kidney BBMV to evaluate peptide stability in order to reduce kidney metabolism of peptides and peptidomimetics (Bendre et al. 2020). Liao et al. used rat kidney and liver homogenate to evaluate the stability, metabolite profile, and cleavage site of exendin-4 (Liao et al. 2015).

4.15 Gastric Intestinal Stability Assays

To achieve oral bioavailability, peptides need to have sufficient stability in the gastrointestinal (GI) tract to be absorbed. Simulated gastric fluid (SGF, pH 1.2) and simulated intestinal fluid (SIF, pH 6.8) or human gastrointestinal fluids are commonly used to study the stability of peptides in GI fluids to evaluate both pH and

enzymatic effects (Di and Kerns 2016; Di 2015). Stability in intestine BBMVs can be used to evaluate stability of peptides in the intestine mucus (Dishon et al. 2018). Intestine microsomes, cytosols, S9, and enterocytes (Ho et al. 2017) can be applied to examine the role of membrane-bound and soluble enzymes in peptide cleavage. Elfgen et al. discussed using SGF, SIF, plasma, and liver microsomes to evaluate stability of the orally available D-peptide RD2 developed for direct elimination of β -amyloid (Elfgen et al. 2019). Dushin et al. reported the use of rat intestinal BBMVs and plasma to compare stability between cyclic and linear peptide inhibitors of the innate immune TLR/IL1R signaling protein MyD88 (Dishon et al. 2018).

4.16 Conclusions

Peptide therapeutics continue to be an integral part of pharmaceutical R&D. The development of peptide drugs can be challenging but also highly rewarding. The peptide drug market is growing twice as fast as the rest of the drug market. Several peptide drugs have achieved blockbuster status. However, low in vivo stability and lack of oral bioavailability are still major ADME (absorption, distribution, metabolism, and excretion) issues for peptides. Other challenges of peptide therapeutics include limitation of disease targets to mostly extracellular space and relatively expensive synthesis. Proteolytic degradation in blood and tissues is one of the key elimination pathways for peptides, along with renal clearance. Structural modification strategies coupled with novel formulation approaches enable the transformation of peptides into effective therapeutic agents. Backbone modifications, protection of N- and C-termini, cyclization, and increasing molecular weight are some of the design principles to improve peptide proteolytic stability. In vitro and in vivo assays have been established to evaluate peptide stability and to develop SAR (structure-activity relationship) to guide peptide design with improved stability. The biomatrices for peptide stability studies include physiological fluids (e.g., serum, plasma, blood, SGF, SIF) and tissue preparations of the liver, kidney, intestine, brain, and other tissues (e.g., BBMVs, microsomes, cytosols, S9, hepatocytes, enterocytes, tissue homogenates). Successful development of peptide drugs requires the integration of multidisciplinary science and technologies. Further innovation in the field will accelerate peptide drug development to bring novel medicines to patients.

References

- Adessi C, Soto C. Converting a peptide into a drug: strategies to improve stability and bioavailability. *Curr Med Chem*. 2002;9:963–78.
- Agerso H, Jensen LB, Elbrond B, Rolan P, Zdravkovic M. The pharmacokinetics, pharmacodynamics, safety and tolerability of NN2211, a new long-acting GLP-1 derivative, in healthy men. *Diabetologia*. 2002;45:195–202.
- Ahren B, Simonsson E, Larsson H, Landin-Olsson M, Torgeirsson H, Jansson P-A, Sandqvist M, Bavenholm P, Efendic S, Eriksson JW, Dickinson S, Holmes D. Inhibition of dipeptidyl pep-

- tidase IV improves metabolic control over a 4-week study period in type 2 diabetes. *Diabetes Care*. 2002;25:869–75.
- Anderson J. Degarelix: a novel gonadotropin-releasing hormone blocker for the treatment of prostate cancer. *Future Oncol*. 2009;5:433–43.
- Aridoss G, Kim D-M, Kim JI, Kang JE. Ziconotide (ω -conotoxin MVIIA)-Efficient solid-phase synthesis of a linear precursor peptide and its strategic native folding. *Pept Sci*. 2021;113(5):e24223.
- Barr RK, Verdile G, Wijaya LK, Morici M, Taddei K, Gupta VB, Pedrini S, Jin L, Nicolazzo JA, Knock E, Fraser PE, Martins RN. Validation and characterization of a novel peptide that binds monomeric and aggregated β -Amyloid and inhibits the formation of neurotoxic oligomers. *J Biol Chem*. 2016;291:547–59.
- Ben-Yedidia T, Beignon A-S, Partidos CD, Muller S, Arnon R. A retro-inverso peptide analogue of influenza virus hemagglutinin B-cell epitope 91-108 induces a strong mucosal and systemic immune response and confers protection in mice after intranasal immunization. *Mol Immunol*. 2002;39:323–31.
- Bendre S, Zhang Z, Kuo H-T, Rousseau J, Zhang C, Merkens H, Roxin A, Benard F, Lin K-S. Evaluation of Met-Val-Lys as a renal brush border enzyme-cleavable linker to reduce kidney uptake of ^{68}Ga -labeled DOTA-conjugated peptides and peptidomimetics. *Molecules*. 2020;25:3854.
- Biber J, Stieger B, Stange G, Murer H. Isolation of renal proximal tubular brush-border membranes. *Nat Protoc*. 2007;2:1356–9.
- Booka AI, Lechanteur A, Fillet M, Piel G. Therapeutic peptides for chemotherapy: trends and challenges for advanced delivery systems. *Eur J Pharm Biopharm*. 2021; Published ahead of print
- Boschi-Muller S, Gand A, Branlant G. The methionine sulfoxide reductases: catalysis and substrate specificities. *Arch Biochem Biophys*. 2008;474:266–73.
- Botterger R, Hoffmann R, Knappe D. Differential stability of therapeutic peptides with different proteolytic cleavage sites in blood, plasma and serum. *PLoS One*. 2017;12:e0178943/1-e43/15.
- Brinckerhoff LH, Kalashnikov VV, Thompson LW, Yamshchikov GV, Pierce RA, Galavotti HS, Engelhard VH, Slingluff CL. Terminal modifications inhibit proteolytic degradation of an immunogenic MART-127-35 peptide: implications for peptide vaccines. *Int J Cancer*. 1999;83:326–34.
- Buckley ST, Vegge A, Pyke C, Ahnfelt-Ronne J, Madsen KG, Scheele SG, Alanentalo T, Kirk RK, Pedersen BL, Skyggebjerg RB, Benie AJ, Strauss HM, Wahlund P-O, Bjerregaard S, Knudsen LB, Baekdal TA, Maarbjerg SJ, Sondergaard FL, Borregaard J, Hartoft-Nielsen M-L, Farkas E, Fekete C, Fekete C. Transcellular stomach absorption of a derivatized glucagon-like peptide-1 receptor agonist. *Sci Transl Med*. 2018;10
- Chen X-W, He Z-X, Zhou Z-W, Yang T, Zhang X, Yang Y-X, Duan W, Zhou S-F. Clinical pharmacology of dipeptidyl peptidase 4 inhibitors indicated for the treatment of type 2 diabetes mellitus. *Clin Exp Pharmacol Physiol*. 2015;42:999–1024.
- Chen X, Mietlicki-Baase EG, Barrett TM, McGrath LE, Koch-Laskowski K, Ferrie JJ, Hayes MR, James Petersson E. Thioamide substitution selectively modulates proteolysis and receptor activity of therapeutic peptide hormones. *J Am Chem Soc*. 2017;139:16688–95.
- Chen Y, Li T, Li J, Cheng S, Wang J, Verma C, Zhao Y, Chuanliu W. Stabilization of peptides against proteolysis through disulfide-bridged conjugation with synthetic aromatics. *Org Biomol Chem*. 2017;15:1921–9.
- Chiu Y-Y, Higaki K, Neudeck BL, Barnett JL, Welage LS, Amidon GL. Human jejunal permeability of cyclosporin A: influence of surfactants on P-glycoprotein efflux in Caco-2 cells. *Pharm Res*. 2003;20:749–56.
- Copley K, McCowen K, Hiles R, Nielsen LL, Young A, Parkes DG. Investigation of exenatide elimination and its in vivo and in vitro degradation. *Curr Drug Metab*. 2006;7:367–74.
- Cossum PA. Role of the red blood cell in drug metabolism. *Biopharm Drug Dispos*. 1988;9:321–36.
- Craik DJ, Fairlie DP, Liras S, Price D. The future of peptide-based drugs. *Chem Biol Drug Des*. 2013;81:136–47.
- Darlak K, Benovitz DE, Spatola AF, Grzonka Z. Dermorphin analogs: resistance to in vitro enzymic degradation is not always increased by additional D-amino acid substitutions. *Biochem Biophys Res Commun*. 1988;156:125–30.

- Dasgupta P, Singh A, Mukherjee R. N-terminal acylation of somatostatin analog with long chain fatty acids enhances its stability and anti-proliferative activity in human breast adenocarcinoma cells. *Biol Pharm Bull.* 2002;25:29–36.
- Davenport AP, Scully CCG, de Graaf C, Brown AJH, Maguire JJ. Advances in therapeutic peptides targeting G protein-coupled receptors. *Nat Rev Drug Discov.* 2020;19:389–413.
- Di L. Strategic approaches to optimizing peptide ADME properties. *AAPS J.* 2015;17:134–43.
- Di L, Kerns EH. *Drug-like properties: concepts, structure design, and methods.* Elsevier: London; 2016.
- Dishon S, Nussbaum G, Schumacher A, Fanous J, Hoffman A, Talhami A, Gilon C, Kassis I, Karussis D. Development of a novel backbone cyclic peptide inhibitor of the innate immune TLR/IL1R signaling protein MyD88. *Sci Rep.* 2018;8:9476.
- Driggers EM, Hale SP, Lee J, Terrett NK. The exploration of macrocycles for drug discovery – an underexploited structural class. *Nat Rev Drug Discov.* 2008;7:608–24.
- Duckworth WC, Gennett RG, Hamel FG. Insulin degradation: progress and potential. *Endocr Rev.* 1998;19:608–24.
- Elbrond B, Jakobsen G, Larsen S, Agero H, Jensen LB, Rolan P, Sturis J, Hatorp V, Zdravkovic M. Pharmacokinetics, pharmacodynamics, safety, and tolerability of a single-dose of NN2211, a long-acting glucagon-like peptide 1 derivative, in healthy male subjects. *Diabetes Care.* 2002;25:1398–404.
- Elfgen A, Bochinsky K, Tusche M, Gering I, Hartmann R, Kutzsche J, Willbold D, Hupert M, de Gonzalez SRME, Huesgen PF, Santiago-Schubel B, Kutzsche J, Sacchi S, Pollegioni L, Sacchi S, Pollegioni L, Willbold D. Metabolic resistance of the D-peptide RD2 developed for direct elimination of amyloid- β oligomers. *Sci Rep.* 2019;9:5715.
- Esposito S, Deventer K, Geldof L, Van Eenoo P. In vitro models for metabolic studies of small peptide hormones in sport drug testing. *J Pept Sci.* 2015;21:1–9.
- Evans BJ, King AT, Katsifis A, Matesic L, Jamie JF. Methods to enhance the metabolic stability of peptide-based PET radiopharmaceuticals. *Molecules.* 2020;25:2314.
- Ferdinand ES, Brazeau P, High K, Procter B, Fennell S, Dubreuil P. Non-clinical pharmacology and safety evaluation of TH9507, a human growth hormone-releasing factor analogue. *Basic Clin Pharmacol Toxicol.* 2007;100:49–58.
- Gentilucci L, De Marco R, Cerisoli L. Chemical modifications designed to improve peptide stability: incorporation of non-natural amino acids, pseudo-peptide bonds, and cyclization. *Curr Pharm Des.* 2010;16:3185–203.
- Goldberg M, Gomez-Orellana I. Challenges for the oral delivery of macromolecules. *Nat Rev Drug Discov.* 2003;2:289–95.
- Goodwin D, Simerska P, Toth I. Peptides as therapeutics with enhanced bioactivity. *Curr Med Chem.* 2012;19:4451–61.
- Green BD, Mooney MH, Gault VA, Irwin N, Bailey CJ, Harriott P, Greer B, O'Harte FPM, Flatt PR. N-terminal His7-modification of glucagon-like peptide-1(7-36) amide generates dipeptidyl peptidase IV-stable analogues with potent antihyperglycaemic activity. *J Endocrinol.* 2004;180:379–88.
- Harris AG. Somatostatin and somatostatin analogues: pharmacokinetics and pharmacodynamic effects. *Gut.* 1994;35:S1–4.
- Ho M-CD, Ring N, Amaral K, Doshi U, Albert PL. Human enterocytes as an in vitro model for the evaluation of intestinal drug metabolism: characterization of drug-metabolizing enzyme activities of cryopreserved human enterocytes from twenty-four donors. *Drug Metab. Dispos.* 2017;45:686–91.
- https://www.accessdata.fda.gov/drugsatfda_docs/label/2011/021567s026lbl.pdf
- <https://www.fda.gov/media/135421/download>
- [https://www.researchandmarkets.com/reports/5265155/peptide-therapeutics-market-growth-trends?utm_source=BW&utm_medium=PressRelease&utm_code=nbjvs&utm_campaign=1507648_PLUS_SPI_-_PLUS_SPI_Global_PLUS_SPI_Peptide_PLUS_SPI_Therapeutics_PLUS_SPI_Market_PLUS_SPI_\(2021_PLUS_SPI_to_PLUS_SPI_2026\)_PLUS_SPI_-_PLUS_SPI_Growth%2c_PLUS_SPI_Trends%2c_PLUS_SPI_COVID-19_PLUS_SPI_Impact%2c_PLUS_SPI_and_PLUS_SPI_Forecasts&utm_exec=jamu273prd](https://www.researchandmarkets.com/reports/5265155/peptide-therapeutics-market-growth-trends?utm_source=BW&utm_medium=PressRelease&utm_code=nbjvs&utm_campaign=1507648_PLUS_SPI_-_PLUS_SPI_Global_PLUS_SPI_Peptide_PLUS_SPI_Therapeutics_PLUS_SPI_Market_PLUS_SPI_(2021_PLUS_SPI_to_PLUS_SPI_2026)_PLUS_SPI_-_PLUS_SPI_Growth%2c_PLUS_SPI_Trends%2c_PLUS_SPI_COVID-19_PLUS_SPI_Impact%2c_PLUS_SPI_and_PLUS_SPI_Forecasts&utm_exec=jamu273prd)

- Irwin N, Green BD, Gault VA, Greer B, Harriott P, Bailey CJ, Flatt PR, O'Harte FPM. Degradation, insulin secretion, and antihyperglycemic actions of two palmitate-derivitized N-terminal pyroglutamyl analogues of glucose-dependent insulintropic polypeptide. *J Med Chem.* 2005;48:1244–50.
- Jing X, Jin K. A gold mine for drug discovery: strategies to develop cyclic peptides into therapies. *Med Res Rev.* 2020;40:753–810.
- John H, Maronde E, Forssmann W-G, Meyer M, Adermann K. N-terminal acetylation protects glucagon-like peptide GLP-1-(7-34)-amide from DPP-IV-mediated degradation retaining cAMP- and insulin releasing capacity. *Eur J Med Res.* 2008;13:73–8.
- Jyrkas J, Tolonen A. Hepatic in vitro metabolism of peptides; Comparison of human liver S9, hepatocytes and Upcyte hepatocytes with cyclosporine A, leuprorelin, desmopressin and cetorelix as model compounds. *J Pharm Biomed Anal.* 2021;196:113921.
- Kapitzka C, Nosek L, Jensen L, Hartvig H, Jensen CB, Flint A. Semaglutide, a once-weekly human GLP-1 analog, does not reduce the bioavailability of the combined oral contraceptive, ethinyl-estradiol/levonorgestrel. *J Clin Pharmacol.* 2015;55:497–504.
- Keefer C, Chang G, Carlo A, Novak JJ, Banker M, Carey J, Cianfrogna J, Eng H, Jagla C, Johnson N, Jones R, Jordan S, Lazzaro S, JianHua Liu R, Obach S, Riccardi K, Tess D, Umland J, Racich J, Varma M, Visswanathan R, Di L. Mechanistic insights on clearance and inhibition discordance between liver microsomes and hepatocytes when clearance in liver microsomes is higher than in hepatocytes. *Eur J Pharm Sci.* 2020;155:105541.
- Knudsen LB, Lau J. The discovery and development of liraglutide and semaglutide. *Front Endocrinol (Lausanne).* 2019;10:155.
- Kyriakakis N, Chau V, Lynch J, Orme SM, Murray RD. Lanreotide autogel in acromegaly – a decade on. *Expert Opin. Pharmacother.* 2014;15:2681–92.
- Lau J, Bloch P, Schaffer L, Pettersson I, Spetzler J, Kofoed J, Madsen K, Knudsen LB, McGuire J, Steensgaard DB, Strauss HM, Gram DX, Knudsen SM, Nielsen FS, Thygesen P, Reedtz-Runge S, Kruse T. Discovery of the once-weekly Glucagon-Like Peptide-1 (GLP-1) Analogue Semaglutide. *J Med Chem.* 2015;58:7370–80.
- Lee AC-L, Harris JL, Khanna KK, Hong J-H. A comprehensive review on current advances in peptide drug development and design. *Int J Mol Sci.* 2019;20:2383/1-83/21.
- Li H, Kem DC, Zhang L, Huang B, Liles C, Benbrook A, Gali H, Veitla V, Scherlag BJ, Cunningham MW, Xichun Y. Novel retro-inverso peptide inhibitor reverses angiotensin receptor autoantibody-induced hypertension in the rabbit. *Hypertension.* 2015;65:793–9.
- Li X, Wang S, Zhu X, Zhangsun D, Yong W, Luo S. Effects of cyclization on activity and stability of α -conotoxin TxIB. *Mar Drugs.* 2020;18:180.
- Liao S, Liang Y, Zhang Z, Li J, Wang J, Wang X, Dou G, Zhang Z, Liu K. In vitro metabolic stability of exendin-4: pharmacokinetics and identification of cleavage products. *PLoS One.* 2015;10:e0116805/1-e05/18.
- Maher S, Mrsny RJ, Brayden DJ. Intestinal permeation enhancers for oral peptide delivery. *Adv Drug Deliv Rev.* 2016;106:277–319.
- Makurvet FD. Biologics vs. small molecules: drug costs and patient access. *Med Drug Discovery.* 2021;9:100075.
- Marastoni M, Salvadori S, Scaranari V, Spisani S, Reali E, Traniello S, Tomatis A. Synthesis and activity of new linear and cyclic peptide T derivatives. *Arzneim-Forsch.* 1994;44:1073–6.
- Mathur D, Singh S, Mehta A, Agrawal P, Raghava GPS. In silico approaches for predicting the half-life of natural and modified peptides in blood. *PLoS One.* 2018;13:e0196829/1-e29/10.
- Mocsai A, Kovacs L, Gergely P. What is the future of targeted therapy in rheumatology: biologics or small molecules? *BMC Med.* 2014;12:43/1-43/9, 9.
- Muttenthaler M, King GF, Adams DJ, Alewood PF. Trends in peptide drug discovery. *Nat Rev Drug Discov.* 2021;20:309–25.
- Nielsen DS, Shepherd NE, Xu W, Lucke AJ, Stoermer MJ, Fairlie DP. Orally absorbed cyclic peptides. *Chem Rev.* 2017;117:8094–128.
- Pollaro L, Heinis C. Strategies to prolong the plasma residence time of peptide drugs. *Med Chem Commun.* 2010;1:319–24.

- Rafferty B, Coy DH, Poole S. Pharmacokinetic evaluation of superactive analogs of growth hormone-releasing factor (1-29)-amide. *Peptides* (Fayetteville, N Y). 1988;9:207–9.
- Ramon J, Saez V, Baez R, Aldana R, Hardy E. PEGylated interferon- α 2b: a branched 40K polyethylene glycol derivative. *Pharm Res*. 2005;22:1374–86.
- Salhanick AI, Clairmont KB, Buckholz TM, Pellegrino CM, Ha S, Lumb KJ. Contribution of site-specific PEGylation to the dipeptidyl peptidase IV stability of glucose-dependent insulinotropic polypeptide. *Bioorg Med Chem Lett*. 2005;15:4114–7.
- Schaffner P, Dard MM. Structure and function of RGD peptides involved in bone biology. *Cell Mol Life Sci*. 2003;60:119–32.
- Schiller PW, Weltrowska G, Berezowska I, Nguyen TMD, Wilkes BC, Lemieux C, Chung NN. The TIPP opioid peptide family: development of δ antagonists, δ agonists, and mixed μ agonist/ δ antagonists. *Biopolymers*. 2000;51:411–25.
- Sharma A, Singla D, Rashid M, Raghava GPS. Designing of peptides with desired half-life in intestine-like environment. *BMC Bioinf*. 2014;15:282/1–8.
- Sharma R, McDonald TS, Eng H, Limberakis C, Stevens BD, Patel S, Kalgutkar AS. In vitro metabolism of the glucagon-like peptide-1 (GLP-1)-derived metabolites GLP-1(9-36)amide and GLP-1(28-36)amide in mouse and human hepatocytes. *Drug Metab Dispos*. 2013;41:2148–57.
- Steinberg M. Degarelix: a gonadotropin-releasing hormone antagonist for the management of prostate cancer. *Clin Ther*. 2009;31:2312–31.
- Stroemstedt AA, Pasupuleti M, Schmidtchen A, Malmsten M. Evaluation of strategies for improving proteolytic resistance of antimicrobial peptides by using variants of EFK17, an internal segment of LL-37. *Antimicrob Agents Chemother*. 2009;53:593–602.
- Sylvestre M, Lv S, Yang LF, Luera N, Peeler DJ, Chen B-M, Roffler SR, Pun SH. Replacement of L-amino acid peptides with D-amino acid peptides mitigates anti-PEG antibody generation against polymer-peptide conjugates in mice. *J Control Release*. 2021;331:142–53.
- Tugyi R, Mezo G, Fellinger E, Andreu D, Hudecz F. The effect of cyclization on the enzymatic degradation of herpes simplex virus glycoprotein D derived epitope peptide. *J Pept Sci*. 2005;11:642–9.
- Veine DM, Yao H, Stafford DR, Fay KS, Livant DL. A d-amino acid containing peptide as a potent, noncovalent inhibitor of α 5 β 1 integrin in human prostate cancer invasion and lung colonization. *Clin Exp Metastasis*. 2014;31:379–93.
- von Hentig N. Atazanavir/ritonavir: a review of its use in HIV therapy. *Drugs Today*. 2008;44:103–32.
- Wakankar AA, Borchardt RT. Formulation considerations for proteins susceptible to asparagine deamidation and aspartate isomerization. *J Pharm Sci*. 2006;95:2321–36.
- Werle M, Bernkop-Schnuerch A. Strategies to improve plasma half life time of peptide and protein drugs. *Amino Acids*. 2006;30:351–67.
- Wu C-Y, Benet LZ, Hebert MF, Gupta SK, Rowland M, Gomez DY, Wacher VJ. Differentiation of absorption and first-pass gut and hepatic metabolism in humans: studies with cyclosporine. *Clin Pharmacol Ther* (St Louis). 1995;58:492–7.
- Wu H, Huang J. Optimization of protein and peptide drugs based on the mechanisms of kidney clearance. *Protein Pept Lett*. 2018;25:514–21.
- Yang JZ, Chen W, Borchardt RT. In vitro stability and in vivo pharmacokinetic studies of a model opioid peptide, H-Tyr-D-Ala-Gly-Phe-D-Leu-OH (DADLE), and its cyclic prodrugs. *J Pharmacol Exp Ther*. 2002;303:840–8.
- Yao J-F, Yang H, Zhao Y-Z, Xue M. Metabolism of peptide drugs and strategies to improve their metabolic stability. *Curr Drug Metab*. 2018;19:892–901.
- Zaman R, Islam RA, Ibat N, Othman I, Zaini A, Lee CY, Chowdhury EH. Current strategies in extending half-lives of therapeutic proteins. *J Control Release*. 2019;301:176–89.
- Zvereva I, Dudko G, Dikunets M. 'Determination of GnRH and its synthetic analogues' abuse in doping control: small bioactive peptide UPLC-MS/MS method extension by addition of in vitro and in vivo metabolism data; evaluation of LH and steroid profile parameter fluctuations as suitable biomarkers. *Drug Test Anal*. 2018;10:711–22.
- Zvereva I, Semenistaya E, Krotov G, Rodchenkov G. Comparison of various in vitro model systems of the metabolism of synthetic doping peptides: proteolytic enzymes, human blood serum, liver and kidney microsomes and liver S9 fraction. *J Proteome*. 2016;149:85–97.

Chapter 5

Therapeutic Peptide Delivery: Fundamentals, Formulations, and Recent Advances



Deepal Vora, Amruta A. Dandekar, and Ajay K. Banga

Contents

5.1 Introduction.....	184
5.2 Peptide Therapeutics Administered via Parenteral Route.....	184
5.3 Peptide Therapeutics Administered via Oral Route.....	187
5.4 Peptide Therapeutics Administered via Pulmonary Route.....	190
5.5 Peptide Therapeutics Administered via Transdermal Route.....	193
5.6 Peptide Therapeutics Administered via Other Routes.....	196
5.7 Summary.....	197
References.....	198

Abstract Recent advances in the field of peptide therapeutics have led to the design and synthesis of many promising peptides. However, research and development to provide safe, stable, efficacious, and patient compliant formulation plays a vital role in bringing peptide therapeutics to market. The conventional parenteral route has made scientific advances to overcome multiple barriers, leading to the approval of many peptide-based products via parenteral route of administration in recent years. In parallel, oral, pulmonary, transdermal, and other delivery routes have been extensively investigated to deliver peptides with improved patient compliance. This includes the use of novel strategies for developing delivery systems that can offer various advantages over conventional dosage forms. This chapter focuses on fundamentals, formulations, and recent advances for successful peptide delivery.

Keywords Peptides · Pharmaceutical market · Formulation development · Delivery systems · Parenteral · Oral · Pulmonary · Transdermal · Controlled release parenteral · Carrier systems

Authors “Deepal Vora” and “Amruta A. Dandekar” have equally contributed to this chapter.

D. Vora · A. A. Dandekar · A. K. Banga (✉)
Center for Drug Delivery Research, Department of Pharmaceutical Sciences,
College of Pharmacy, Mercer University, Atlanta, GA, USA
e-mail: banga_ak@mercer.edu

5.1 Introduction

Peptides and polypeptides are small-sized compounds (fewer than 50 amino acids) and a separate category of drugs positioned between small organic molecules and large proteins (Lau and Dunn 2018). Peptides offer several advantages such as better efficiency, specificity, and selectivity as compared to small molecules and are more economical and less immunogenic as compared to large proteins (D'Hondt et al. 2014). Peptides of natural and synthetic origin have been involved in various biological roles such as hormones, enzyme substrate, antibiotics, regulators, and inhibitors (Tesauro et al. 2019). Recent years have witnessed the authorization of over 25 peptides and oligonucleotide drugs (Al Musaimi et al. 2021). In 2020, the US Food and Drug Administration (FDA) had approved 53 new drug entities, 6 of which fall in the peptides and oligonucleotides category. This number is almost 10% of the total drug authorizations and reflects the consolidation of peptides in the market. Although peptides cover a wide spectrum of therapeutic areas, three major areas for peptide development include metabolic disease, cardiovascular disease, and oncology (Lau and Dunn 2018). These numbers reflect the potential interest in peptides (D'Hondt et al. 2014). The rise in scientific publications in the last decade related to peptide research indicates increased interest in this area of research which aligns with the current trend of marketed peptide therapeutics. However, the physical and chemical instability, enzymatic breakdown, and short half-life are the major barriers to the use of peptides (Sachdeva et al. 2016). Various routes and newer delivery systems have thus been explored for therapeutic peptide delivery. While the most preferred and traditional parenteral route of administration overcomes some of these challenges, other routes of administration namely oral, pulmonary, transdermal are widely researched and investigated for safe and efficient delivery of peptide therapeutics (Fig. 5.1). This chapter focuses on different routes of administration for peptide therapeutics, formulation development, examples of marketed formulations, and recent advances in the field of peptide delivery such as new drug delivery systems and technologies investigated.

5.2 Peptide Therapeutics Administered via Parenteral Route

In the past few decades, tremendous advances have been made in bringing peptide therapeutics to the market. Despite efforts put into other routes of administration, the parenteral route remains the primary and most promising route for the administration of peptides. However, their delivery can be challenging due to factors such as susceptibility to denaturation, degradation by enzymes, and short half-life, ultimately leading to poor bioavailability (Banga 2005; Agarwal and Rupenthal 2013). Parenteral route of drug administration refers to injection directly into the body, which bypasses the skin and mucous membranes, and common routes of parenteral administration are intramuscular, subcutaneous, and intravenous.

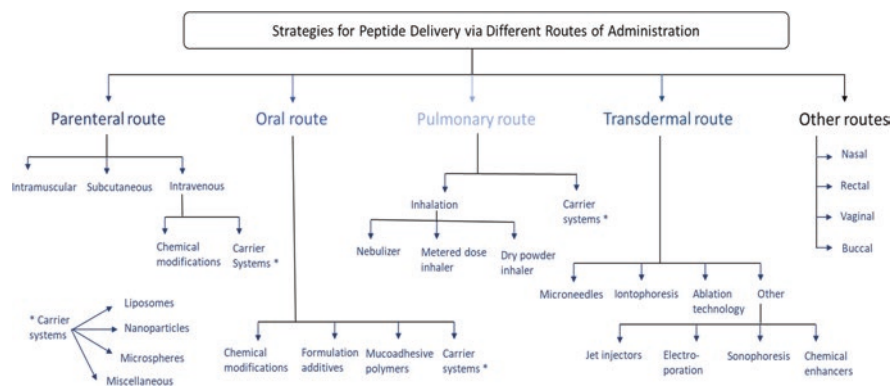


Fig. 5.1 Summary of strategies for peptide delivery via different routes of administration

The intramuscular route refers to the administration of medication deep into the muscles. Most of the vaccines, like flu shots, are administered via the intramuscular route. Interferon alfa-2b to treat hairy cell leukemia, malignant melanoma is commonly administered via the intramuscular route. Plenaxis (abarelix), a product by Praecis Pharmaceuticals, which is an injectable gonadotropin-releasing hormone antagonist, approved by the FDA in 2003 is administered via intramuscular route (https://www.accessdata.fda.gov/drugsatfda_docs/label/2011/021567s026lbl.pdf).

The subcutaneous route refers to the administration of drug using a short needle into the tissue layers between the skin and the muscle. Absorption from the subcutaneous route is usually slower than intravenous route. This route also permits self-administration, thus freeing up the healthcare practitioners' time for other issues beyond administration (Banga 2005; Ibeanu et al. 2020). Insulin, the first polypeptide to be administered therapeutically, is given via the subcutaneous route. Adipose tissue peptide, leptin to treat obesity, octreotide to treat acromegaly, and peptide neurotrophic factors such as nerve growth factor (NGF) used in the therapy of neurodegenerative disorders are some examples of peptides or polypeptides commonly administered via the subcutaneous routes. Imcivree™ (Setmelanotide), a product by Molina Healthcare for use as anti-obesity medication and delivered via subcutaneous route, was approved by the FDA in 2020.

The intravenous route facilitates the administration of drug directly into the systemic circulation. A number of approved and under clinical trial peptides via intravenous routes have increased exponentially in recent years (Asfour 2021). For example, Parsabiv (Etelcalcetide) is a product currently owned by Amgen and Ono Pharmaceuticals and was approved by FDA in 2017 for chronic idiopathic hyperparathyroidism and is administered via the intravenous route. Another recently FDA-approved product via intravenous route is Polivy® (polatuzumab vedotin-piiq), a product by Genentech Inc., indicated to treat relapsed or refractory diffuse large B-cell lymphoma. Padcev® (enfortumab vedotin-ejfv) was also approved by FDA in 2019 for adult patients with locally advanced or metastatic urothelial cancer and is administered via the intravenous route.

In order to improve stability, prolong delivery, and decrease clearance of biologics after intravenous administration, chemical modifications such as PEGylation, hyperglycosylation, mannosylation, stapled peptides, and colloidal carriers (liposomes, microspheres, and nanoparticles) are widely explored strategies. PEGylation can be used as a strategy to alter properties such as molecular weight, solubility, and steric hindrance. Thus, it leads to improved stability and pharmacokinetic activity of peptides. Oncaspar® (pegaspargase) is an FDA-approved product with L-asparaginase (Lasparagine amidohydrolase) covalently conjugated to monomethoxypolyethylene glycol (mPEG). It is indicated as a component of a multiagent chemotherapeutic regimen and is administered via the intravenous route. Hyperglycosylation refers to a co- or post-translational enzymatic process that conjugates proteins, lipids, or other organic molecules with polysaccharides to form a glycoconjugate. This technique has been explored for peptide delivery via intravenous route and offers advantages such as prolonged systemic circulation and reduced immunogenicity (Patel et al. 2014). Mannosylation refers to mannose receptor-targeted delivery of proteins and peptides by conjugation with mannose to achieve improved pharmacokinetic and pharmacological properties (Patel et al. 2014). Other technologies such as stapled peptides have been developed to bypass the intrinsic problems related to peptides and to enhance their pharmacological performance. Stapled peptides are short peptides that are constrained in their alpha helical conformation by a synthetic brace, also called “staples.” They allow targeting the peptides inside the cells with high selectivity and efficiency (Moiola et al. 2019). ALRN-6924 is a stabilized cell-permeating peptide designed by Aileron Therapeutics, Inc. to disrupt the interaction between the p53 tumor suppressor protein and its predominant endogenous inhibitors. As of July 2020, ALRN-6924 is in Phase 2a clinical trial study that evaluates the antitumor effects in patients with advanced solid tumors or lymphomas (Kim and Jacobsen 2020). These technologies may thus play a pivotal role in the future of peptide therapeutics in the pharmaceutical field. Several delivery issues exist for peptide products, such as instability during storage and processing, instability in biological fluids, and poor intracellular delivery (Swaminathan and Ehrhardt 2012). Carrier systems such as liposomes, microspheres, and nanoparticles are some of the approaches to overcome these challenges in peptide delivery by conventional intravenous therapy. Liposomes are vesicles composed of one or more phospholipid bilayers with an aqueous internal cavity. Liposomes can encapsulate hydrophilic as well lipophilic components and hence have multiple applications as drug delivery systems. They can be formulated to be of different sizes, compositions, charge, and lamellarity (Bulbake et al. 2017). Microspheres can be used as depot systems for controlled and localized delivery of therapeutic peptides. Microspheres-based delivery systems are commonly fabricated by three methods: polymerization, emulsion-solvent extraction evaporation, and extrusion (Ramteke et al. 2012). There have been recent advances in the use of polymeric microspheres for the delivery of single-shot vaccines, plasmid DNA, and therapeutic peptides (Ramteke et al. 2012). Nanoparticles are microscopic particles with at least one dimension smaller than 100 nm. Nanoparticulate delivery systems have gained attention in delivering pharmaceutical actives in recent years.

Various products based on these controlled release technologies have been approved by FDA for parenteral administration of peptides. Lupron Depot® (TAP Pharmaceuticals) is a biodegradable, biocompatible polymeric microspheres-based product indicated for targeted and localized delivery of leuprolide acetate (a synthetic nonapeptide) to treat the symptoms of advanced prostate cancer, endometriosis, and uterine leiomyomata. Signifor Lar® (Novartis) is another product approved by the FDA in 2014 for pasireotide pamoate in PLGA microspheres for intravenous administration to treat acromegaly and Cushing's disease. Bydureon Bcise® (Exenatide) is an extended-release injectable suspension from AstraZeneca indicated to improve blood sugar control in adults with type 2 diabetes.

5.3 Peptide Therapeutics Administered via Oral Route

Peptides are typically administered through the parenteral route because of their poor bioavailability orally. While each route has its advantages and disadvantages, the oral route of administration is often the most preferred route of drug administration due to factors such as noninvasive, painless delivery, lower manufacturing cost, and high patient compliance (Drucker 2020). Glucagon-like peptide-1 (GLP-1) is an intestinal hormone that exerts profound effects in the regulation of glycemia, stimulating glucose-dependent insulin secretion, and various other pathways (Lim and Brubaker 2006). The FDA recently approved Rybelsus (Semaglutide) oral tablets to improve blood sugar control in adult patients with type 2 diabetes. This is the first GLP-1 product approved for oral use and is a landmark in peptide delivery by oral administration (Drucker 2020; Fda 2019). Research has been done for delivering other peptides such as calcitonin, insulin, human growth hormone, parathyroid hormone, and many other peptide therapeutics via the oral route (Liu and Dinh 2011). Trulance™ (Plecanatide), developed by Synergy Pharmaceuticals, is another peptide product that received FDA and global approval in 2017 for chronic idiopathic constipation and is administered via oral route (Al-Salama and Syed 2017).

Orally administered peptides follow the same route as ingested food and hence undergo extensive degradation and metabolism and have to cross other barriers before being absorbed into the bloodstream from the small intestine. Barriers in oral peptide delivery include enzymatic barriers of the digestive system, mucus and epithelial barriers of the intestine, paracellular transport, efflux pumps, and interindividual variability (Drucker 2020; Patel et al. 2014).

The gastric acid present in the stomach destabilizes and exposes the peptide bonds for enzymatic degradation in the small intestine. Various proteolytic enzymes such as pepsin, trypsin, and aminopeptidases in the stomach and small intestine promote the digestion of peptides to amino acids, thus complicating the delivery of peptides via oral route (Dan et al. 2020; Brown et al. 2020). Various factors such as molecular weight, structural flexibility, hydrophilicity, and a number of enzyme susceptible groups affect the extent of proteolysis in the peptide therapeutics. Primarily, the function of the gut epithelium is recognition and exclusion of any foreign

substances such as viruses, bacteria, and other pathogens which makes the lining of the gastrointestinal tract poses an even greater challenge for the delivery of peptides (Dan et al. 2020). The gut epithelium consists of a single layer of columnar epithelial cells supported by the lamina propria and the muscularis mucosae layers. The mucus secreted by these intestinal goblet cells, which cover the mucosal surfaces, comprises water and mucins, which are high molecular weight and heavily glycosylated proteins. This mucus membrane effectively traps larger macromolecules and blocks their access to the underlying epithelial cells and thus another barrier to peptide absorption (Drucker 2020; Dan et al. 2020).

Paracellular transport refers to the movement of molecules through spaces between epithelial cells regulated by tight junctions by passive diffusion. The size and charge of peptides generally exclude them from paracellular transport. On the other hand, the transcellular pathway has a large surface area of the brush border membrane available for absorption. Hence, most oral drugs are absorbed passively via the transcellular pathway (Dan et al. 2020). Efflux pumps pose another significant barrier to peptide absorption (Bruno et al. 2013). These are proteins of the ATP-binding cassettes superfamily and present on mature epithelial cells and play a role in multidrug resistance in humans. These efflux pumps can pump the peptide back to the lumen of the gastrointestinal tract (Bruno et al. 2013). Interindividual variability refers to the differences in the physiology of the gastrointestinal tract between individuals, including factors like the extent of mucus, enzyme production and expression, gastrointestinal motility, and gastric emptying time (Drucker 2020). Lastly, even after the peptide drug is absorbed, the first-pass metabolism in the gut and liver extensively reduces the fraction of the drug that reaches the systemic circulation (Bruno et al. 2013).

Multiple strategies have been pursued to overcome the barriers in the oral delivery of peptides. Zizzari et al. (2021) and Liu et al. (2011) have reviewed and highlighted many promising attempts to improve the oral bioavailability of peptides. Multiple oral delivery products of insulin, calcitonin, and parathyroid hormone which are in clinical development and are based on various strategies to enhance oral bioavailability of peptides have been described in these reviews (Liu and Dinh 2011; Dan et al. 2020; Zizzari et al. 2021).

Chemical modifications of peptide drugs can alter their physiological properties and improve enzyme stability and membrane permeation and minimize immunogenicity (Liu and Dinh 2011; Shaji and Varkey 2012). Such modifications include alteration of amino acid side chains, alteration of carbohydrate moieties in glycoproteins, and conjugation to fatty acid and other lipophilic moieties to increase hydrophobicity and introduction of protective groups to prevent degradation (Dan et al. 2020). For example, the addition of alpha-lipoic acid and palmitoyl derivative moiety to insulin resulted in protection against digestion by trypsin (Hashimoto et al. 1989). Another example of structural modification is the modification by deaminating and substituting a protective group in vasopressin to produce desmopressin (Dan et al. 2020). Desmopressin (1-desamino-8-D-arginine vasopressin; DDAVP) developed by Ferring Pharmaceuticals is a synthetic analog of human

hormone vasopressin. Desmopressin has greatly simplified the management of diabetes insipidus by offering ease of administration, safety, and tolerability (Kim et al. 2004).

Formulation additives can enhance peptide drug gastrointestinal absorption (Liu and Dinh 2011). The use of enzyme inhibitors resists degradation by enzymes in the gastrointestinal tract, while absorption enhancers improve membrane permeability. Protease inhibitors inhibit the enzymatic activity of the protease enzyme. The choice of protease inhibitors should depend on the amino acids in the peptide therapeutics since these inhibitors are sequence specific (Brown et al. 2020). Another approach to inhibit the enzymes is to alter the pH to inactivate the local digestive enzymes. For example, digestive enzymes such as trypsin, chymotrypsin, and elastase can be inactivated by a sufficient amount of pH lowering buffer that can lower the local pH of the intestine to below 4.5 (Liu and Dinh 2011; Shaji and Patole 2008). Absorption enhancers are components of the formulation that disrupt the intestinal barrier and aid the peptides in crossing the biological membrane and reaching systemic circulation (Shaji and Patole 2008). Numerous compounds such as surfactants, bile salts, chelating agents, fatty acids, alkanoylcholines, mucoadhesive polymers, acyl carnitine, lectins, and chitosans are used as absorption enhancers for peptide therapeutics (Bruno et al. 2013; Zizzari et al. 2021; Shaji and Patole 2008; Aungst 2000). Some chelating agents such as ethylenediaminetetraacetic acid (EDTA) and ethylene glycol tetraacetic acid (EGTA) can sequester metal ions to enhance the paracellular transport of peptides (Brown et al. 2020).

Strategies such as PEGylation and various formulation vehicles such as liposomes, nanoparticles, and microspheres discussed earlier for parenteral delivery have multiple applications in the oral delivery of peptides as well (Bruno et al. 2013; Fasano 1998; Cao et al. 2019). However, the renewal of the mucus layer due to periodic turnover every 4–6 hours can lead to the rapid clearing of entrapped particulate system. For example, BioOral system (BioSante Pharmaceuticals) is based on calcium phosphate-based nanotechnology (CAP) for facilitating vaccine delivery. CAP was observed as a safer adjuvant to improve the efficacy of vaccines as compared to other approved adjuvants (BioSante 2004; Dan et al. 2020). Insulin was formulated CAP-PEG-insulin-casein (CAPIC), to improve the oral bioavailability by protecting insulin against the low pH in stomach where casein encapsulates the insulin-PEG formulation and acts as an enteric coating. The use of CAP in cancer, infectious disease, autoimmune disease vaccines, malaria vaccine, and anthrax vaccine is investigated (BioSante 2004).

Hydrophilic mucoadhesive polymers (polyacrylates, cellulose, and chitosan) and thiolated polymers have also been used as peptide carrier systems (Bruno et al. 2013). Nanoemulsions are another type of carrier system specifically for oral peptide delivery. They are typically oil-in-water (o/w) or water-in-oil (w/o) emulsions with mean droplet diameters ranging from 50 to 1000 nm (Bruno et al. 2013). Another type of formulation in development are self-emulsifying drug delivery systems (SEDDS) that are composed of surfactants, lipids, and co-solvents. These systems have gained focus in recent years to improve oral delivery of peptides owing to their mechanism of forming emulsions and microemulsions when dispersed in GI

fluids (Zizzari et al. 2021). Sandimmune/Neoral® (Cyclosporin A) by Novartis Pharma AG is an FDA-approved marketed SEDDS-based product to prevent transplant rejection and treat patients with severe rheumatoid arthritis (Zizzari et al. 2021).

Hydrogels have high water content and a cross-linked network of hydrophilic polymer yet insoluble in water and can also be tailored for site-specific sustained oral drug delivery (Bindu Sri et al. 2012; Peppas et al. 2004). Hydrogels are biocompatible and have a high drug loading. Poly (methyl methacrylic acid), alginates, and chitosans are commonly used polymers for formulating hydrogels (Brown et al. 2020). Mucoadhesive polymeric systems are another approach for delivering peptides. These delivery systems adhere to the mucin layer and increase residence time at drug absorption site, thus decreasing drug clearance rate from the absorption site (Shaji and Patole 2008). Examples of mucoadhesive polymers include semi-natural polymers such as xanthan gum, polyvinyl pyrrolidone, polyvinyl alcohol, and synthetic polymers such as polyacrylic acid-based polymers and cellulose derivatives (Shaji and Patole 2008). In conclusion, despite being extremely challenging, oral delivery of peptides evolved with newer technologies. It can be an alternate and a more patient-compliant route of peptide therapeutics administration.

5.4 Peptide Therapeutics Administered via Pulmonary Route

In recent years, pulmonary route due to its unique and versatile features (Smith 1997) including (1) large surface area of the respiratory tract and lungs allowing simultaneous exposure of drug, (2) increased blood flow with thinner alveolar epithelium, (3) lesser metabolic activity, and (4) no hepatic metabolism has become an alternate route for delivery of therapeutic peptides via noninvasive route. As opposed to the oral route of delivery, pulmonary route has shown better bioavailability for peptide therapeutics with providing both local and systemic effects. However, delivery of peptides via respiratory tract has its challenges (Kwok et al. 2011). Mucociliary clearance is the primary challenge as it is the mechanism by which foreign particles get restricted from entering the body. Besides, the geometry and morphology of the airway and the presence of pulmonary peptidases are the other challenges for absorption of peptides (Smith 1997; Banga 2011). With a better understanding of these challenges, various approaches have been tried to deliver therapeutic peptides with better efficiency and reproducibility while maintaining their stability. One major technology for pulmonary delivery of peptides is the use of aerosols that is via inhalation (Shoyele and Slowey 2006). Nebulizers, metered-dose inhalers (MDI), and dry powder inhalers (DPI) are the most popular inhaled formulations in the market. The drug inhaled from aerosol gets deposited by gravitational sedimentation, inertial impaction, and diffusion (Shoyele and Slowey 2006). The aerodynamic behavior of particles, breathing patterns, and airway anatomy affects the distribution and absorption of drugs in the lungs (Liang et al. 2020). Although aerosols are most popular for pulmonary delivery of peptides, newer formulation

approaches such as liposomes and microparticulate systems have been investigated recently for controlled and targeted delivery of peptides (Wan et al. 2012).

Nebulizers are commonly used to deliver a large volume of dose for inhalation in the form of droplets. Depending on the mechanism of generation of droplets, nebulizers can be classified as (1) jet nebulizer which uses compressed air to produce droplets or (2) ultrasonic nebulizer in which a piezoelectric crystal below the liquid reservoir generates ultrasonic waves and leads to the creation of aerosol droplets. The creation of an air-water interface and repeated stress can lead to physical instability in jet nebulizers, whereas thermal stress can be an issue in the case of ultrasonic nebulizers (Banga 2005). Advances in nebulizer technologies led to the development of devices such as AERx (Aradigm, Hayward, CA) and Respimat (Boehringer, Germany), which are based on mechanical extrusion of liquid from unit dose, and more recent AeroDose (Aerogen Inc., Mountain View, CA) based on vibrating mesh technology which all are investigated for pulmonary delivery of peptides (Schuster et al. 1997; Perera et al. 2002; Cryan 2005). However, the only marketed formulation administered via jet nebulizer is Pulmozyme® containing dornase alfa to treat cystic fibrosis (Cryan 2005).

As the name suggests, MDIs are formulations containing the active substance and other excipients dissolved or suspended in a propellant system which, upon actuation, delivers a measured dose of drug in the form of aerosol spray. Advantages of MDI include their low cost, portability, dose reproducibility, and disposability (Shoyele and Slowey 2006). Propellants such as hydrofluoroalkanes (HFA) have replaced chlorofluorocarbons (Banga 2005). HFA-based MDI was investigated to deliver a high dose of cyclosporine using ethanol as co-solvent (Myrdal et al. 2004). In MDI, the inert propellant vapor atmosphere and sealed container protect the active from oxidative degradation and microbiological contamination, but the stability of peptides in these propellants remains a challenge (Banga 2005; Shoyele and Slowey 2006). Thus, many MDIs are formulated as a suspension. There is reported use of nonionic soluble surfactants to create a uniform suspension using freeze-dried protein and surfactant particles to formulate MDIs (Kwok et al. 2011). Synthetic nonapeptide leuprolide acetate is reported to be formulated as MDI both as solution and suspensions (Adjei and Gupta 1994). Studies conducted with human subjects with these formulations showed a three–four fold increase in plasma concentration from suspension as compared to a solution.

DPI is a system inhaled as a cloud of fine particles, and they do not contain propellant (Shoyele and Slowey 2006). These can be single or multiple doses where the drug is preloaded or loaded in the form of hard capsule (e.g., Spinhaler from Fisons Pharmaceuticals, Rochester, NY) or foil blister discs (e.g., Diskhaler from GSK, RTP, NC). Thus, they are easy to operate and inexpensive and have an advantage over nebulizers and MDI for administration of peptides (Banga 2005). Newer devices such as Spiros (Dura Pharmaceuticals, San Diego, CA) and Nektar T-326/TOBI®Podhaler (now owned by Novartis AG) use powder inhaler system which generates aerosol independent of inspiratory rate and volume, enhancing the efficiency of DPI for drug delivery (Cryan 2005). However, formulating a peptide as DPI faces other challenges, including powder flowability, dispersibility, and

stability (Banga 2005). One of the advances made into overcoming these challenges is drying the protein to an amorphous glassy state which inhibits denaturation. While this technology takes care of biochemical stability, the hygroscopic nature of dried particles can lead to physical instability (Kwok et al. 2011). Other ways of producing dry powders of peptides include milling, spray drying, spray freeze-drying, and supercritical fluid (Cryan 2005). Caveolin scaffolding domain peptide was investigated recently for treating pulmonary fibrosis as DPI using excipient free jet milling (Zhang et al. 2020). Cetrorelix acetate (decapeptide) was also reported to be formulated as DPI to study the role of particle engineering in detail (Zijlstra et al. 2004). Previously marketed inhalable insulin Exubera by Pfizer was the first-in-class product manufactured using spray drying and was administered as DPI with low-dose powder filing technology. However, it was withdrawn from the market due to several other factors, including its price, low sales, bulky device, and lack of insurance reimbursement (Banga 2005).

Like parenteral and oral routes, other strategies researched for pulmonary delivery of peptides include liposomes and microparticulate system (Wan et al. 2012). Due to their ability to encapsulate peptides, reduced local irritation, and toxicity with possible sustained release, liposomes are widely explored (Cryan 2005). They can be formulated as a liquid to be used with a nebulizer or in the form of dry powder for DPI. As mentioned earlier, essential parameters for formulation include selection and composition of lipids used and their ratio, charge, and size, which can be modulated to control drug release. Pulmonary insulin delivery has been previously investigated via liposomal carriers where insulin absorption and retention was increased using liposomes (Liu et al. 1993). Other peptides investigated for liposomal pulmonary delivery include cyclosporine, interleukin-2, and enzyme catalase (Cryan 2005). Microspheres and large porous particles are different microparticulate delivery systems explored for pulmonary delivery of peptides due to their ability to be formulated as powder and better stability. Polylactic acid (PLA), polylactic-co-glycolic acid (PLGA), chitosan, dextran, and gelatin are some of the polymers used to fabricate the microparticles. Large porous particles have an aerodynamic diameter of less than 5 μm owing to their low density and hence can be used for deep lung delivery (Banga 2005).

Factors to be considered while developing a peptide-based formulation for pulmonary delivery include pH of the formulation, buffering agents used, solubility, osmolarity, and special excipients such as absorption enhancers, protease inhibitors, and surfactants (Kwok et al. 2011). All these factors affect the conformation, stability, and absorption of peptides from the lungs. Absorption enhancers such as oleic acid, polyoxyethylene oleyl ether, bile acids, bile salts, fatty acids, and surfactants are reported to enhance absorption of peptides (Banga 2005). The use of protease inhibitors is investigated mainly for delivery of proteins rather than peptides. In a study conducted using rat lung homogenate, bacitracin showed highest effectiveness among the various protease inhibitors (sodium glycocholate, soybean trypsin inhibitor, aprotinin, and bacitracin) tested for delivery of insulin (Shin et al. 1994). In another study, pulmonary transport of polypeptide arginine vasopressin was enhanced with use of protease inhibitors (Hiroshi et al. 1994). Another critical

factor that needs to be considered for the formulation is the immunogenicity of the excipients used, where certain absorption enhancers and enzyme inhibitors can lead to immunogenic response (Kwok et al. 2011). Liposomes discussed earlier can help to encapsulate peptides, thus avoiding clearance triggered by the immunogenic response. The safety of excipients used in the formulation is of concern as only phosphatidylcholine is approved for inhalation as of now, whereas safety of other absorption enhancers remains to be tested (Cryan 2005).

5.5 Peptide Therapeutics Administered via Transdermal Route

Topical/transdermal delivery of peptides has gained attention due to its advantages, which include (1) large surface area, (2) bypassing hepatic metabolism and minimizing enzymatic degradation, (3) ease of administration leading to better patient compliance, and (4) controlled and sustained delivery (Herwadkar and Banga 2011). Thus, transdermal route is one of the routes explored as an alternative to parenteral administration, especially for peptides. Stratum corneum, the outermost layer of skin, is a major barrier to diffusion of drugs. An ideal drug administered via skin passively is a moderately lipophilic ($\log P \sim 1-3$) molecule with molecular weight <500 Da (Banga 2005). On the other hand, peptides are hydrophilic macromolecules, making passive delivery via the skin a challenge (Schuetz et al. 2005). Apart from permeation across skin, proteolytic enzymes in the skin are an enzymatic barrier to delivering peptides.

To overcome these challenges, various physical and chemical enhancement techniques are employed to enhance delivery of peptides into and across skin (Banga 2005). The physical enhancement techniques researched so far include microneedles, iontophoresis, electroporation, sonophoresis, laser ablation, thermal and radio-frequency ablation, jet injectors, and their combination. On the other hand, chemical enhancement technique uses various chemicals to alter permeation of molecules across skin.

Microneedles (MN) are micron-sized needles that create hydrophilic microchannels in the skin by offering a needle-free, pain-free mode of administration (Herwadkar and Banga 2011). This disruption of stratum corneum thus bypasses the major barrier for transdermal delivery enabling delivery of hydrophilic macromolecules such as peptides. Material of fabrication for microneedles, microneedle length, number of microneedles per unit area, microneedle structure/geometry, and type of application are among the critical factors determining successful delivery (Kirkby et al. 2020). Transdermal delivery of insulin via MN and the effect of MN lengths on delivery have been investigated (Zhou et al. 2010). Transdermal delivery of four hydrophilic peptides having different molecular weights was investigated using solid MN across porcine ear skin, and microneedles enhanced delivery of all four peptides (Zhang et al. 2014). Delivery of peptide salmon calcitonin via coated microneedles was found to be similar to subcutaneous administration in hairless rat

models (Banga 2005). The ability to deliver hydrophobic autoantigen peptides has been demonstrated recently using coated microneedles (Zhao et al. 2017). Another study used dissolving MN to deliver monoclonal IgG (Mönkäre et al. 2015) and interferon- α -2b (Chen et al. 2016). Swelling hydrogel-forming microneedles are among the new type of microneedles investigated for delivery of insulin and protein-based drug bevacizumab (Courtenay et al. 2018). With many examples investigated so far, MN is a promising approach for delivery of peptides; however, its commercialization will depend on developments in large-scale manufacturing to ensure safety, stability, and regulatory compliance.

Iontophoresis is another technique that uses a physiologically acceptable amount of current to deliver molecules into and across skin. As opposed to other techniques, iontophoresis works on the drug rather than on skin layers, with electro-osmosis and electro-repulsion being the two driving forces to deliver molecules (Banga 2005; Bakshi et al. 2020). Factors that need to be considered for iontophoretic delivery of peptides include physicochemical properties of peptide (isoelectric point, charge at physiological pH), type, amount, and duration of current applied. Peptides with a high charge-to-mass ratio can be delivered more efficiently. Due to skin pH between 4 and 7, peptides with isoelectric point below 4 and above 7.4 are better suited for transdermal delivery (Herwadkar and Banga 2011). The effect of pH on delivery of peptide leuprolide was investigated, where a pH of 7.4 was found to be optimal for delivery (Kochhar and Imanidis 2004). In another study, iontophoresis enabled delivery of leuprolide in vivo in humans to achieve a response similar to subcutaneous injection (Meyer et al. 1990). Iontophoretic delivery of various peptides, including cyclosporine, angiotensin, octreotide, arginine vasopressin, nafarelin, luteinizing hormone-releasing hormone (LHRH), and thyrotropin-releasing hormone, has been investigated (Schuetz et al. 2005). While mild symptoms including redness and erythema associated with iontophoresis are well-tolerated, success of the available marketed iontophoretic device will determine development of iontophoretic delivery of peptides in the near future.

Laser ablation, thermal microporation, and radiofrequency-assisted ablation are different ablation technologies explored for drug delivery, including peptides (Benson and Namjoshi 2008). Laser ablation involves using a high laser beam that leads to water evaporation on the skin surface, thereby creating microchannels. These have been investigated for delivery of various macromolecules. For example, Nelson et al. observed a 2.1-fold increase in delivery of peptide INF- γ using an erbium-YSGG laser across porcine skin (Nelson et al. 1991). As the name suggests, thermal ablation uses heat for a short amount of time to create micropores. Delivery of interferon- α -2b using thermal ablation across hairless rats was found to be comparable to subcutaneous injection (Badkar et al. 2007). The use of the Passport™ system using thermal ablation to deliver insulin has also been reported (Benson and Namjoshi 2008). Similar to thermal ablation, radio frequency-assisted thermal ablation uses radiofrequency generated energy for ablation. The ViaDor™ system developed by Trans Pharma uses this technique to deliver drugs into and across skin which is tested for human growth hormone (Levin et al. 2005). For all the ablation

technologies, safety and patient compliance remain to be tested for successful delivery system.

Other technologies to deliver peptides via transdermal route include jet injectors, electroporation, sonophoresis, and chemical enhancers. A jet injector is a velocity-based technology that uses a high velocity (>100 m/s) jet to penetrate skin. There are different jet injectors developed for liquid and powder formulations (Schuetz et al. 2005). Some of the jet injectors developed include Vitajet™, Medi-jector®, and Zomajet®, which are investigated for delivery of insulin, human growth hormone, and erythropoietin (Benson and Namjoshi 2008). Other dry powder formulation injector uses finer particles of drug in supersonic flow to penetrate the skin. Successful delivery of salmon calcitonin *in vivo* in rabbits was reported using PowderJect® device, which was further tested for insulin delivery in rats (Benson and Namjoshi 2008). The maximum dose that can be delivered using the jet injectors and stability of formulations are some of the barriers that limit the application of this technology for peptide delivery.

Electroporation uses high voltage pulses (60–1000 V) for a short duration (μ to ms) to create pores in lipid bilayer (Herwadkar and Banga 2011). Physiological properties of drug and formulation and the electrical parameters such as pulse rate, duration, and voltage affect peptide delivery. Peptide vaccine has been investigated for delivery using electroporation in mice (Zhao et al. 2006). Electroporation-assisted delivery of luteinizing hormone-releasing hormone has been reported *in vitro* across porcine skin in combination with iontophoresis, where increased delivery was observed with an increase in number of pulses of electroporation (Riviere et al. 1995). A combination of electroporation and iontophoresis was also used for delivery of salmon calcitonin (Chang et al. 2000) and human parathyroid hormone (Medi and Singh 2003). The nature of micropores (reversibility) and concern about damage to skin limit application of electroporation for peptide delivery.

Sonophoresis, as the name suggests, uses ultrasonic perturbation to drive molecules into and across skin. It employs piezoelectric crystal, which produces acoustic waves. The resulting acoustic cavitation is believed to play a major role in enhancing drug delivery (Dragicevic and Maibach 2017). The ultrasonic frequency, duration of application, intensity, and pulse length are the parameters that need to be optimized for achieving target delivery. A low frequency is preferred for permeation enhancement. A study was conducted to deliver cyclosporin A using sonophoresis alone and in combination with electroporation and chemical enhancers. It was observed that the skin delivery of cyclosporin was enhanced with sonophoresis alone and in combination with chemical enhancers as compared to passive diffusion. However, a significant enhancement in systemic delivery was observed only when electroporation was combined with sonophoresis and chemical enhancers (Liu et al. 2006). For sonophoresis-assisted delivery of insulin, the effect of intensity, duration, and threshold energy to increase permeation has been tested in several studies (Dragicevic and Maibach 2017). Other peptides and proteins explored for delivery using this technology include vasopressin, interferon- γ , and erythropoietin. Although various devices such as SonoDerm™ Technology and Sonoprep®

were developed, heavy and oversized instrumentation and flexibility limit the use of sonophoresis in transdermal delivery (Benson and Namjoshi 2008).

The use of chemical enhancers is another strategy traditionally used to deliver small molecules, and it is also utilized to deliver peptides transdermally. These chemicals, when applied, change the permeability of skin either by increasing solubility of drug or chemical enhancer in lipid layer or by disrupting the barrier of skin. Polyalcohols, esters, fatty acids, pyrrolidones, sulfoxides, amines, amides, surfactants, and phospholipids are various chemical enhancers that can be used to enhance transdermal delivery of peptides (Herwadkar and Banga 2011). For example, ethanol alone and combined with cineole is reported to enhance the delivery of thyrotropin-releasing hormone analog (M-TRH) across human epidermal membrane (Magnusson and Runn 1999). Another strategy is the use of synthetic peptides to be administered along with large protein molecules. For example, a short synthetic peptide was found to enhance permeation of insulin and human growth hormone (Chen et al. 2006). Although chemical enhancers are well known for enhancement, their use for enhancing delivery of hydrophilic molecules such as peptides is limited.

Formulation approaches for transdermal delivery of peptides include encapsulation technologies, use of protease inhibitors, and prodrug/chemical modification. The encapsulation technology uses a carrier system to deliver peptide transdermally, whereas use of protease inhibitors is the same as mentioned earlier. In the prodrug approach, hydrophilic peptide molecules are modified by conjugating it with lipid moiety or by derivatization, enabling delivery via skin. For example, a 2.5–5-fold increase in delivery of INF- α was observed when converted to acyl derivative for in vitro permeation across human skin (Foldvari et al. 1999). Similarly, delivery of insulin and thyrotropin hormone was also tested with fatty acid derivative (Herwadkar and Banga 2011). As this approach leads to creation of a new chemical entity, additional testing might be required from a regulatory perspective. Other novel and emerging technologies yet to be explored for transdermal delivery of peptides are photochemical waves, heat-assisted drug delivery, microscissioning, and ionic liquids (Schuetz et al. 2005).

5.6 Peptide Therapeutics Administered via Other Routes

Other routes for administration of peptides include nasal, rectal, vaginal, and buccal routes. The nasal route for peptide administration offers advantages such as rapid absorption, bypassing presystemic clearance, and ease of administration. Two approaches are commonly used for nasal delivery of peptides, namely, use of absorption enhancers to modify the permeability of the nasal membrane and use of mucoadhesive systems to decrease mucociliary clearance, and thereby increase the contact time at the site of absorption (Jitendra et al. 2011). Another potential route for administering peptide therapeutics is via the rectum. The lower part of the rectum is connected directly to systemic circulation, which also offers an opportunity

to enter directly into the systemic circulation. In addition, it has many lymphatic vessels and thus can be a target to the lymphatic system route (Sanders 1990). Vaginal route can also be a favorable alternative for systemic drug delivery due to rich blood supply and large surface area of the vagina. Research has been conducted to deliver peptides such as calcitonin, human growth hormone, oxytocin, and insulin via the vaginal route (Jitendra et al. 2011). However, factors such as cultural sensitivity, personal hygiene, local irritation, and gender specificity can limit the vaginal route of administration of peptide therapeutics (Jitendra et al. 2011). Finally, the buccal route of administration is known to be an alternative to the conventional oral route, wherein the formulation can stick to the buccal mucosa and maintain a prolonged contact due to the mucoadhesive polymers. The formulation for delivery via buccal route also contains additives such as penetration enhancers to improve peptide permeation across buccal mucosa and enzyme inhibitors to protect the peptides from mucosal enzymes (Jitendra et al. 2011).

5.7 Summary

- *Parenteral Route:* Primary and most promising with choice of IV/IM/SC administration but faces challenges including stability, short half-life, and poor bioavailability for peptide delivery. Various formulation approaches including chemical modifications and colloidal carriers have been explored for successful delivery with products in market.
- *Oral Route:* Noninvasive and highly patient compliant route of administration. However, enzymatic degradation in GI tract along with other factors leads to poor bioavailability of peptides. Chemical modification and use of enzyme inhibitors and absorption enhancers along with several other approaches such as mucoadhesive and carrier systems have been utilized for therapeutic peptide delivery.
- *Pulmonary Route:* A noninvasive route providing both local and systemic effects, with improved bioavailability compared to oral administration. Morphology and geometry of airway, mucociliary clearance, and immunogenicity limits peptide delivery. Various microparticulate systems and liposomes have been explored with strategic use of absorption enhancers and protease inhibitors for enhanced peptide delivery via aerosolized systems.
- *Transdermal Route:* Another noninvasive route which can provide sustained and controlled delivery. However, permeation of hydrophilic peptides across skin is a major challenge. Various physical enhancement techniques such as microneedles, iontophoresis, and sonophoresis, encapsulation technologies, and chemical modifications have been explored for therapeutic peptide delivery.
- *Future Aspects:* Extensive research and technological advances in various routes of administration have shown promising results, and this may broaden the reach of peptide therapeutics in the future.

References

- Adjei A, Gupta P. Pulmonary delivery of therapeutic peptides and proteins. *J Control Release*. 1994;29:361–73.
- Agarwal P, Rupenthal ID. Injectable implants for the sustained release of protein and peptide drugs. *Drug Discov Today*. 2013;18:337–49.
- Al Musaimi O, Al Shaer D, Albericio F, de la Torre BG. 2020 FDA TIDES (peptides and oligonucleotides) harvest. *Pharmaceuticals*. 2021;14:1–14.
- Al-Salama ZT, Syed YY. Plecanatide: first global approval. *Drugs*. 2017;77:593–8.
- Asfour MH. Advanced trends in protein and peptide drug delivery: a special emphasis on aquasomes and microneedles techniques. *Drug Deliv Transl Res*. 2021;11:1–23.
- Aungst BJ. Intestinal permeation enhancers. *J Pharm Sci*. 2000;89:P429–42.
- Badkar AV, Smith AM, Eppstein JA, Banga AK. Transdermal delivery of interferon alpha-2b using microporation and iontophoresis in hairless rats. *Pharm Res*. 2007;24:1389–95.
- Bakshi P, Vora D, Hemmady K, Banga AK. Iontophoretic skin delivery systems: success and failures. *Int J Pharm*. 2020;586:119584.
- Banga AK. Therapeutic peptides and proteins formulation, processing, and delivery systems. Boca Raton: CRC Press; 2005.
- Banga AK. Transdermal and intradermal delivery of therapeutic agents: application of physical technologies, vol. 38. Boca Raton: CRC Press; 2011. p. 1–282.
- Benson HAE, Namjoshi S. Proteins and peptides: strategies for delivery to and across the skin. *J Pharm Sci*. 2008;97:3591–610.
- Bindu Sri M, Ashok V, Arkendu C. As a review on hydrogels as drug delivery in the pharmaceutical field. *Int J Pharm Chem Sci*. 2012;1:642–61.
- BioSante. CAP nanotechnology product development – new product. February 5, 2004.
- Brown TD, Whitehead KA, Mitragotri S. Materials for oral delivery of proteins and peptides. *Nat Rev Mater*. 2020;5:127–48.
- Bruno BJ, Miller GD, Lim CS. Basics and recent advances in peptide and protein drug delivery. *Ther Deliv*. 2013;4:1443–67.
- Bulbake U, Doppalapudi S, Kommineni N, Khan W. Liposomal formulations in clinical use: an updated review. *Pharmaceutics*. 2017;9:1–33.
- Cao SJ, Xu S, Wang HM, Ling Y, Dong J, Xia RD, Sun XH. Nanoparticles: oral delivery for protein and peptide drugs. *AAPS PharmSciTech*. 2019;20:1–11.
- Chang SL, Hofmann GA, Zhang L, Defetos LJ, Banga AK. The effect of electroporation on iontophoretic transdermal delivery of calcium regulating hormones. *J Control Release*. 2000;66:127–33.
- Chen Y, Shen Y, Guo X, Zhang C, Yang W, Ma M, Liu S, Zhang M, Wen LP. Transdermal protein delivery by a coadministered peptide identified via phage display. *Nat Biotechnol*. 2006;24:455–60.
- Chen J, Qiu Y, Zhang S, Gao Y. Dissolving microneedle-based intradermal delivery of interferon- α -2b. *Drug Dev Ind Pharm*. 2016;42:890–6.
- Courtenay AJ, McCrudden MTC, McAvoy KJ, McCarthy HO, Donnelly RF. Microneedle-mediated transdermal delivery of bevacizumab. *Mol Pharm*. 2018;15:3545–56.
- Cryan SA. Carrier-based strategies for targeting protein and peptide drugs to the lungs. *AAPS J*. 2005;7:E20–41.
- D'Hondt M, Bracke N, Taevernier L, Gevaert B, Verbeke F, Wynendaele E, De Spiegeleer B. Related impurities in peptide medicines. *J Pharm Biomed Anal*. 2014;101:2–30.
- Dan N, Samanta K, Almoazen H. An update on pharmaceutical strategies for Oral delivery of therapeutic peptides and proteins in adults and pediatrics. *Children*. 2020;7:307.
- Dragicevic N, Maibach HI. Percutaneous penetration enhancers physical methods in penetration enhancement. Berlin Heidelberg: Springer; 2017;1:1–508.
- Drucker DJ. Advances in oral peptide therapeutics. *Nat Rev Drug Discov*. 2020;19:277–89.
- Fasano A. Innovative strategies for the oral delivery of drugs and peptides. *Trends Biotechnol*. 1998;16:152–7.

- FDA. FDA approves first oral GLP-1 treatment for type 2 diabetes inquiries. FDA News Release. 2-4;2019.
- Foldvari M, Baca-Estrada ME, He Z, Hu J, Attah-Poku S, King M. Dermal and transdermal delivery of protein pharmaceuticals: lipid-based delivery systems for interferon alpha. *Biotechnol Appl Biochem*. 1999;30:129-37.
- Hashimoto M, Takada K, Kiso Y, Muranishi S. Synthesis of palmitoyl derivatives of insulin and their biological activities. *Pharm Res*. 1989;6:171-6.
- Herwadkar, Anushree, and Ajay K. Banga. 2011. *Transdermal delivery of peptides and proteins* (Elsevier).
- Hiroshi Y, Morimoto K, Lee VHL, Kwang-Jin K. Effect of protease inhibitors on vasopressin transport across rat alveolar epithelial cell monolayers. *Pharm Res*. 1994;11:1617-22.
https://www.accessdata.fda.gov/drugsatfda_docs/label/2011/021567s0261bl.pdf.
- Ibeanu N, Egbu R, Onyekuru L, Javaheri H, Khaw PT, Williams GR, Brocchini S, Awwad S. Injectables and depots to prolong drug action of proteins and peptides. *Pharmaceutics*. 2020;12:1-42.
- Jitendra PK, Sharma SB, Banik A. Noninvasive routes of proteins and peptides drug delivery. *Indian J Pharm Sci*. 2011;73:367-75.
- Kim A, Jacobsen E. Anaplastic large cell lymphoma. In: *Novel therapeutics for rare lymphomas*. Cham: Springer; 2020.
- Kim RJ, Malattia C, Allen M, Moshang T Jr, Maghnie M. Vasopressin and desmopressin in central diabetes insipidus: adverse effects and clinical considerations. *Pediatr Endocrinol Rev*. 2004;2:115-23.
- Kirkby M, Hutton ARJ, Donnelly RF. Microneedle mediated transdermal delivery of protein, peptide and antibody based therapeutics: current status and future considerations. *Pharm Res*. 2020;37:1-18.
- Kochhar C, Imanidis G. In vitro transdermal iontophoretic delivery of leuprolide under constant current application. *J Control Release*. 2004;98:25-35.
- Kwok L, Chi P, Chan HK. *Pulmonary delivery of peptides and proteins*. Elsevier; Amsterdam, 2011.
- Lau JL, Dunn MK. Therapeutic peptides: historical perspectives, current development trends, and future directions. *Bioorg Med Chem*. 2018;26:2700-7.
- Levin G, Gershonowitz A, Sacks H, Stern M, Sherman A, Rudaev S, Zivin I, Phillip M. Transdermal delivery of human growth hormone through RF-microchannels. *Pharm Res*. 2005;22:550-5.
- Liang W, Pan HW, Vllasaliu D, Lam JKW. Pulmonary delivery of biological drugs. *Pharmaceutics*. 2020;12:1-28.
- Lim GE, Brubaker PL. Glucagon-like peptide 1 secretion by the L-cell: the view from within. *Diabetes*. 2006;55:S70-7.
- Liu P, Dinh S. Oral delivery of protein/peptide therapeutics. In: *Oral bioavailability: basic principles, advanced concepts, and applications*. Hoboken: Wiley; 2011. p. 371-80.
- Liu FY, Shao Z, Kildsig DO, Mitra AK. Pulmonary delivery of free and liposomal insulin. *Pharm Res*. 1993;10(2):228-32.
- Liu H, Li S, Pan W, Wang Y, Han F, Yao H. Investigation into the potential of low-frequency ultrasound facilitated topical delivery of Cyclosporin A. *Int J Pharm*. 2006;326:32-8.
- Magnusson BM, Runn P. Effect of penetration enhancers on the permeation of the thyrotropin releasing hormone analogue pGlu-3-methyl-His-Pro amide through human epidermis. *Int J Pharm*. 1999;178:149-59.
- Medi BM, Singh J. Electronically facilitated transdermal delivery of human parathyroid hormone (1-34). *Int J Pharm*. 2003;263:25-33.
- Meyer BR, Kreis W, Eschbach J, O'Mara V, Rosen S, Sibalis D. Transdermal versus subcutaneous leuprolide: a comparison of acute pharmacodynamic effect. *Clin Pharmacol Ther*. 1990;48:340-5.
- Moiola M, Memeo MG, Quadrelli P. Stapled peptides-a useful improvement for peptide-based drugs. *Molecules*. 2019;24:3654

- Mönkäre J, Reza Nejadnik M, Baccouche K, Romeijn S, Jiskoot W, Bouwstra JA. IgG-loaded hyaluronan-based dissolving microneedles for intradermal protein delivery. *J Control Release*. 2015;218:53–62.
- Myrdal PB, Karlage KL, Stein SW, Brown BA, Haynes A. Optimized dose delivery of the peptide cyclosporine using hydrofluoroalkane-based metered dose inhalers. *J Pharm Sci*. 2004;93:1054–61.
- Nelson JS, McCullough JL, Wright GTC, Liaw LH, Wh JSL. Mid-infrared laser ablation of stratum corneum enhances in vitro percutaneous transport of drugs. *J Invest Dermatol*. 1991;97:874–9.
- Patel A, Cholkar K, Mitra AK. Recent developments in protein and peptide parenteral delivery approaches. *Ther Deliv*. 2014;5:337–65.
- Peppas NA, Wood KM, Blanchette JO. Hydrogels for oral delivery of therapeutic proteins. *Expert Opin Biol Ther*. 2004;4:881–7.
- Perera AD, Kapitza C, Nosek L, Fishman RS, David a Shapiro, Tim Heise, and Lutz Heinemann. Absorption and metabolic effect of inhaled insulin. *Diabetes Care*. 2002;25:2276–81.
- Ramteke KH, Jadhav VB, Dhole SN. Microspheres: as carriers used for novel drug delivery system. *IOSRPHR*. 2012;2:44–8.
- Riviere JE, Monteiro-Riviere NA, Rogers RA, Bommannan D, Tamada JA, Potts RO. Pulsatile transdermal delivery of LHRH using electroporation: drug delivery and skin toxicology. *J Control Release*. 1995;36:229–33.
- Sachdeva S, Lobo S, Tarun G. What is the future of noninvasive routes for protein- and peptide-based drugs? *Ther Deliv*. 2016;7:355–7.
- Sanders LM. Drug delivery systems and routes of administration of peptide and protein drugs. *Eur J Drug Metab Pharmacokinet*. 1990;15:95–102.
- Schuetz YB, Naik A, Guy RH, Kalia YN. Emerging strategies for the transdermal delivery of peptide and protein drugs. *Expert Opin Drug Deliv*. 2005;2:533–48.
- Schuster J, Rubsamen R, Lloyd P, Lloyd J. The AER(x)(TM) aerosol delivery system. *Pharm Res*. 1997;14:354–7.
- Shaji J, Patole V. Protein and peptide drug delivery: Oral approaches. *Indian J Pharm Sci*. 2008;70:269–77.
- Shaji J, Varkey D. Recent advances in physical approaches for transdermal penetration enhancement. *Curr Drug Ther*. 2012;7:184–97.
- Shin Y-S, Mori T, Okita M, Gemma T, Kai C, Mikami T. NII-electronic library service. *Chem Pharm Bull*. 1994;17:1466–2.
- Shoyele SA, Slowey A. Prospects of formulating proteins/peptides as aerosols for pulmonary drug delivery. *Int J Pharm*. 2006;314:1–8.
- Smith PL. Peptide delivery via the pulmonary route: a valid approach to local and systemic delivery. *J Control Release*. 1997;46:99–106.
- Swaminathan J, Ehrhardt C. Liposomal delivery of proteins and peptides. *Expert Opin Drug Deliv*. 2012;9:1489–503.
- Tesauro D, Accardo A, Diaferia C, Milano V, Guillon J, Ronga L, Rossi F. Peptide-based drug-delivery systems in biotechnological applications: recent advances and perspectives. *Molecules*. 2019;24:1–27.
- Wan F, Møller EH, Yang M, Jørgensen L. Formulation technologies to overcome unfavorable properties of peptides and proteins for pulmonary delivery. *Drug Discov Today Technol*. 2012;9:e141–e46.
- Zhang S, Qiu Y, Gao Y. Enhanced delivery of hydrophilic peptides in vitro by transdermal microneedle pretreatment. *Acta Pharm Sin B*. 2014;4:100–4.
- Zhang Y, Mackenzie B, Koleng JJ, Maier E, Warnken ZN, Williams RO. Development of an excipient-free peptide dry powder inhalation for the treatment of pulmonary fibrosis. *Mol Pharm*. 2020;17:632–44.
- Zhao YL, Murthy SN, Manjili MH, Guan LJ, Sen A, Hui SW. Induction of cytotoxic T-lymphocytes by electroporation-enhanced needle-free skin immunization. *Vaccine*. 2006;24:1282–90.

- Zhao X, Coulman SA, Hanna SJ, Susan Wong F, Dayan CM, Birchall JC. Formulation of hydrophobic peptides for skin delivery via coated microneedles. *J Control Release*. 2017;265:2–13.
- Zhou CP, Liu YL, Wang HL, Zhang PX, Zhang JL. Transdermal delivery of insulin using microneedle rollers in vivo. *Int J Pharm*. 2010;392:127–33.
- Zijlstra GS, Hinrichs WLJ, de Boer AH, Frijlink HW. The role of particle engineering in relation to formulation and de-agglomeration principle in the development of a dry powder formulation for inhalation of cetorelix. *Eur J Pharm Sci*. 2004;23:139–49.
- Zizzari AT, Pliatsika D, Gall FM, Fischer T, Riedl R. New perspectives in oral peptide delivery. *Drug Discov Today*. 2021;26:1097–105.

Chapter 6

Liposome Nanocarriers for Peptide Drug Delivery



Jafrin Jobayer Sonju, Achyut Dahal, and Seetharama D. Jois

Contents

6.1	Introduction.....	204
6.1.1	Limitations of Peptides as Therapeutics.....	205
6.2	Nanoparticle in Drug Delivery.....	206
6.2.1	Dendrimers.....	207
6.2.2	Micelles.....	207
6.2.3	Carbon Nanotubes.....	207
6.2.4	Quantum Dots.....	208
6.2.5	Liposomes.....	208
6.3	Liposome.....	209
6.4	Formulation and Manufacturing Strategies of Liposome.....	210
6.5	Characterization of Liposomes.....	212
6.6	Liposome Stabilization Strategy with Lyophilization.....	214
6.7	Liposomal Nanocarrier System for Peptide Drug Delivery.....	215
6.8	Types of Liposome.....	221
6.8.1	Active Targeting Liposomes.....	221
6.8.2	Stimuli-Responsive Liposomes.....	221
6.8.3	Temperature-Sensitive Liposomes (Thermosensitive Liposomes).....	222
6.8.4	pH-Sensitive Liposomes.....	222
6.8.5	Magnetic Field-Responsive Liposomes.....	223
6.8.6	Ultrasound-Responsive Liposomes.....	223
6.8.7	Light-Sensitive Liposomes.....	223
6.9	Limitation of Liposomes.....	224
6.9.1	Reticuloendothelial System (RES) and Liposome Clearance.....	224
6.9.2	Accelerated Blood Clearance Phenomenon.....	224
6.9.3	High Serum Protein Binding.....	224
6.9.4	Masking of Surface Ligands by Polymers.....	225

Authors Jafrin Jobayer Sonju and Achyut Dahal have equally contributed to this chapter

J. J. Sonju · A. Dahal · S. D. Jois (✉)
School of Basic Pharmaceutical and Toxicological Sciences, College of Pharmacy, University
of Louisiana at Monroe, Monroe, LA, USA
e-mail: jois@ulm.edu

6.9.5	Difficulty in the Accurate Characterization of Surface-Functionalized Liposomes.....	225
6.9.6	Large-Scale Production of Surface-Functionalized Liposomes.....	225
6.9.7	Stringent Storage Conditions.....	225
6.9.8	Aggregation of Liposomes.....	226
6.9.9	High Cost.....	226
6.9.10	Recent Examples of Liposome-Mediated Peptide Drug Delivery in Clinical Trials.....	226
6.10	Summary.....	226
	References.....	227

Abstract Peptide drugs have limitations, including low membrane permeability, inability to cross the blood-brain barrier, low stability towards enzymatic degradation, lower plasma half-life, and low oral bioavailability. Nanoparticles can effectively encapsulate various hydrophobic and hydrophilic drugs, protein-based drugs, peptides, and nucleic acids. Entrapment of these drugs can improve their solubility and stability. Nanoparticles can be developed to release the drug at the target site using stimulus trigger release. The different nanoparticles-based system serves this purpose, and they can be broadly classified as dendrimers, micelles, liposomes, carbon nanotubes, and quantum dots. Therapeutic peptides can be incorporated inside the liposomes to enhance stability and better tumor accumulation. Peptide-based liposomes can successfully target the tumor cells and lower the off-target effects of chemotherapeutics. Toxicity to normal cells caused by anticancer therapeutics has been dramatically reduced by the use of peptide-based liposomes. This chapter covers the fundamentals of incorporating peptides in liposomal particles and characterizing them using different methods. Examples of peptide-based liposomal delivery is also discussed.

Keywords Nanoparticles · Liposomes · Peptides · Drug delivery · Targeted liposome

6.1 Introduction

Nanotechnology can be defined as a branch of science based on the development of technology in the synthesis, manipulation, and study of materials and devices in the nanometer size range. Nanotechnology is applicable to a wide range of disciplines from basic materials science to personal care and therapeutic and diagnostic applications (Park 2007). Significant research and development for the medical application of nanotechnology (nanomedicine) in the last decades has provided a wide range of biomedical applications (Kawasaki and Player 2005), including diagnosis and as a tool for the treatment of various diseases (De Jong and Borm 2008). The development of novel approaches to effective drug delivery is one of the most promising applications of nanomedicine. The prospect of nanoparticles from adenoviral

vector to lipid capsules as nanocarrier in vaccines against Covid-19 (Tenchov et al. 2021; Shin et al. 2020) has shown the tremendous impact and success of nanomedicine.

Approval of medicinal use of insulin in 1922 as the first peptide-based treatment for diabetes had opened the immense potential of peptides as therapeutics in different diseases. Several peptide drugs are on the market, and many more are in clinical development (Muttenthaler et al. 2021; Henninot et al. 2018). Peptide drugs offer several advantages such as ease of synthesis, low costs, low immunogenicity, natural biological messengers of various pathways, and targeting of protein-protein interactions.

6.1.1 Limitations of Peptides as Therapeutics

Along with so many advantages of peptides, their limitations such as low membrane permeability, inability to cross the blood-brain barrier, low stability towards enzymatic degradation, lower plasma half-life, and low oral bioavailability are the hurdles for peptide-based drug development (Cao et al. 2019). *In vivo* stability is one of the major barriers to peptide drug delivery. Peptides' bioavailability is reduced through degradation by various protease enzymes along with other digestive enzymes. Various processes are being explored to increase the plasma stability of peptides to render enough therapeutic efficacy (Otvos Jr and Wade 2014). Peptides have only 1% oral bioavailability, with few exceptions (e.g., cyclosporine A) (Zhou and Po 1991). Along with cellular proteases, proteases such as trypsin, pancreatic esterase, and α -chymotrypsin, secreted from the pancreas, are present in large quantities in the small intestinal lumen. These proteases are responsible for the degradation of the majority of peptides (Vlieghe et al. 2010). Due to the size of the peptide molecules and the polar nature of peptide bonds, permeability across the cell membrane becomes challenging. Absorption of a peptide across the intestinal barrier can occur by (i) passive diffusion through the lipid layer, (ii) the paracellular pathway, and (iii) transporters [e.g., peptide transporter 1 (PEPT1), vitamin B12 transport system] (Edmonds and Price 2013). Various methods are being investigated to increase peptides' plasma stability to have enough biological effect. PEGylation, lipidation, and glycosylation processes have shown a sufficient effect on peptide stability (Morimoto 2017). Several formulation strategies have been considered to tackle the poor oral bioavailability of peptides and proteins (Vlieghe et al. 2010). One of these strategies is the use of substances that can assist the absorption of drugs, enhancing oral bioavailability. Various absorption enhancers have been examined for the improvement of peptide and protein absorption; these can be categorized into cationic and anionic agents, surfactants, bile salts, fatty acids, chelating agents, acylcarnitines, and their derivatives (Renkuntla et al. 2013). Synergistic effects can be observed from combination of these enhancers rather than a single enhancer. Co-administration of protease inhibitors can prevent the degradation of protein and peptides in the gastrointestinal tract and reduce enzymatic blockade.

Inhibitors for major digestive enzymes such as aprotinin and inhibitors for aminopeptidase, namely bestatin, puromycin, and boroleucine, have been used widely to prevent the degradation of peptides *in vivo* as well as for oral delivery of peptides. Physiologically responsive hydrogels can protect peptide degradation through their three-dimensional mesh-like structure and are also capable of reacting to surrounding stimuli such as ionic strength difference, temperature, and pH alterations (Lowman et al. 1999). The oral bioavailability of protein and peptide therapeutics can be supported through mucoadhesive polymer systems, which contain natural or synthetic polymers that enable them to adhere to the mucin layer on the mucosal epithelium. Incorporation of cyclodextrins with a hydrophobic interior and a hydrophilic outer side (Challa et al. 2005; Kanwar et al. 2011) has the potential to interact with guest molecules and serve as a drug delivery vehicle for large molecules like peptide or protein (Irie and Uekama 1999; Renukuntla et al. 2013).

6.2 Nanoparticle in Drug Delivery

Nanoparticles can effectively encapsulate various hydrophobic and hydrophilic drugs, protein-based drugs, peptides, and nucleic acids. Entrapment of these drugs can improve their solubility and stability. Nanoparticles can be developed to release the drug in the target site by the use of stimulus trigger release. Nanoparticles can also be functionalized on their surface by peptides, antibodies, and aptamers for active targeting and for diagnostic purposes (Kim et al. 2006; Sonju et al. 2021). In addition to this, nanoparticles can be designed to circulate in the blood for a longer time, improving the biodistribution properties of the drugs. Due to their size, nanoparticles can easily pass through the endothelium and accumulated in inflammatory sites like tumors (Moghimi et al. 2001). This property of nanoparticles makes them effective nanocarriers and also reduces the toxicity of free drugs resulting from off-target effects (Singh and Lillard Jr. 2009). The different nanoparticles-based systems can be broadly classified as dendrimers, micelles, liposomes, carbon nanotubes, and quantum dots.

Advantages of Nanoparticles as Carriers of Peptides

- The incorporation of peptides in nanoparticles improves the stability as it protects the peptide from enzymatic degradation
- Nanoparticles as nanocarriers of a peptide can be used for control release and for targeting specific effects.
- Oral delivery of peptides using a nanoparticle drug delivery system is a promising platform for peptide therapeutics.
- Nanoparticles as nanocarrier of peptide drugs can change/enhance the biodistribution and pharmacokinetic properties.
- Nanoparticles can incorporate peptides as cargo or as a targeting agent.

6.2.1 *Dendrimers*

Dendrimers are radially symmetric well-defined artificial molecules with a symmetric core, an inner shell, and an outer shell. Dendrimers are hyper-branched structures characterized by a high number of functional groups and a compact molecular architecture. The molecular structure of dendrimer consists of the central core with a single atom or group of atoms from which the branches of atoms, also called dendrons, are produced by various chemical reactions to form a homogeneous and the monodisperse structure consisting of tree-like arms or branches (Tomalia and Fréchet 2002; Abbasi et al. 2014). End groups in dendrimers can easily be functionalized and facilitate modifications of physicochemical and biological properties. Dendrimers have emerged as a new class of nano-sized molecules with tremendous application in anticancer therapies and diagnostic imaging (Srinivasa-Gopalan and Yarema 2007; Stiriba et al. 2002). Peptide dendrimers are being widely researched for their applications in various fields, including as biomedical diagnostic agents and as delivery vehicles for vaccines, drugs, and genes (Sadler and Tam 2002).

6.2.2 *Micelles*

Conventional micelles are defined as a collection of amphiphilic surfactants that aggregate spontaneously in an aqueous solution to form a vesicle with a hydrophobic core. It can incorporate hydrophobic drugs in the inner core (Rangel-Yagui et al. 2005). Alternatively, polymeric micelles are formed by the spontaneous arrangement of amphiphilic co-polymers in an aqueous solution with a hydrophobic core and a hydrophilic shell. Polymeric micelles can incorporate a hydrophobic drug in the core and can also be coupled with the targeting ligands such as peptide antibodies on its shell for specific cell targeting and enhancing the cellular uptake of the incorporated drugs (Amin et al. 2017). These nanoparticles have been widely studied for their role as anticancer drug delivery system. Phospholipid micelles are widely used for peptide drug delivery. Glucagon-like peptide 1, glucose-dependent insulinotropic peptide, and neuropeptide Y are some of the peptides that are delivered using the micellar nanocarrier (Esparza et al. 2019). A chitosan-based micelle using N-octyl-N-arginine chitosan (OACS) was developed for insulin oral delivery (Zhang et al. 2013).

6.2.3 *Carbon Nanotubes*

Carbon nanotubes (CNTs) are cylindrical molecules composed of carbon atoms and can be described as graphene sheets rolled into a single or multiwall seamless cylinder. The diameters of CNTs vary from a few to hundreds of nanometers with a

length/diameter ratio of higher than 106 with exceptional thermal, mechanical, optical, and electrical properties (Roldo 2016). The single-walled CNTs have gained wide popularity as a drug delivery system due to their high cargo loading capacity, intrinsic stability, prolonged circulation time, and enhancement of bioavailability of the incorporated drug. Single-walled CNT-based nanomaterials have been reported to be drug delivery vehicles for nucleic acids, proteins, and drug molecules. These nanotubes have also been functionalized by antibodies for enhancing the uptake and site-specific anticancer drug delivery (Mahajan et al. 2018). Along with the carbon nanotubes, peptide-based nanotubes are also gaining popularity. The use of synthetic polypeptides, short Fmoc-dipeptides, cyclic peptides of alternating D- and L-amino acids, and preassembled bundles of α -helices forming peptide-based nanotubes has been reported (Burgess et al. 2015; Ghadiri et al. 1993; Rho et al. 2019; Hartgerink et al. 1996).

6.2.4 *Quantum Dots*

Quantum dots (QD) are inorganic fluorescent semiconductor nanoparticles with very unique optical and electronic properties. They exhibit high photostability, size-dependent optical properties, high brightness, and large Stokes shift, making them a better choice over organic dyes. Quantum dots consist of an ultra-small core with a size ranging from 1.5 to 10 nm of a semiconductor material (e.g., cadmium selenide (CdSe)) that is surrounded by another layer of semiconductor usually made of zinc sulfide (ZnS). The inner core and semiconductor layer are encapsulated by a cap from the outside that is made of different materials (Lombardo et al. 2019). QDs are used as a fluorescent agent for disease diagnosis and in various cellular and in vivo assays (Maxwell et al. 2020). Quantum dots have several features such as small size versatile surface chemistry with unique optoelectrical properties, which make them an excellent agent for real-time monitoring and tracking of nanoparticles in an in vivo model without significant alteration of nanocarrier (Probst et al. 2013). Recently, QDs conjugated with peptides have been developed for the enhancement of activity, site-specific action, and drug delivery. Peptide nanofibers with graphene quantum dots are evaluated for both targeting and imaging of tumor cells (Su et al. 2015). A peptide-carbon QDs conjugate derived from human retinoic acid receptor responder protein 2 has been used against both antibiotic-resistant gram-positive and gram-negative pathogenic bacteria (Mazumdar et al. 2020).

6.2.5 *Liposomes*

Liposomes are widely studied and one of the most well-characterized nanoparticle-based systems. A hydrophobic molecule can be incorporated into the lipid bilayer, whereas a hydrophilic drug can be entrapped into the core of the liposome.

Liposomes are used as nanocarriers of drugs for different diseases such as cancer, hepatitis A, and fungal and bacterial infections (Beltrán-Gracia et al. 2019). Surface-functionalized liposomes are used for site-specific delivery of drugs. Peptides, proteins, antibodies, and carbohydrates molecules can be coupled on the surface of liposomes for targeting designated cell types which overexpress specific types of receptors. Recently approved vaccines for COVID-19 from Pfizer-BioNTech, Moderna uses liposome technology for the delivery of RNA-based vaccines (Polack et al. 2020; Dagan et al. 2021). In this chapter, we will focus on peptide-based liposomes.

6.3 Liposome

Liposomes as a delivery vehicle for peptide and protein drugs and proteins are extensively studied. Liposome surface can be easily functionalized by peptides for targeting specific cell types, or they can be used as cargo for hydrophilic and hydrophobic peptides for specific site delivery. Peptide-based liposomes have the ability to lower the off-target effects, enhancing the stability of peptides and having better cells and tissue permeability.

Liposomes are spherical nanoparticles with an aqueous core and a lipid bilayer. These are formed naturally when lipids are stirred into an aqueous media, resulting in a population of vesicles with diameters from nanometers to micrometers. The water molecules reject the hydrocarbon tails, which point in the same direction; however, the lipid head groups are drawn to water molecules and organize themselves in such a way that they point into the aqueous compartment (Fig. 6.1) (Lopes 2013).

The inner layer's head groups point in the direction of the intravesicular fluid, while the tails point away. As a result, one layer's hydrocarbon tails point toward the outer layer's hydrocarbon tails, creating a natural bilipid membrane (Raffa et al. 2010). Liposomes can contain drugs with a wide variety of lipophilicities in the

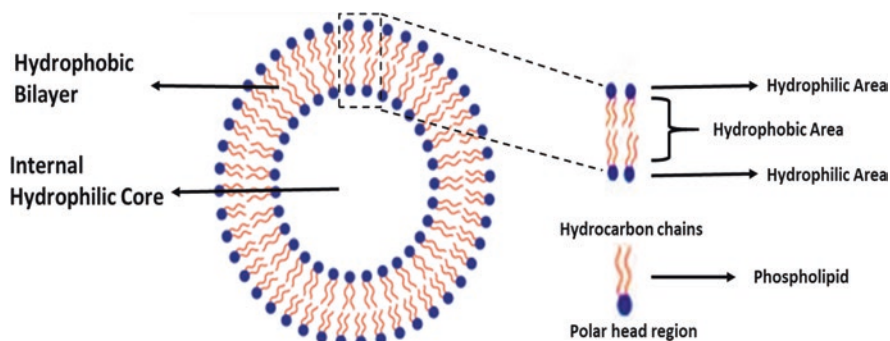


Fig. 6.1 Basic structure and composition of liposomes

lipid bilayer, the enclosed aqueous volume, or the bilayer interface (Huwlyer et al. 2008). Liposomes are usually prepared from natural or synthetic lipids, and the ingredients of liposomes are not limited to lipids and can also be formed from polymers (Meerovich and Dash 2019). Liposomes are biocompatible, biodegradable nanostructures made out of natural or synthetic lipids or polymers that can be used in biomedical research. One of the remarkable properties of liposomes is the ability to compartmentalize and dissolve both hydrophilic and hydrophobic molecules. Because of these properties, liposomes find applications in drug delivery (Çağdaş et al. 2014).

6.4 Formulation and Manufacturing Strategies of Liposome

Liposomes are made up mostly of phospholipids. These biomolecules are also key components in the construction of biological membranes. They have a polar head (water-soluble hydroxy groups) and an insoluble base, making them amphiphilic molecules. Liposomes may be zwitterionic, charged positively or negatively, or uncharged. The polar head charge is fully responsible for this. Liposomes are usually prepared for two kinds of lipids: naturally occurring or synthetic lipids (consisting of a phosphorus polar head and a glycerol backbone) and sterols (e.g., cholesterol) (Chowdhury 2008).

Liposomes also contain cholesterol, which is an important ingredient. It affects the characteristics of the lipid bilayer in liposomes in a modulatory way. It can modulate the stoutness of the liposome structure and enhance the packing between phospholipid molecules (Briuglia et al. 2015), resulting in a more ordered conformation and lower micropolarity in the aliphatic tail region (Liu et al. 2017); neighboring molecules (particularly water-soluble compounds) have less bilayer flexibility (Tarun and Goyal Amit 2014) and increased bilayer microviscosity (Olusanya et al. 2018). Cholesterol is also needed for liposomal membrane structural stability, which helps to keep the liposome stable in intestinal environmental stress (Liu et al. 2017).

Liposomes can be functionalized with varieties of biomolecules or small molecules (PEGs, aptamers, antibodies, proteins, peptides, ligands, sugars, or small molecules) for targeting effects (Fig. 6.2). To render special targeting properties, surface functionalization of liposomes can be used (Riaz et al. 2018). The preparation process has a significant impact on the characteristics of the processed liposomes. While liposome formation may occur spontaneously, mechanical agitation is often necessary. Various preparation procedures have been created in order to have control over the size and form of the liposomes generated, increase the efficacy of trapping the target molecules, and avoid the eventual leaking of encapsulated compounds from liposomes (Çağdaş et al. 2014).

Before one selects the type of liposome and components required, there are a few factors to consider: (1) liposomal components and the physicochemical properties of the material to be entrapped, (2) the ideal concentration of the encapsulated item

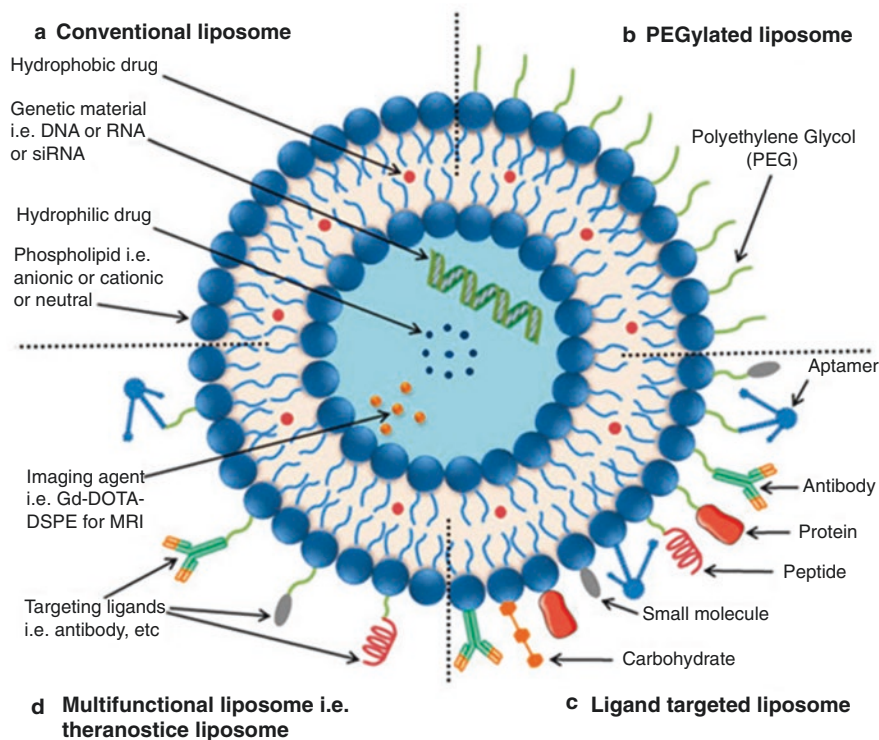


Fig. 6.2 Liposomes: Conventional and functionalized liposomes: (a) Phospholipid-based liposomes, (b) PEGylated/stealth liposomes with a layer of polyethylene glycol (PEG), (c) targeted liposomes with a specific ligand to target a disease site, and (d) multifunctional liposomes that can be used for diagnosis and treatment. (Reproduced from Creative Commons Attribution License) (Riaz et al. 2018)

and its possible toxicity, (3) the features of the media in which the liposomes are suspended (4) extra processes engaged during application (liposome transport), and (5) the optimal size, polydispersity, and shelf-life of the liposomes, (6) the possibility processing and the reproducibility of effective and efficient liposomal products across batches (Gomez-Hens and Fernandez-Romero 2006; Dua et al. 2012). Liposome size is an important factor in achieving effective drug accumulation at the target location and preserving liposome circulation half-life for in vivo drug delivery. The size and quantity of bilayers in the produced liposome are also related to the volume of encapsulated medication. Depending on the goal of the formulation, several liposome preparation methods may be utilized. Lipid hydration and the replacement of organic solvents with an aqueous medium are the most extensively utilized liposome production procedures (reverse-phase evaporation and organic solvent injection). According to Bangham's procedure, lipids are dissolved in a suitable organic solvent such as chloroform or methanol. The solvent is then evaporated

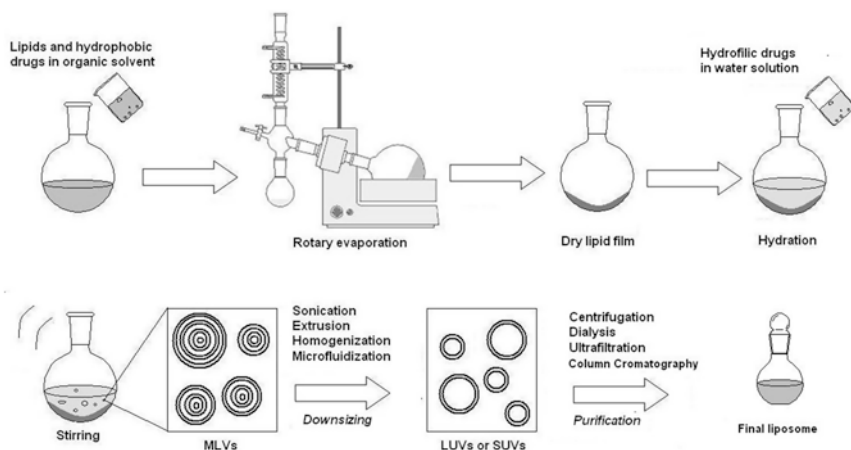


Fig. 6.3 Lipid hydration accompanied by vortex or manual stirring represents liposome production. (Reproduced from Creative Commons Attribution License) (Lopes 2013)

using rotary evaporation under reduced pressure before a thin layer is produced. The thin film is then hydrated in an aqueous solution at a temperature above phase transition to producing multi-lamellar vesicle (MLV) liposomes (Fig. 6.3). Lipophilic drugs are typically incorporated by co-dissolution with lipids (Wagner and Vorauer-Uhl 2011). Drugs that are hydrophilic dissolve in an aqueous medium or buffers. Amphiphilic drugs can be dissolved in both mediums. The creation of large vesicles (MLV) with a heterogeneous size distribution will occur during the liposome preparation process; and large liposomes can be made into small unilamellar vesicles by using a vesicle size reduction process.

6.5 Characterization of Liposomes

Characterization of liposomes involves various attributes such as encapsulation of drug, nanoparticle morphology, shape, size, surface charge (zeta potential), physical and chemical stability, and release of encapsulated drugs by *in vitro* studies (Fig. 6.4). Dialysis, ultra-centrifugation, ultrafiltration, and solid-phase extraction aid in the removal of the unencapsulated drug. Then, the encapsulated drug can be quantified using various methods such as fluorescent-based spectroscopy, RP-HPLC, capillary electrophoresis (CE), and field-flow fractionation (FFF).

The morphology of liposomes is determined by transmission electron microscopy (TEM), scanning electron microscopy (SEM), and cryo-TEM. TEM is the most frequently used microscopy technique to study the morphology of nanoparticles (Henry 2005; Kuntsche et al. 2011). These techniques are used to study the spherical shape of liposomes as well as detailed structural information of lipid layers (Tonggu and Wang 2020). Microscopy techniques are great tools to determine

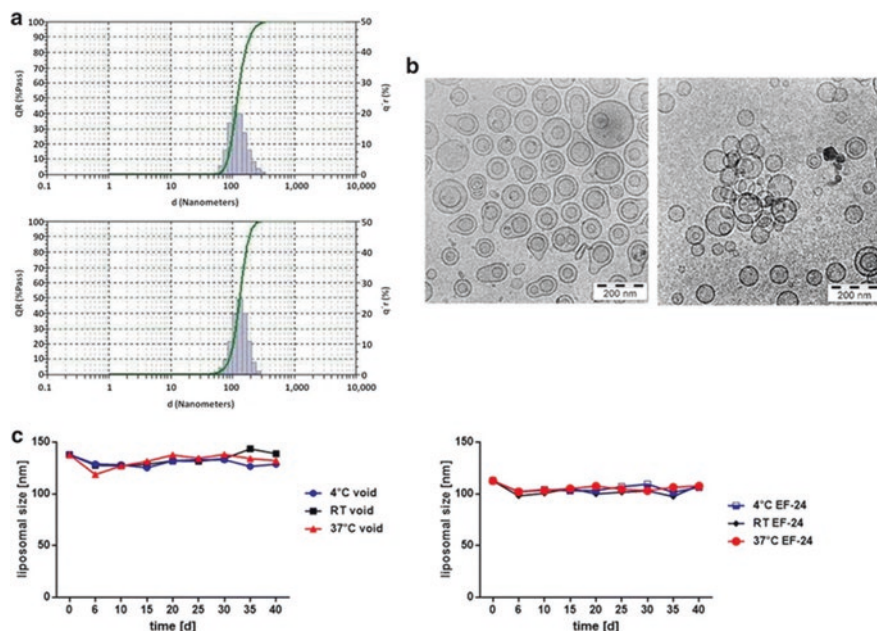


Fig. 6.4 Synthesis and characterization of void and EF24-containing PEGylated liposomes. Pegylated liposomes synthesized using a lipid hydration method were further characterized using DLS and TEM. (a) DLS of void and EF24-loaded liposomes revealed a narrow size distribution with an average diameter of less than 150 nm, (b) transmission electron microscopy of void (left panel) and EF24-containing liposomes (right panel) demonstrated spherical morphology and an average diameter of around 120 nm, in line with the data obtained by DLS, (c) the stabilities of void and EF24-loaded liposomes were determined at three different temperatures (4, 20, and 37 °C) using DLS over 40 days. (Reproduced from Creative Common Attribution License) (Bisht et al. 2016)

particle morphology but provide very little information about particle size and its distribution (Robson et al. 2018; Fan et al. 2021).

The particle size of liposomes is usually determined by using dynamic light scattering (DLS). DLS aid in the characterization of liposomes by providing information about mean particle size, particle size distribution, zeta potential, and polydispersity index. The particle size of the liposome is an important parameter as it has to be optimized depending on the targeted site and delivery method. Liposomes for antitumor drug delivery are usually in size range of ≤ 100 nm for better drug permeation into the tumor micro-environment and local tumor tissues (Nagayasu et al. 1999; Danaei et al. 2018). The optimum size of liposomes depends on the targeted tumor size, stage, and location. The large particle size of liposomes may not show the intended therapeutic effect because of poor permeability and phagocytosis by immune system cells. The particle size of liposomes can be controlled by using the techniques such as, sonication, homogenization, and extrusion (Mozafari

2005a; Mozafari 2005b). These techniques ensure the liposomes with small particle size (<200nm) and uniform size distribution.

Another important attribute of liposomes is the surface charge. Surface charge of the liposomes is determined by the phospholipid head groups that can incorporate positive or negative charges. The surface potential of nanoparticles provides information about the intraparticle interactions strength, adsorption of counterions, and particle stability (Fan et al. 2021). Surface potential, also termed as zeta potential, should be optimum to maintain the repulsion between the particles and uniform suspension of particles. Usually zeta potential of < -30 mV or > 30 mV is considered optimum for preventing the aggregation of particles in suspension (Samimi et al. 2019). Liposomes should be able to retain the drug during storage and in vivo circulation before delivering the drug to the desired site (Shen and Burgess 2013; Wang et al. 2014).

In vitro drug release from the liposomes can be determined by employing different methods such as sampling and separation (SS), dialysis membrane (DM), continuous flow (CF), or combined approaches (Fan et al. 2021). Among them, the dialysis membrane method is widely used for the determination of the drug release from the liposome. DM approach employs a dialysis sac in which the liposomes are kept, and this dialysis sac is immersed in the buffer. The dialysis membrane with an appropriate molecular weight cut-off (MWCO) should be selected based on the entrapped drug molecular weight. The amount of drug in the buffer outside of the dialysis sac is quantified at different time intervals to obtain the in vitro release of the drug from liposomes (Fan et al. 2021).

6.6 Liposome Stabilization Strategy with Lyophilization

Liposome stability is a major concern when developing them for pharmaceutical applications. Liposomes' increased bilayer permeability can lead to drug leakage, vesicle aggregation/fusion, and precipitation (Du Plessis et al. 1996). The most frequent strategy to improve liposome stability is to design an appropriate formulation, which entails selecting the right lipid composition and concentration as well as incorporating other substances to lengthen shelf life. Cholesterol and its derivatives, for example, can decrease lipid bilayer permeability. Antioxidants and metal chelators can be added to enhance the stability since unsaturated lipids are prone to peroxidation (Abdelwahed et al. 2006a, b; Chen et al. 2010). Furthermore, the presence of oxygen, both in the form of dissolved oxygen and in the container's headspace, must be avoided. When liposomes are hydrolyzed in aqueous dispersion, lysophospholipids and fatty acids are produced. This reaction is catalyzed by hydroxyl and hydrogen ions, and it can be slowed down by adjusting the pH or by adding a neutral buffer (Zhang and Pawelchak 2000). Liposome stabilization can be achieved using a variety of techniques, including lyophilization, and spray-drying, as well as formulation optimization. Lyophilization is the most popular procedure for increasing

the shelf life of liposomes, especially for thermosensitive drugs contained inside them (Chen et al. 2010).

Lyophilization, also known as freeze-drying, is a sophisticated drying method used to transform labile material solutions into solids that are stable enough for distribution and storage. Lyophilization is an industrial procedure that involves sublimating and desorbing the water from a frozen sample using a vacuum system. Nonetheless, when performed without the required stabilizers, this process causes a wide range of stress, including fusion and drug degradation, during the freezing and drying steps (Abdelwahed et al. 2006a, b; Abdelwahed et al. 2006a). Cryoprotectants, such as saccharides and their derivatives (e.g., sucrose, trehalose, hydroxypropyl—cyclodextrin (HP—CD), are used to keep the vesicles stable throughout freeze-drying (Bendas et al. 1996; Mohammed et al. 2006). Sugars stabilize membranes by replacing water, which is the most well-known and widely accepted mechanism. The protective effect is thought to be based on complicated and particular interactions between phospholipids and sugars. Experiments have shown that interactions take place by a hydrogen bond between the sugars' hydroxyl groups and the phosphate groups on the bilayer surface. In summary, sugars limit water-phospholipid interactions (Chen et al. 2010; Santivarangkna et al. 2008). Among different sugars, trehalose has been detected to enter the membrane and establish hydrogen bonds with the carbonyl groups of phospholipids (Diaz et al. 1999; Luzardo et al. 2000; Villarreal et al. 2004). As a result, trehalose seems to have a greater affinity for phospholipid bonding and is considered an excellent cryoprotectant (freeze-protectant) for liposomes. When liposomes were freeze-dried in the presence of adequate amounts of trehalose (a carbohydrate found in high concentrations in organisms), it was observed that they preserved up to 100% of their original contents. Proper controlled lyophilization of liposomes can be achieved using freeze-driers in various sizes, from portable laboratory versions to large commercial devices (Akbarzadeh et al. 2013; Awada et al. 2004).

6.7 Liposomal Nanocarrier System for Peptide Drug Delivery

Peptide drugs can be incorporated into the liposome in various ways, and it depends on the function of the peptide as well as the characteristics of the transported peptide drug. Hydrophobic peptides are entrapped into the liposomal bilayer and typically dissolved into the organic solvent before forming a thin lipid layer in the thin layer hydration method. Hydrophilic peptides can be entrapped into the hydrophilic compartment of the liposome. Targeting peptides are attached to the outer layer to facilitate the selective delivery of the liposomal carrier. One such example is the study reported on peptide S1 that could specifically bind to the vascular epithelial growth factor receptor 2 (VEGFR2) used to target liposomes to the cells that over-express VEGFR2. S1 peptide incorporated liposomal formulation was found to be an efficient nanoscale drug delivery device *in vitro* and *in vivo* (Han et al. 2016).

The peptide entrapped in the liposome can be measured through the process called the entrapment efficiency method. In this process, the liposome solution can be dissolved in a suitable organic solvent to extract the entrapped peptide and then analyzed with suitable detection methods and compared with the loading solution to get the percentage of the entrapped peptide content into the liposomal compartment. Entrapment efficiency can be measured using the following formula:

$$\text{Entrapment efficiency} = \frac{\text{Drug content in the liposome formulation}}{\text{Drug content in the loading solution}} \times 100\%$$

The release profile of the liposome is an important characteristic of the liposomal formulation that can be assessed through different release study methods, and the dialysis process can be a reliable way for this process. Specific molecular cutout dialysis bags can be used, and buffer media can be utilized for the assessment of the release property of the liposome solution. Rezaei et al. performed the dialysis method to investigate the *in vitro* release profile of the peptide from encapsulated. The cumulative percentage release of the liposomal samples is given as a summary of their release behavior. Under the test conditions, the peptide's release rate was gradual and desirable (Rezaei et al. 2020).

Liposomal drug release can be increased or decreased with the utilization of different phospholipid properties and concentrations through maintaining peptide to lipid ratio (Li et al. 2015). To test the pharmacological availability of liposomes with phosphatidylethanol, Kisel et al. developed three types of liposomal formulations with insulin: (i) dipalmitoyl phosphatidylcholine/dipalmitoyl phosphatidylethanol (1:1 w/w) liposomes, (ii) dipalmitoyl phosphatidylcholine/dipalmitoyl phosphatidylethanol/palmitoyl–stearoyl sucrose (1:1:0.2) liposomes, and (iii) liposomes composed of natural phosphatidylcholine and phosphatidylinositol (1:1). Hyperinsulinemia was seen after oral administration of all liposomal species in animal model studies. Hyperinsulinemia was accompanied by a drop in blood glucose content when liposomes containing dipalmitoyl phosphatidylethanol were used (Kisel et al. 2001). In male Wistar rats, Katayama et al. found that with intragastric delivery of positively charged double liposomes (DL) prepared with stearylamine (SA) and phosphatidylserine (PS) using the glass filter method as insulin carriers in combination with aprotinin, a protease inhibitor resulted in notable hypoglycemic effects (Katayama et al. 2003). Aprotinin was conjugated with chitosan to create a polymer–protease inhibitor with a positive charge. Liposomes were prepared with 1- α -distearoylphosphatidylcholine (DSPC), dicetylphosphate (DCP), and cholesterol (molar ratio: 8:2:1) to acquire negative charges. A formulation was prepared with polyelectrolyte complexes between negatively charged multilamellar vesicles (MLV) and positively charged chitosan–aprotinin conjugate to improve systemic uptake of therapeutic peptides after oral administration. *In vitro*, it was shown that chitosan–aprotinin inhibited trypsin substantially at doses of 0.05% and 0.1%, but no inhibition was shown in the presence of 0.1% chitosan (Werle and Takeuchi 2009). In another study, a new form of liposome containing tetraether lipids (TELs) produced from archaea bacteria may enhance oral peptide delivery. All liposomal

formulations were prepared by the film method with a mixture of host lipids, e.g., egg phosphatidylcholine (EPC) or dipalmitoyl phosphatidylcholine (DPPC), tetraetherlipids (TEL), and other lipids, e.g., cholesterol. The film technique was used to make liposomes, which were then extruded. Photon correlation spectroscopy revealed the presence of vesicles with sizes ranging from 130 to 207 nm (Parmentier et al. 2011). Niu et al. showed that insulin oral bioavailability was considerably improved when liposomes containing bile salts (BS-liposomes) were used for the formulation of recombinant human insulin (rhINS). By using the reversed-phase evaporation method, BS-liposomes containing sodium glycocholate (SGC), sodium taurocholate (STC), or sodium deoxycholate (SDC) were prepared. These bile salt containing liposomes exhibited increased residence duration and penetration through biomembranes and increased absorption in a model system. Based on their studies Niu et al. conclude that enhanced absorption of insulin-loaded liposomes may be due to mechanisms of trans-enterocytic internalization of liposomes. (Niu et al. 2014). In another study, liposomal vasoactive intestinal peptide (VIP) inhalation therapy was used as a potential therapeutic option for severe lung disorders. It was observed that in the lungs, VIP has a relatively brief time of action. To enhance the activity and duration of action of VIP in the lungs, Stark et al. created a liposomal drug delivery method for VIP and evaluated it for its ability to protect VIP against enzymatic cleavage. The liposomal formulation components were polyethyleneglycol conjugated distearyl-phosphatidylethanolamine (DSPE-PEG2000)-lyso-stearyl-phosphatidylglycerol (lyso-PG-palmitoyl-oleoyl-phosphatidylcholine (POPC) (1,7.5:11). When these formulations were evaluated in a model system, it was found that the free VIP was quickly digested, but liposomal-associated VIP showed relatively enhanced stability (Stark et al. 2008) (Table 6.1).

Liposomes as a nanocarrier are being used for peptide drug delivery in various ways. Oral administration of liposomes with peptide cargos was evaluated using various techniques. Muramatsu et al. found that soybean sterol containing insulin liposomes were able to lower blood glucose levels in rats for up to 21 h after injection (Muramatsu et al. 1996). Liposomes with peptide loads for buccal distribution have been the subject of research. In normotensive/spontaneously hypertensive hamsters, Suzuki et al. found significant vasodilation after administration of free vasoactive intestinal peptide (VIP) and significantly prolonged vasodilation with VIP liposomes, indicating that VIP-induced vasodilation is impaired in situ in essential hypertension and is restored by encapsulation into liposomes (Suzuki et al. 1996; Sejourne et al. 1997). Liposomes with peptide cargos for pulmonary administration also had been investigated. In the rat lung, Huang et al. found that pulmonary distribution of insulin-loaded liposomes resulted in prolonged effective hypoglycemia, which was not seen with a mix of free insulin and empty liposomes (Huang and Wang 2006; Bi et al. 2008). Intravenous administration of liposomal formulation containing peptides was also investigated. Using iodine-labeled VIP liposomes, Refai et al. found that VIP liposomes were better absorbed by the rat lung than free VIP (Refai et al. 1999). The inhalation of vasoactive intestinal peptide (VIP) has been proposed as a viable therapeutic option for a variety of lung disorders, including asthma and pulmonary hypertension. Due to fast enzymatic

Table 6.1 Peptide liposomes composition, preparation, and characteristics

Name	Composition	Preparation method	Particle size	Indication	Ref.
Liposome-entrapped insulin	Liposomes were prepared by mixing phosphatidylcholine and phosphatidylinositol ('fluid' 1:1 w:w) or dipalmitoyl phosphatidylcholine and dipalmitoyl phosphatidylethanol ('solid', 1:1 w:w) or the latter supplemented with palmitoyl–stearoyl sucrose (1:1:0.2 w:w:w) in chloroform	Thin-film hydration method	50–250 nm	Hypoglycemic effect	Kisel et al. (2001)
Double liposomes (DLs) containing insulin	26 mM H-soyaPC was dissolved alone or with 2.6 mM SA or PS as lipids with electrical charges in chloroform	Glass filter method	2–9 μ m	Hypoglycemic effect	Katayama et al. (2003)
Oral peptide delivery with chitosan–aprotinin-coated liposome	Anionic multilamellar liposomes (MLV) consisting of DSPC, DCP, and Chol (molar ratio: 8:2:1)	Thin-film hydration method	3 and 4.5 μ m	Novel polymer protease inhibitor-based delivery systems	Werle and Takeuchi (2009)
Octreotide peptide oral delivery with tetraether lipids (TELs)	A mixture of host-lipids, e.g., EPC or DPPC, TEL, and other lipids, e.g., cholesterol was dissolved in chloroform/methanol (8:1)	Thin-film hydration method	130–207 nm	Acromegaly, Gastrointestinal disorders, and psoriasis	Parmentier et al. (2011)
Recombinant human insulin (rhINS)-loaded BS-liposome	Soybean phosphatidylcholine (SPC) and bile salts (SGC/STC/SDC) were dissolved in 5 mL absolute ether with a molar ratio of 4:1, into which 1 mL rhINS solution in citric–Na ₂ HPO ₄ buffer (4 mg/mL, pH = 3.0) was added	Reversed-phase evaporation method	150 nm	Hypoglycemic effect	Niu et al. (2014)
Inhalation administration of vasoactive intestinal peptide (VIP)	Stock solutions of DSPE PEG2000, lyso-PG, and POPC in pure chloroform or chloroform-methanol mixtures, respectively, were made	Thin-film rehydration method	95 nm	Cystic fibrosis, ulcerative colitis, and primary pulmonary hypertension	Stark et al. (2008)

breakdown in the airways, peptides have a short half-life, which limits their medicinal utility. As a result, Hajos et al. created unilamellar nano-sized VIP-loaded liposomes (VLL) and found that by creating a “dispersible peptide depot” in the bronchi, the liposomes have the potential to enhance VIP inhalation treatment. As a result, exposure to cells, i.e., direct ligand-receptor interactions, might stimulate the release of VIP from liposomes. A schematic diagram of the model proposed for the stability of VIP against enzymatic degradation is shown in Fig. 6.5 (Hajos et al. 2008).

Peptides encapsulated in liposomes are shielded from the inactivating effects of environmental factors, thus causing no adverse side effects. Liposomes provide a unique method of delivering peptides into cells or even individual cellular compartments. Liposome size, charge, and surface qualities can easily be changed by

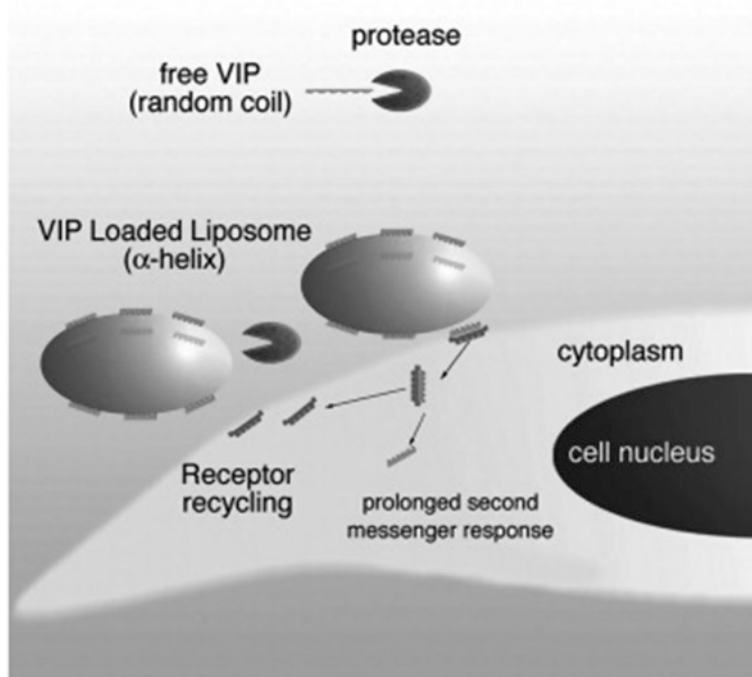


Fig. 6.5 Suggested fusion model of VLL Free VIP (random coiled) can easily be degraded by proteases on the way to the receptor. VIP from VLL is protected against proteases. VLL show no peptide leakage during storage, but electrostatically bound VIP may become released by direct contact and binding to the receptor. Following the ligand-receptor-complex internalization, and intracytoplasmic complex disintegration, the receptor is recycled to the cell surface to bind new upcoming VIP molecules. In addition to the protection by liposomes per se, the alpha helical conformation of VIP induced by negatively charged liposomes may convey further degradation protection; moreover, it is preferred VIP-conformation for receptor binding. Reprinted from Publication Inhalable liposomal formulation for vasoactive intestinal peptide Hajos et al. Vol357 p286-294 (2008) International Journal of Pharmaceutics, with permission from Elsevier (Hajos et al. 2008)

simply adding new chemicals to the lipid mixture before liposome formation and/or using different preparation procedures (Adibzadeh et al. 1992). The characteristics of the carrier are then exploited to modulate drug distribution rather than the physicochemical properties of the medicinal ingredient (Çağdaş et al. 2014). Liposomes have different advantages in peptide drug delivery and peptide-based targeted therapy leading to enhanced efficacy with reduced toxicity (Swaminathan and Ehrhardt 2012). They play a role in improving drug solubility (Mohammed et al. 2004), working as a sustained release system (Mukherjee et al. 2007), offering targeted drug delivery (Sonju et al. 2021), reducing drug toxicity (Naik et al. 2021), providing support against drug degradation (Cristiano et al. 2017), enhancing API circulation half-life (Allen et al. 2006), overcoming multidrug resistance (Matsuo et al. 2001), improving the therapeutic index of the entrapped drug (Wang et al. 2010), and shielding APIs against their neighboring environment (Park et al. 2011a).

Positive surface charge insulin-loaded liposomes display greater resistance to trypsin digestion than negative and neutral surface charged insulin-loaded liposomes, according to Kato et al. (Kato et al. 1993). Corona-Ortega et al. found that cationic/positively charged liposomes had better cell adhesion than neutral/negative IL-2-encapsulated liposomes (Corona-Ortega et al. 2009). Law et al. conducted a thorough investigation into formulation production parameters, with the liposomal charge being a key parameter under investigation in salmon calcitonin-loaded liposomes, where a negative surface charge allowed greater cargo encapsulation over neutral and positive surface charges (Law and Shih 2001). Some payloads exhibit interaction with lipid membrane because they are either lipophilic or amphiphilic. Peptides belong to the latter group and studies demonstrating their interaction with a lipid membrane (Stark et al. 2007; Neville et al. 2000; Joffret et al. 1990; Arien et al. 1995). Following cholate-induced disruption of sCT liposomes, Arien et al. tried to explain why the oral administration of calcitonin (CT)-loaded liposomes resulted in a hypocalcemia in animal model. Based on their studies the authors conclude that CT-lipid complex formed protects the peptide from trypsin digestion (Arien et al. 1995). VIP interaction with lipid bilayer was reported by Stark et al. (Stark et al. 2007), and IL-2 interaction with lipid bilayer was reported by Neville et al. (Neville et al. 2000), all of which were verified by freeze-fracture microscopy. Joffret et al. hypothesized that IL-2 contact with liposomal bilayers resulted in increased proliferation of cytotoxic T cells after administration of IL-2 liposomes (Joffret et al. 1990). The drug loading process determines the payload's encapsulation efficiency and is thus a relevant parameter. The association of insulin with the liposomal membrane is promoted in the presence of a transmembrane gradient, according to Hwang et al., and this is not optimal for insulin loading (Hwang et al. 1999). In contrast to the transmembrane gradient process, the reversed step evaporation method resulted in a twofold rise in insulin encapsulation into liposomes (Hwang et al. 2000). In vitro studies have shown that including permeation enhancers into liposomal formulations has a positive influence. Degim et al. and Maitani et al. demonstrated increased insulin permeability through Caco-2 cell monolayers and rabbit nasal mucosa linked to diffusion cells prepared with Caco-2 cell monolayers (Degim et al. 2004; Maitani et al. 1992). In the presence of sodium

taurodeoxycholate, Song et al. demonstrated improved sCT permeation through Caco-2 cell monolayers (Song et al. 2005). PEGylation has many benefits, including the avoidance of opsonization (in vivo) and, in several situations (formulation related), increased payload encapsulation (Immordino et al. 2006). Park et al. found that PEGylated liposomes encapsulated insulin more effectively than non-PEG counterparts (Park et al. 2011a; b). According to Kedar et al., IL-2-encapsulated PEG liposomes interacted better with cells in vitro than their non-PEG equivalent (Kedar et al. 2000).

6.8 Types of Liposome

6.8.1 Active Targeting Liposomes

The main limitation of conventional liposomes is the off-target effects. This problem can be addressed by the development of actively targeted liposomes. Liposomes serve as an attractive, active targeting drug delivery system as their surface can be modified by various ligands such as small molecules, aptamers, antibodies, and peptides (Byrne et al. 2008). These ligands can specifically target the cancer cell, which overexpresses the corresponding receptors. Tumor cells are found to overexpress receptors like EGFR, HER-2, transferrin receptor, folate receptors, integrins, and so on, depending on tumor types. Targeting these overexpressed receptors by conjugating ligands on the liposomal surface can result in tumor-specific targeting and drug delivery (Torchilin 2007). Actively targeting liposomes can be developed to target tumor cell surface receptors like EGFR, HER-2, or tumor microenvironment/vasculature like VEGF, matrix metalloproteinases, and $\alpha\beta$ -integrins (Deshpande et al. 2013). Peptides serve as an excellent targeting agent because of their specificity, easy synthesis, low costs, lower immunogenicity, and ease of conjugation on the liposomal surface. Additionally, antitumor peptides as ligands on liposomal surfaces incorporating cytotoxicity can have better tumor accumulation and better cytotoxicity towards tumor cells. Use of cell-penetrating peptides in liposomes aid in better penetration of the drug into tumor cells (Ye et al. 2016).

6.8.2 Stimuli-Responsive Liposomes

Conventional liposomes have some limitations in terms of drug release, as the drug may get released before it reaches to target site and may not accumulate at the target site. This challenge is addressed by the development of stimuli-responsive liposomes (An and Gui 2017). The stimuli may be presented to liposomes by the target site microenvironment (e.g., pH, redox potential, enzyme) or applied externally from the outside (e.g., hyperthermia, ultrasound, magnetism). Various

stimuli-responsive liposomes such as pH-sensitive, thermosensitive, magnetic field-sensitive, and ultrasound-sensitive liposomes are developed (Torchilin 2009; An et al. 2010; Amstad et al. 2011). Stimuli-responsive liposomes, after reaching the target site, undergo changes in the composition or structure of the bilayer in the presence of the stimuli leading to the release of drug at the target site. Thus reducing the premature release of the drug and increasing the site-specific targeting with reduced off-target effects of liposomes (Andresen et al. 2005).

6.8.3 Temperature-Sensitive Liposomes (Thermosensitive Liposomes)

Thermosensitive lipids and polymers with a low critical solution temperature (LCST) are used to prepare the thermosensitive liposomes. These liposomes can be used in the site-specific delivery to the tumor sites where hyperthermia is observed compared to normal tissues (Danhier et al. 2010). An increase in temperature changes disrupts the lipid bilayer structure of liposomes, releasing the drug in the tumor sites (Chountoulesi et al. 2017). Heat can also be applied externally on tumor sites which trigger the release of drugs from thermosensitive liposomes.

6.8.4 pH-Sensitive Liposomes

The tumor microenvironment has lower pH compared to the normal cell environment due to the high rate of glycolysis in cancer cells generating excessive metabolites like lactic acid and CO₂ (Cardone et al. 2005). This lower pH of the tumor microenvironment can be exploited for effective delivery and targeting of the targets by the development of pH-sensitive liposomes. pH-sensitive liposomes are composed of pH-sensitive lipids that show structural integrity in normal pH but destabilize the lipid bilayer as soon as it is exposed to lower pH around 6–6.5 (Felber et al. 2012; Cardone et al. 2005). This facilitates the drug release more in tumor cells from pH-sensitive liposomes. A pH-sensitive liposome is composed of pH-sensitive lipids like DOPE (1,2-dioleoyl-sn-glycero-3-phosphoethanolamine) and a weakly acidic amphiphile, such as cholesteryl hemisuccinate (CHEMS) with polymorphic phase behavior, and adopts the hexagonal state at lower pH compromising the lipid bilayer integrity and facilitates the drug delivery (Soares et al. 2011). Dual targeting liposomes by the use of a pH-sensitive approach are studied for their target-specific effect and delivery. The use of pH-dependent peptides, cell-penetrating peptides, and antimicrobial peptides coupled to the pH-sensitive lipids/polymers to trigger the drug release from liposomes are also being developed for targeting tumor cells (Zhao et al. 2016; Ding et al. 2017; Zhang et al. 2016).

6.8.5 Magnetic Field-Responsive Liposomes

Magento-liposomes (MLs) incorporate magnetic and metallic nanoparticles such as Fe_3O_4 and can serve as an excellent agent for diagnostic and therapeutic purposes. MLs are responsive to the externally applied magnetic field and help to maintain the liposomes at the target site for complete release of the drug and enhance the site-specific delivery of the drug. MLs which incorporate metallic nanoparticles can play a dual role as a hyperthermia agent or as a nanocarrier for drug delivery or a combination therapy for cancer treatment (Anilkumar et al. 2019). Magnetic nanoparticle-loaded thermosensitive liposomes with near-infrared (NIR) laser-triggered release of doxorubicin in tumor cells showed that MLs can be used as combined photothermal–chemotherapy of tumors (Shen et al. 2019).

6.8.6 Ultrasound-Responsive Liposomes

Ultrasound-responsive liposomes encapsulate the drug along with a small amount of air which makes them acoustically active on ultrasound stimulation. These liposomes are suitable for ultrasound imaging and for controlled localized drug delivery on ultrasound stimulation in different conditions like cancer, thrombus, arterial stenosis, myocardial infarction, and angiogenesis (Huang 2010). These echogenic liposomes are prepared from different methods, including lyophilization and pressurization. The application of high-intensity-focused ultrasound (HIFU) induces local hyperthermia resulting in the release of drug contents from liposomes. Various studies of ultrasound-responsive liposomes, including cell-penetrating peptides and conjugated doxorubicin, showed enhanced cytotoxicity towards cancer cells upon ultrasound stimulation (Xie et al. 2016; Lin et al. 2016).

6.8.7 Light-Sensitive Liposomes

Light of a certain wavelength can trigger the release of drug contents from light-sensitive liposomes with a high level of control. These liposomes consist of photosensitizer (photoactive molecules) that generates singlet oxygen and other reactive oxygen species upon exposure to a particular wavelength of light resulting in disruption of lipid bilayer membrane and release of entrapped drug (Miranda and Lovell 2016; Prasmickaite et al. 2002). Photosensitizers like porphyrin derivatives, chlorins, phthalocyanines, and porphycenes are incorporated in the lipid bilayer or conjugated to the lipids for the light-sensitive effect (Enzian et al. 2020; Yavlovich et al. 2010). Various adjustable factors like wavelength, duration, light intensity, and photosensitizer concentrations make these photosensitive liposomes an attractive drug delivery system with a high level of temporal and spatial control of drug release

and easy optimization with a broad range of applications (Yavlovich et al. 2010). Various photosensitive liposomes incorporating drugs are studied for precise and control delivery to the target site by adjustable near-infrared (NIR) light irradiation (Sun et al. 2016). Photosensitive liposomes consisting of gold nanoparticles are gaining popularity as they can use NIR as a light source for deep tissue penetration and lower phototoxicity (Mathiyazhakan et al. 2018).

6.9 Limitation of Liposomes

6.9.1 *Reticuloendothelial System (RES) and Liposome Clearance*

The reticuloendothelial system is present in primary organs such as the liver, spleen, kidney, lungs, lymph nodes, and bone marrow. Maximum liposomal uptake is observed in the liver, followed by the spleen, which aids in the removal of liposomes from circulation. Macrophages present in the RES eliminate the liposomes. RES elimination of liposomes is addressed by the conjugation of PEG polymers to the lipid membrane. PEG helps prolong the circulation time of liposomes and prevents elimination by RES by steric stabilization of liposomes (Ishida et al. 2001).

6.9.2 *Accelerated Blood Clearance Phenomenon*

Repeated injection of PEGylated liposomes may result in loss of their long circulation time properties, which in turn leads to rapid blood clearance. This phenomenon is known as accelerated blood clearance (ABC) (Dams et al. 2000; Ishida et al. 2003). The exact mechanism of ABC of repeated dosing of PEGylated liposomes is still unclear, but this may be a great hurdle for clinical approval and application of PEGylated liposomes. This will be particularly challenging for surface-functionalized liposomes as most of these liposomes use PEG. ABC phenomenon is affected by PEG density, lipid dose, and dosing intervals (Ishida and Kiwada 2008; Sercombe et al. 2015). Such factors need to be carefully optimized to prevent this ABC phenomenon.

6.9.3 *High Serum Protein Binding*

Liposomes show a higher affinity to the serum protein that may result in the masking of ligands to the receptors. It may also affect the liposome biodistribution and pharmacokinetic properties lowering the targeting and drug delivery efficiency of the liposomes (Sercombe et al. 2015).

6.9.4 Masking of Surface Ligands by Polymers

Polymers like PEG are used for enhancing the circulation time of liposomes, but they might have a limitation in actively targeted liposomes. PEG and other polymers may mask the ligand by steric hindrance and may prevent the interaction of the ligand with the receptor. This will greatly affect the target-specific delivery of liposomes. This limitation can be overcome by careful selection of the length of PEG.

6.9.5 Difficulty in the Accurate Characterization of Surface-Functionalized Liposomes

There are well-defined methods for the characterization of conventional liposomes, but when it comes to surface-functionalized liposomes, there are limited guidelines and characterization methods. This makes it very difficult for accurate quantification of ligands and drugs in the functionalized liposomes. Proper technique to quantify the ligand attached is limited. Highly reliable biochemical and biophysical methods for quantification of the peptide and proteins attached on the surface of liposomes are still lacking. Hence batch-to-batch variation in production is very hard to address.

6.9.6 Large-Scale Production of Surface-Functionalized Liposomes

Production of conventional liposomes that are FDA approved has been well characterized, and large-scale production is relatively smooth due to their simple composition. However, the surface-functionalized liposomes with ligands such as peptides and antibodies face challenges in characterization, and hence industrial-scale production is an uphill task. These surface-engineered liposomes are studied on a small scale in a laboratory setting, and industrial scaling of these liposomes might have a very high batch-to-batch variation due to the difficulty in accurately quantifying. Thus, the physiochemical, pharmacokinetic, and biological properties of the peptide-functionalized liposomes may be compromised during industrial scaling (Wagner and Vorauer-Uhl 2011).

6.9.7 Stringent Storage Conditions

Liposomes require very strict storage conditions as high temperature leads to instability, and freezing them may result in ice crystal formation, which compromises the structural integrity of the formulation. Hence, liposomes need to be stored in a refrigerator at all times.

6.9.8 Aggregation of Liposomes

The high density of the ligands such as peptides, antibodies may lead to the aggregation of the liposomes resulting in loss of membrane integrity of liposomes. Proper optimization and study of the optimum ligand density, particle size and composition, and cell type to target may help in the prevention of aggregation (Barenholz 2001).

6.9.9 High Cost

Surface-functionalized liposomes require high costs for development. Unlike conventional liposomes, surface engineering of liposomes by ligands requires expertise and various optimization process, which results in higher cost.

6.9.10 Recent Examples of Liposome-Mediated Peptide Drug Delivery in Clinical Trials

Until now, some peptides containing liposomal formulations are in clinical trials, for example, Mepact® containing Mifamurtide which is a synthesized derivative of muramyl dipeptide (MDP), the smallest naturally occurring immunological stimulatory component of mycobacterium species' cell walls. The formulation exhibits similar immunostimulatory properties to natural MDP but with a longer plasma half-life. This formulation was approved in Europe in March 2009 (Kager et al. 2010). Xemys, an immunodominant MBP peptide encapsulated in mannosylated liposomes, is in a clinical trial to treat multiple sclerosis (MS) (Lomakin et al. 2016). MUC-1 is a cell surface glycoprotein that is significantly expressed in lung cancer (Turner 2008; Gandhi et al. 2018). A mucin-1 peptide-based vaccine was formulated with a synthetic lipopeptide and liposome (Tecemotide). The vaccine was designed to induce a cellular immune response to cancer cells that express MUC-1. There are attempts to deliver insulin via different routes of administration using insulin-liposomal formulations (El-Wakeel and Dawoud 2019; Degim et al. 2006; Akimoto et al. 2019; Huang and Wang 2006).

6.10 Summary

Liposomes serve as an attractive platform for drug delivery. Peptide-based liposomes can successfully target the tumor cells and can lower the off-target effects of chemotherapeutics. Toxicity to normal cells by anticancer therapeutics has been dramatically improved by the use of peptide-based liposomes. Therapeutic

peptides can be incorporated inside the liposomes to enhance stability and better tumor accumulation. Apart from that, development of peptide-functionalized liposome to specifically target the desired site has several advantages. Different peptide-functionalized liposomal formulations are being developed for various diseases and for diagnostic applications. Peptide-based liposomes have increased the dimensions of nanomedicine for clinical applications. Surface functionalization of liposomes by peptides is a relatively new technology that requires further study to generate clinically approved therapeutics.

With the ever-growing field of nanotechnology and the success of Covid-19 mRNA vaccines encapsulating the mRNA in lipid capsules developed by Pfizer-BioNtech and Moderna, liposomes as drug delivery agents are getting wide attention. Covid-19 mRNA vaccines were developed using liposomal technology to deliver the mRNA. This overwhelming success of the vaccines further bolsters liposomes as an effective drug delivery agent for susceptible drugs like proteins, peptides, mRNA, DNA, and viral vectors. Peptide-based liposomes for delivery of peptides and for specific targeting are being developed. Various challenges of peptide-based liposomes like difficulty in characterization and quantification of attached ligands, stability issues, masking of ligands, and accelerated blood clearances should be addressed. The development of a different biophysical and biochemical method for accurate quantification and reduction of the batch-to-batch variation for industrial production of surface-functionalized liposomes would help these liposomes to be clinically approved. Further research, including in vitro, pre-clinical, and clinical trials of the peptide-based liposomes, would certainly forward the field of nanomedicine.

References

- Abbasi E, Aval SF, Akbarzadeh A, Milani M, Nasrabadi HT, Joo SW, Hanifehpour Y, Nejati-Koshki K, Pashaei-Asl R. Dendrimers: synthesis, applications, and properties. *Nanoscale Res Lett.* 2014;9:247.
- Abdelwahed W, Degobert G, Fessi H. Investigation of nanocapsules stabilization by amorphous excipients during freeze-drying and storage. *Eur J Pharm Biopharm.* 2006a;63:87–94.
- Abdelwahed W, Degobert G, Stainmesse S, Fessi H. Freeze-drying of nanoparticles: formulation, process and storage considerations. *Adv Drug Deliv Rev.* 2006b;58:1688–713.
- Adibzadeh M, Weder HG, Rehbein A, Schwulera U, Obermeier J, Pawelec G. Activity of liposomal interleukin-2 in vitro. *Mol Biother.* 1992;4:24–8.
- Akbarzadeh A, Rezaei-Sadabady R, Davaran S, Joo SW, Zarghami N, Hanifehpour Y, Samiei M, Kouhi M, Nejati-Koshki K. Liposome: classification, preparation, and applications. *Nanoscale Res Lett.* 2013;8:102.
- Akimoto H, Fukuda-Kawaguchi E, Duramad O, Ishii Y, Tanabe K. A novel liposome formulation carrying both an insulin peptide and a ligand for invariant natural killer T cells induces accumulation of regulatory T cells to islets in nonobese diabetic mice. *J Diabetes Res.* 2019;2019:9430473.
- Allen TM, Cheng WW, Hare JI, Laginha KM. Pharmacokinetics and pharmacodynamics of lipidic nano-particles in cancer. *Anti Cancer Agents Med Chem.* 2006;6:513–23.

- Amin MC, Mohd I, Butt AM, Amjad MW, Kesharwani P. Chapter 5 – Polymeric micelles for drug targeting and delivery. In: Mishra V, Kesharwani P, Amin MCIM, Iyer A, editors. *Nanotechnology-based approaches for targeting and delivery of drugs and genes*. Academic; 2017.
- Amstad E, Kohlbrecher J, Muller E, Schweizer T, Textor M, Reimhult E. Triggered release from liposomes through magnetic actuation of iron oxide nanoparticle containing membranes. *Nano Lett.* 2011;11:1664–70.
- An X, Gui R. Chapter 28 – Stimuli-responsive liposome and control release drug. In: Andronescu E, Grumezescu AM, editors. *Nanostructures for drug delivery*. Elsevier; 2017.
- An X, Zhang F, Zhu Y, Shen W. Photoinduced drug release from thermosensitive AuNPs-liposome using a AuNPs-switch. *Chem Commun (Camb)*. 2010;46:7202–4.
- Andresen TL, Jensen SS, Jorgensen K. Advanced strategies in liposomal cancer therapy: problems and prospects of active and tumor specific drug release. *Prog Lipid Res.* 2005;44:68–97.
- Anilkumar TS, Shalumon KT, Chen JP. Applications of magnetic liposomes in cancer therapies. *Curr Pharm Des.* 2019;25:1490–504.
- Arien A, Toulme-Henry N, Dupuy B. Cholate-induced disruption of calcitonin-loaded liposomes: formation of trypsin-resistant lipid-calcitonin-cholate complexes. *Pharm Res.* 1995;12:1289–92.
- Awada A, Gil T, Sales F, Dubuisson M, Vereecken P, Klastersky J, Moerman C, de Valeriola D, Piccart MJ. Prolonged schedule of temozolomide (Temodal) plus liposomal doxorubicin (Caelyx) in advanced solid cancers. *Anti-Cancer Drugs.* 2004;15:499–502.
- Barenholz Y. Liposome application: problems and prospects. *Curr Opin Colloid Interface Sci.* 2001;6:66–77.
- Beltrán-Gracia E, López-Camacho A, Higuera-Ciajara I, Velázquez-Fernández JB, Vallejo-Cardona AA. Nanomedicine review: clinical developments in liposomal applications. *Cancer Nanotechnology.* 2019;10:11.
- Bendas G, Wilhelm F, Richter W, Nuhn P. Synthetic glycolipids as membrane-bound cryoprotectants in the freeze-drying process of liposomes. *Eur J Pharm Sci.* 1996;4:211–22.
- Bi R, Shao W, Wang Q, Zhang N. Spray-freeze-dried dry powder inhalation of insulin-loaded liposomes for enhanced pulmonary delivery. *J Drug Target.* 2008;16:639–48.
- Bisht S, Schlesinger M, Rupp A, Schubert R, Nolting J, Wenzel J, Holdenrieder S, Brossart P, Bendas G, Feldmann G. A liposomal formulation of the synthetic curcumin analog EF24 (Lipo-EF24) inhibits pancreatic cancer progression: towards future combination therapies. *J Nanobiotechnology.* 2016;14:57.
- Briuglia ML, Rotella C, McFarlane A, Lamprou DA. Influence of cholesterol on liposome stability and on in vitro drug release. *Drug Deliv Transl Res.* 2015;5:231–42.
- Burgess NC, Sharp TH, Thomas F, Wood CW, Thomson AR, Zaccai NR, Brady RL, Serpell LC, Woolfson DN. Modular design of self-assembling peptide-based nanotubes. *J Am Chem Soc.* 2015;137:10554–62.
- Byrne JD, Betancourt T, Brannon-Peppas L. Active targeting schemes for nanoparticle systems in cancer therapeutics. *Adv Drug Deliv Rev.* 2008;60:1615–26.
- Çağdaş M, Sezer AD, Bucak S. Liposomes as potential drug carrier systems for drug delivery. Application of nanotechnology in drug delivery. 2014:1–100.
- Cao SJ, Xu S, Wang HM, Ling Y, Dong J, Xia RD, Sun XH. Nanoparticles: oral delivery for protein and peptide drugs. *AAPS PharmSciTech.* 2019;20:190.
- Cardone RA, Casavola V, Reshkin SJ. The role of disturbed pH dynamics and the Na PLUS_SPI /H PLUS_SPI exchanger in metastasis. *Nat Rev Cancer.* 2005;5:786–95.
- Challa R, Ahuja A, Ali J, Khar RK. Cyclodextrins in drug delivery: an updated review. *AAPS PharmSciTech.* 2005;6:E329–E57.
- Chen C, Han D, Cai C, Tang X. An overview of liposome lyophilization and its future potential. *J Control Release.* 2010;142:299–311.
- Chountoules M, Kyrili A, Pippa N, Meristoudi A, Pispas S, Demetzos C. The modulation of physicochemical characterization of innovative liposomal platforms: the role of the grafted thermoresponsive polymers. *Pharm Dev Technol.* 2017;22:330–5.

- Chowdhury DF Pharmaceutical nanosystems: manufacture, characterization, and safety. In: *Pharmaceutical manufacturing handbook*. 2008.
- Corona-Ortega T, Rangel-Corona R, Hernández-Jiménez M, Baeza I, Ibáñez M, Weiss-Steider B. Characterization of cationic liposomes having IL-2 expressed on their external surface, and their affinity to cervical cancer cells expressing the IL-2 receptor. *J Drug Target*. 2009;17:496–501.
- Colton, J. S., S. D. Erickson, T. J. Smith, and R. K. Watt. “Sensitive detection of surface-and size-dependent direct and indirect band gap transitions in ferritin.” *Nanotechnology 2* (2014): 25(13) 135703.
- Cristiano MC, Cosco D, Celia C, Tudose A, Mare R, Paolino D, Fresta M. Anticancer activity of all-trans retinoic acid-loaded liposomes on human thyroid carcinoma cells. *Colloids Surf B: Biointerfaces*. 2017;150:408–16.
- Dagan N, Barda N, Kepten E, Miron O, Perchik S, Katz MA, Hernán MA, Lipsitch M, Reis B, Balicer RD. BNT162b2 mRNA Covid-19 vaccine in a nationwide mass vaccination setting. *N Engl J Med*. 2021;384:1412–23.
- Dams ET, Laverman P, Oyen WJ, Storm G, Scherphof GL, van Der Meer JW, Corstens FH, Boerman OC. Accelerated blood clearance and altered biodistribution of repeated injections of sterically stabilized liposomes. *J Pharmacol Exp Ther*. 2000;292:1071–9.
- Danaei M, Dehghankhold M, Ataei S, Hasanzadeh Davarani F, Javanmard R, Dokhani A, Khorasani S, Mozafari MR. Impact of particle size and polydispersity index on the clinical applications of lipidic nanocarrier systems. *Pharmaceutics*. 2018;10
- Danhier F, Feron O, Preat V. To exploit the tumor microenvironment: passive and active tumor targeting of nanocarriers for anti-cancer drug delivery. *J Control Release*. 2010;148:135–46.
- De Jong WH, Borm PJ. Drug delivery and nanoparticles: applications and hazards. *Int J Nanomedicine*. 2008;3:133–49.
- Degim Z, Unal N, Essiz D, Abbasoglu U. The effect of various liposome formulations on insulin penetration across Caco-2 cell monolayer. *Life Sci*. 2004;75:2819–27.
- Degim IT, Gumusel B, Degim Z, Ozcelikay T, Tay A, Guner S. Oral administration of liposomal insulin. *J Nanosci Nanotechnol*. 2006;6:2945–9.
- Deshpande PP, Biswas S, Torchilin VP. Current trends in the use of liposomes for tumor targeting. *Nanomedicine (London)*. 2013;8:1509–28.
- Diaz S, Amalfa F, de Lopez AB, Disalvo EA. Effect of water polarized at the carbonyl groups of phosphatidylcholines on the dipole potential of lipid bilayers. *Langmuir*. 1999;15:5179–82.
- Ding Y, Cui W, Sun D, Wang GL, Hei Y, Meng S, Chen JH, Xie Y, Wang ZQ. In vivo study of doxorubicin-loaded cell-penetrating peptide-modified pH-sensitive liposomes: biocompatibility, bio-distribution, and pharmacodynamics in BALB/c nude mice bearing human breast tumors. *Drug Des Devel Ther*. 2017;11:3105–17.
- Du Plessis J, Ramachandran C, Weiner N, Müller DG. The influence of lipid composition and lamellarity of liposomes on the physical stability of liposomes upon storage. *Int J Pharm*. 1996;127:273–8.
- Dua JS, Rana AC, Bhandari AK. Liposome: methods of preparation and applications. *Int J Pharm Stud Res*. 2012;3:14–20.
- Edmonds DJ, Price DA. Oral GLP-1 modulators for the treatment of diabetes. *Annu Rep Med Chem*. 2013;48:119–30.
- El-Wakeel NM, Dawoud MHS. Topical insulin-liposomal formulation in management of recurrent aphthous ulcers: a randomized placebo-controlled trial. *J Investig Clin Dent*. 2019;10:e12437.
- Enzian P, Schell C, Link A, Malich C, Pries R, Wollenberg B, Rahmzadeh R. Optically controlled drug release from light-sensitive liposomes with the new photosensitizer 5,10-DiOH. *Mol Pharm*. 2020;17:2779–88.
- Esparza K, Jayawardena D, Onyuksel H. Phospholipid micelles for peptide drug delivery. *Methods Mol Biol*. 2019;2000:43–57.
- Fan Y, Marioli M, Zhang K. Analytical characterization of liposomes and other lipid nanoparticles for drug delivery. *J Pharm Biomed Anal*. 2021;192:113642.

- Felber AE, Dufresne MH, Leroux JC. pH-sensitive vesicles, polymeric micelles, and nanospheres prepared with polycarboxylates. *Adv Drug Deliv Rev.* 2012;64:979–92.
- Gandhi L, Vansteenkiste JF, Shepherd FA. 50 – Immunotherapy and lung cancer. In: Pass HI, Ball D, Scagliotti GV, editors. *IASLC thoracic oncology*. 2nd ed. Elsevier: Philadelphia; 2018.
- Ghadiri MR, Granja JR, Milligan RA, McRee DE, Khazanovich N. Self-assembling organic nanotubes based on a cyclic peptide architecture. *Nature.* 1993;366:324–7.
- Gomes, Ariane C., Mona Mohsen, and Martin F. Bachmann. “Harnessing nanoparticles for immunomodulation and vaccines.” *Vaccines* 5.1 (2017): 5(1) 6.
- Gomez-Hens A, Fernandez-Romero JM. Analytical methods for the control of liposomal delivery systems. *TrAC Trends Anal Chem.* 2006;25:167–78.
- Hajos F, Stark B, Hensler S, Prassl R, Mosgoeller W. Inhalable liposomal formulation for vasoactive intestinal peptide. *Int J Pharm.* 2008;357:286–94.
- Han Q, Jia X, Qian Y, Wang Z, Yang S, Jia Y, Wang W, Hu Z. Peptide functionalized targeting liposomes: for nanoscale drug delivery towards angiogenesis. *J Mater Chem B.* 2016;4:7087–91.
- Hartgerink JD, Granja JR, Milligan RA, Reza Ghadiri M. Self-Assembling peptide nanotubes. *J Am Chem Soc.* 1996;118:43–50.
- Henninot A, Collins JC, Nuss JM. The current state of peptide drug discovery: back to the future? *J Med Chem.* 2018;61:1382–414.
- Henry CR. Morphology of supported nanoparticles. *Prog Surf Sci.* 2005;80:92–116.
- Huang SL. Ultrasound-responsive liposomes. *Methods Mol Biol.* 2010;605:113–28.
- Huang YY, Wang CH. Pulmonary delivery of insulin by liposomal carriers. *J Control Release.* 2006;113:9–14.
- Hudson, Richard DA. “Ferrocene polymers: current architectures, syntheses and utility.” *Journal of Organometallic Chemistry* 637 (2001): 47-69.
- Huwlyer J, Drewe J, Krahenbuhl S. Tumor targeting using liposomal antineoplastic drugs. *Int J Nanomedicine.* 2008;3:21–9.
- Hwang SH, Maitani Y, Qi XR, Takayama K, Nagai T. Remote loading of diclofenac, insulin and fluorescein isothiocyanate labeled insulin into liposomes by pH and acetate gradient methods. *Int J Pharm.* 1999;179:85–95.
- Hwang SH, Maitani Y, Takayama K, Nagai T. High entrapment of insulin and bovine serum albumin into neutral and positively-charged liposomes by the remote loading method. *Chem Pharm Bull (Tokyo).* 2000;48:325–9.
- Immordino ML, Dosio F, Cattel L. Stealth liposomes: review of the basic science, rationale, and clinical applications, existing and potential. *Int J Nanomedicine.* 2006;1:297–315.
- Irie T, Uekama K. Cyclodextrins in peptide and protein delivery. *Adv Drug Deliv Rev.* 1999;36:101–23.
- Ishida T, Kiwada H. Accelerated blood clearance (ABC) phenomenon upon repeated injection of PEGylated liposomes. *Int J Pharm.* 2008;354:56–62.
- Ishida T, Harashima H, Kiwada H. Interactions of liposomes with cells in vitro and in vivo: opsonins and receptors. *Curr Drug Metab.* 2001;2:397–409.
- Ishida T, Masuda K, Ichikawa T, Ichihara M, Irimura K, Kiwada H. Accelerated clearance of a second injection of PEGylated liposomes in mice. *Int J Pharm.* 2003;255:167–74.
- Joffret ML, Morgeaux S, Laclerc C, Oth D, Zanetti C, Sureau P, Perrin P. Enhancement of interleukin-2 activity by liposomes. *Vaccine.* 1990;8:385–9.
- Otvos L Jr, Wade JD. Current challenges in peptide-based drug discovery. *Front Chem.* 2014;2:–62.
- Kaur, Jashandeep, Gurlal Singh Gill, and Kiran Jeet. “Applications of carbon nanotubes in drug delivery: A comprehensive review.” *Characterization and biology of nanomaterials for drug delivery* (2019): 113–135.
- Kager L, Pötschger U, Bielack S. Review of mifamurtide in the treatment of patients with osteosarcoma. *Ther Clin Risk Manag.* 2010;6:279–86.
- Kanwar JR, Long BM, Kanwar KR. The use of cyclodextrins nanoparticles for oral delivery. *Curr Med Chem.* 2011;18:2079–85.
- Katayama K, Kato Y, Onishi H, Nagai T, Machida Y. Double liposomes: hypoglycemic effects of liposomal insulin on normal rats. *Drug Dev Ind Pharm.* 2003;29:725–31.

- Kato Y, Hosokawa T, Hayakawa E, Ito K. Influence of liposomes on tryptic digestion of insulin. II. *Biol Pharm Bull.* 1993;16:740–4.
- Kawasaki ES, Player A. Nanotechnology, nanomedicine, and the development of new, effective therapies for cancer. *Nanomedicine.* 2005;1:101–9.
- Kedar E, Gur H, Babai I, Samira S, Even-Chen S, Barenholz Y. Delivery of cytokines by liposomes: hematopoietic and immunomodulatory activity of interleukin-2 encapsulated in conventional liposomes and in long-circulating liposomes. *J Immunother.* 2000;23:131–45.
- Kim K, Lee M, Park H, Kim JH, Kim S, Chung H, Choi K, Kim IS, Seong BL, Kwon IC. Cell-permeable and biocompatible polymeric nanoparticles for apoptosis imaging. *J Am Chem Soc.* 2006;128:3490–1.
- Kisel MA, Kulik LN, Tsybovsky IS, Vlasov AP, Vorob'yov MS, Kholodova EA, Zabarovskaya ZV. Liposomes with phosphatidylethanol as a carrier for oral delivery of insulin: studies in the rat. *Int J Pharm.* 2001;216:105–14.
- Kuntsche J, Horst JC, Bunjes H. Cryogenic transmission electron microscopy (cryo-TEM) for studying the morphology of colloidal drug delivery systems. *Int J Pharm.* 2011;417:120–37.
- Law SL, Shih CL. Characterization of calcitonin-containing liposome formulations for intranasal delivery. *J Microencapsul.* 2001;18:211–21.
- Li J, Wang X, Zhang T, Wang C, Huang Z, Luo X, Deng Y. A review on phospholipids and their main applications in drug delivery systems. *Asian Journal of Pharmaceutical Sciences.* 2015;10:81–98.
- Lin W, Xie X, Deng J, Liu H, Chen Y, Fu X, Liu H, Yang Y. Cell-penetrating peptide-doxorubicin conjugate loaded NGR-modified nanobubbles for ultrasound triggered drug delivery. *J Drug Target.* 2016;24:134–46.
- Liu W, Wei F, Ye A, Tian M, Han J. Kinetic stability and membrane structure of liposomes during *in vitro* infant intestinal digestion: effect of cholesterol and lactoferrin. *Food Chem.* 2017;230:6–13.
- Lomakin Y, Belogurov A, Glagoleva I, Stepanov A, Zakharov K, Okunola J, Smirnov I, Genkin D, Gabibov A. Administration of myelin basic protein peptides encapsulated in mannosylated liposomes normalizes level of serum TNF- α and IL-2 and chemoattractants CCL2 and CCL4 in multiple sclerosis patients. *Mediat Inflamm.* 2016;2016:2847232.
- Lombardo D, Kiselev MA, Caccamo MT. Smart nanoparticles for drug delivery application: development of versatile nanocarrier platforms in biotechnology and nanomedicine. *J Nanomater.* 2019;2019:3702518.
- Lopes SC d A. Liposomes as carriers of anticancer drugs. In: dos Santos Giuberti C, editor. . Rijeka: IntechOpen; 2013.
- Lowman AM, Morishita M, Kajita M, Nagai T, Peppas NA. Oral delivery of insulin using pH-responsive complexation gels. *J Pharm Sci.* 1999;88:933–7.
- Luzardo MC, Amalfa F, Nunez AM, Diaz S, Biondi De Lopez AC, Disalvo EA. Effect of trehalose and sucrose on the hydration and dipole potential of lipid bilayers. *Biophys J.* 2000;78:2452–8.
- Mahajan S, Patharkar A, Kuche K, Maheshwari R, Deb PK, Kalia K, Tekade RK. Functionalized carbon nanotubes as emerging delivery system for the treatment of cancer. *Int J Pharm.* 2018;548:540–58.
- Maitani Y, Asano S, Takahashi S, Nakagaki M, Nagai T. Permeability of insulin entrapped in liposome through the nasal mucosa of rabbits. *Chem Pharm Bull (Tokyo).* 1992;40:1569–72.
- Mathiyazhakan M, Wiraja C, Xu C. A concise review of gold nanoparticles-based photo-responsive liposomes for controlled drug delivery. *Nano Lett.* 2018;10:10.
- Matsuo H, Wakasugi M, Takanaga H, Ohtani H, Naito M, Tsuruo T, Sawada Y. Possibility of the reversal of multidrug resistance and the avoidance of side effects by liposomes modified with MRK-16, a monoclonal antibody to P-glycoprotein. *J Control Release.* 2001;77:77–86.
- Maxwell T, Campos MGN, Smith S, Doomra M, Thwin Z, Santra S. Chapter 15 – Quantum dots. In: Chung EJ, Leon L, Rinaldi C, editors. *Nanoparticles for biomedical applications.* Elsevier; 2020.

- Mazumdar A, Haddad Y, Milosavljevic V, Michalkova H, Guran R, Bhowmick S, Moulick A. Peptide-carbon quantum dots conjugate, derived from human retinoic acid receptor responder protein 2, against antibiotic-resistant gram positive and gram negative pathogenic bacteria. *Nanomaterials* (Basel). 2020;10
- Meerovich I, Dash AK. Polymersomes for drug delivery and other biomedical applications. In: *Materials for Biomedical Engineering*. Elsevier; 2019.
- Miranda D, Lovell JF. Mechanisms of light-induced liposome permeabilization. *Bioeng Transl Med*. 2016;1:267–76.
- Moghipi SM, Hunter AC, Murray JC. Long-circulating and target-specific nanoparticles: theory to practice. *Pharmacol Rev*. 2001;53:283–318.
- Mohammed AR, Weston N, Coombes AG, Fitzgerald M, Perrie Y. Liposome formulation of poorly water soluble drugs: optimisation of drug loading and ESEM analysis of stability. *Int J Pharm*. 2004;285:23–34.
- Mohammed AR, Bramwell VW, Coombes AG, Perrie Y. Lyophilisation and sterilisation of liposomal vaccines to produce stable and sterile products. *Methods*. 2006;40:30–8.
- Morimoto BH. Enhancing the pharmaceutical properties of peptides. *Pharma Horizon*. 2017;1:29–31.
- Mozafari MR. Liposomes: an overview of manufacturing techniques. *Cell Mol Biol Lett*. 2005a;10:711–9.
- Mozafari RM. *Nanoliposomes: from fundamentals to recent developments*. Trafford; 2005b.
- Mukherjee B, Patra B, Layek B, Mukherjee A. Sustained release of acyclovir from nano-liposomes and nano-niosomes: an in vitro study. *Int J Nanomedicine*. 2007;2:213–25.
- Muramatsu K, Maitani Y, Nagai T. Dipalmitoylphosphatidylcholine liposomes with soybean-derived sterols and cholesterol as a carrier for the oral administration of insulin in rats. *Biol Pharm Bull*. 1996;19:1055–8.
- Muttenthaler M, King GF, Adams DJ, Alewood PF. Trends in peptide drug discovery. *Nat Rev Drug Discov*. 2021;20:309–25.
- Nagayasu A, Uchiyama K, Kiwada H. The size of liposomes: a factor which affects their targeting efficiency to tumors and therapeutic activity of liposomal antitumor drugs. *Adv Drug Deliv Rev*. 1999;40:75–87.
- Naik H, Sonju JJ, Singh S, Chatzistamou I, Shrestha L, Gauthier T, Jois S. Lipidated peptidomimetic ligand-functionalized HER2 targeted liposome as nano-carrier designed for doxorubicin delivery in cancer therapy. *Pharmaceuticals* (Basel). 2021;14
- Neville ME, Boni LT, Pflug LE, Popescu MC, Robb RJ. Biopharmaceutics of liposomal interleukin 2, oncolipin. *Cytokine*. 2000;12:1691–701.
- Niu M, Tan Y, Guan P, Hovgaard L, Lu Y, Qi J, Lian R, Li X, Wu W. Enhanced oral absorption of insulin-loaded liposomes containing bile salts: a mechanistic study. *Int J Pharm*. 2014;460:119–30.
- Olusanya TOB, Ahmad RRH, Ibegbu DM, Smith JR, Elkordy AA. Liposomal drug delivery systems and anticancer drugs. *Molecules*. 2018;23:907.
- Park K. Nanotechnology: what it can do for drug delivery. *J Control Release*. 2007;120:1–3.
- Park K, Kwon IC, Park K. Oral protein delivery: current status and future prospect. *React Funct Polym*. 2011a;71:280–7.
- Park SJ, Choi SG, Davaa E, Park JS. Encapsulation enhancement and stabilization of insulin in cationic liposomes. *Int J Pharm*. 2011b;415:267–72.
- Parmentier J, Thewes B, Gropp F, Fricker G. Oral peptide delivery by tetraether lipid liposomes. *Int J Pharm*. 2011;415:150–7.
- Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, Perez JL, Marc GP, Moreira ED, Zerbini C. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N Engl J Med*. 2020;384(16):1577.
- Prasmickaite L, Hogset A, Selbo PK, Engesaeter BO, Hellum M, Berg K. Photochemical disruption of endocytic vesicles before delivery of drugs: a new strategy for cancer therapy. *Br J Cancer*. 2002;86:652–7.

- Probst CE, Zrazhevskiy P, Bagalkot V, Gao X. Quantum dots as a platform for nanoparticle drug delivery vehicle design. *Adv Drug Deliv Rev.* 2013;65:703–18.
- Raffa V, Vittorio O, Riggio C, Cuschieri A. Progress in nanotechnology for healthcare. *Minim Invasive Ther Allied Technol.* 2010;19:127–35.
- Rangel-Yagui CO, Pessoa A Jr, Tavares LC. Micellar solubilization of drugs. *J Pharm Sci.* 2005;8:147–65.
- Refai E, Jonsson C, Andersson M, Jacobsson H, Larsson S, Kogner P, Hassan M. Biodistribution of liposomal 131I-VIP in rat using gamma camera. *Nucl Med Biol.* 1999;26:931–6.
- Renunkuntla J, Vadlapudi AD, Patel A, Boddu SHS, Mitra AK. Approaches for enhancing oral bioavailability of peptides and proteins. *Int J Pharm.* 2013;447:75–93.
- Rezaei N, Mehrnejad F, Vaezi Z, Mosslim Sedghi S, Asghari M, Naderi-Manesh H. Encapsulation of an endostatin peptide in liposomes: stability, release, and cytotoxicity study. *Colloids Surf B: Biointerfaces.* 2020;185:110552.
- Rho JY, Cox H, Mansfield EDH, Ellacott SH, Peltier R, Brendel JC, Hartlieb M, Waigh TA, Perrier S. Dual self-assembly of supramolecular peptide nanotubes to provide stabilisation in water. *Nat Commun.* 2019;10:4708.
- Riaz MK, Riaz MA, Zhang X, Lin C, Wong KH, Chen X, Zhang G, Lu A, Yang Z. Surface functionalization and targeting strategies of liposomes in solid tumor therapy: a review. *Int J Mol Sci.* 2018;19:195.
- Robson AL, Dastoor PC, Flynn J, Palmer W, Martin A, Smith DW, Woldu A, Hua S. Advantages and limitations of current imaging techniques for characterizing liposome morphology. *Front Pharmacol.* 2018;9:80.
- Roldo M. Carbon nanotubes in drug delivery: just a carrier? *Ther Deliv.* 2016;7:55–7.
- Sadler K, Tam JP. Peptide dendrimers: applications and synthesis. *Rev Mol Biotechnol.* 2002;90:195–229.
- Samimi S, Maghsoudnia N, Eftekhari RB, Dorkoosh F. Chapter 3 – Lipid-based nanoparticles for drug delivery systems. In: Mohapatra SS, Ranjan S, Dasgupta N, Mishra RK, Thomas S, editors. *Characterization and Biology of Nanomaterials for Drug Delivery.* Elsevier; 2019.
- Santivarangkna C, Higl B, Foerst P. Protection mechanisms of sugars during different stages of preparation process of dried lactic acid starter cultures. *Food Microbiol.* 2008;25:429–41.
- Sejourne F, Suzuki H, Alkan-Onyuksel H, Gao XP, Ikezaki H, Rubinstein I. Mechanisms of vasodilation elicited by VIP in sterically stabilized liposomes in vivo. *Am J Phys.* 1997;273:R287–92.
- Sercombe L, Veerati T, Moheimani F, Wu SY, Sood AK, Hua S. Advances and challenges of liposome assisted drug delivery. *Front Pharmacol.* 2015;6:286.
- Shen J, Burgess DJ. In Vitro dissolution testing strategies for nanoparticulate drug delivery systems: recent developments and challenges. *Drug Deliv Transl Res.* 2013;3:409–15.
- Shen S, Huang D, Cao J, Chen Y, Zhang X, Guo S, Ma W, Qi X, Ge Y, Wu L. Magnetic liposomes for light-sensitive drug delivery and combined photothermal-chemotherapy of tumors. *J Mater Chem B.* 2019;7:1096–106.
- Shin MD, Shukla S, Chung YH, Beiss V, Chan SK, Ortega-Rivera OA, Wirth DM, Chen A, Sack M, Pokorski JK, Steinmetz NF. COVID-19 vaccine development and a potential nanomaterial path forward. *Nat Nanotechnol.* 2020;15:646–55.
- Singh R, Lillard JW Jr. Nanoparticle-based targeted drug delivery. *Exp Mol Pathol.* 2009;86:215–23.
- Soares DC, de Oliveira MC, dos Santos RG, Andrade MS, Vilela JM, Cardoso VN, Ramaldes GA. Liposomes radiolabeled with 159Gd-DTPA-BMA: preparation, physicochemical characterization, release profile and in vitro cytotoxic evaluation. *Eur J Pharm Sci.* 2011;42:462–9.
- Song KH, Chung SJ, Shim CK. Enhanced intestinal absorption of salmon calcitonin (sCT) from proliposomes containing bile salts. *J Control Release.* 2005;106:298–308.
- Sonju JJ, Dahal A, Singh SS, Jois SD. Peptide-functionalized liposomes as therapeutic and diagnostic tools for cancer treatment. *J Control Release.* 2021;329:624–44.
- Srinivasa-Gopalan S, Yarema KJ. *Nanotechnologies for the life sciences: dendrimers in cancer treatment and diagnosis.* New York: Wiley; 2007.

- Stark B, Debbage P, Andraea F, Mosgoeller W, Prassl R. Association of vasoactive intestinal peptide with polymer-grafted liposomes: structural aspects for pulmonary delivery. *Biochim Biophys Acta*. 2007;1768:705–14.
- Stark B, Andraea F, Mosgoeller W, Edetsberger M, Gaubitzer E, Koehler G, Prassl R. Liposomal vasoactive intestinal peptide for lung application: protection from proteolytic degradation. *Eur J Pharm Biopharm*. 2008;70:153–64.
- Stiriba SE, Frey H, Haag R. Dendritic polymers in biomedical applications: from potential to clinical use in diagnostics and therapy. *Angew Chem Int Ed Eng*. 2002;41:1329–34.
- Su Z, Shen H, Wang H, Wang J, Li J, Nienhaus GU, Shang L, Wei G. Motif-designed peptide nanofibers decorated with graphene quantum dots for simultaneous targeting and imaging of tumor cells. *Adv Funct Mater*. 2015;25:5472–8.
- Sun Y, Ji Y, Haiyan Y, Wang D, Cao M, Wang J. Near-infrared light-sensitive liposomes for controlled release. *RSC Adv*. 2016;6:81245–9.
- Suzuki H, Noda Y, Gao XP, Sejourne F, Alkan-Onyuksel H, Paul S, Rubinstein I. Encapsulation of VIP into liposomes restores vasorelaxation in hypertension in situ. *Am J Phys*. 1996;271:H282–7.
- Swaminathan J, Ehrhardt C. Liposomal delivery of proteins and peptides. *Expert Opin Drug Deliv*. 2012;9:1489–503.
- Tarun G, Goyal Amit K. Liposomes: targeted and controlled delivery system. *Drug Delivery Letters*. 2014;4:62–71.
- Tenchov R, Bird R, Curtze AE, Zhou Q. Lipid nanoparticles—from liposomes to mRNA vaccine delivery, a landscape of research diversity and advancement. *ACS Nano*. 2021;15(11):16982–7015.
- Tomalia DA, Fréchet JMJ. Discovery of dendrimers and dendritic polymers: a brief historical perspective. *J Polym Sci A Polym Chem*. 2002;40:2719–28.
- Tonggu L, Wang L. Cryo-EM sample preparation method for extremely low concentration liposomes. *Ultramicroscopy*. 2020;208:112849.
- Torchilin VP. Targeted pharmaceutical nanocarriers for cancer therapy and imaging. *AAPS J*. 2007;9:E128–47.
- Torchilin V. Multifunctional and stimuli-sensitive pharmaceutical nanocarriers. *Eur J Pharm Biopharm*. 2009;71:431–44.
- Turner LS. Stimuvax. In: Enna SJ, Bylund DB, editors. *xPharm: the comprehensive pharmacology reference*. New York: Elsevier; 2008.
- van Rijt, Sabine H., Thomas Bein, and Silke Meiners. “Medical nanoparticles for next generation drug delivery to the lungs.” (2014): 44(3) 765–774.
- Villarreal MA, Diaz SB, Disalvo EA, Montich GG. Molecular dynamics simulation study of the interaction of trehalose with lipid membranes. *Langmuir*. 2004;20:7844–51.
- Vlieghe P, Lisowski V, Martinez J, Khrestchatisky M. Synthetic therapeutic peptides: science and market. *Drug Discov Today*. 2010;15:40–56.
- Wagner A, Vorauer-Uhl K. Liposome technology for industrial purposes. *J Drug Deliv*. 2011;2011:591325.
- Wang CX, Li CL, Zhao X, Yang HY, Wei N, Li YH, Zhang L, Zhang L. Pharmacodynamics, pharmacokinetics and tissue distribution of liposomal mitoxantrone hydrochloride. *Yao Xue Bao*. 2010;45:1565–9.
- Wang WT, Liu JG, Sun WH, Wang WY, Wang SL, Zhu NH. Widely tunable single bandpass microwave photonic filter based on Brillouin-assisted optical carrier recovery. *Opt Express*. 2014;22:29304–13.
- Werle M, Takeuchi H. Chitosan–aprotinin coated liposomes for oral peptide delivery: development, characterisation and in vivo evaluation. *Int J Pharm*. 2009;370:26–32.
- Wu, Wei, Quanguo He, and Changzhong Jiang. “Magnetic iron oxide nanoparticles: synthesis and surface functionalization strategies.” *Nanoscale research letters* 3.11 (2008): 397–415.
- Xie X, Lin W, Liu H, Deng J, Chen Y, Liu H, Fu X, Yang Y. Ultrasound-responsive nanobubbles contained with peptide-campthothecin conjugates for targeted drug delivery. *Drug Deliv*. 2016;23:2756–64.

- Yavlovich A, Smith B, Gupta K, Blumenthal R, Puri A. Light-sensitive lipid-based nanoparticles for drug delivery: design principles and future considerations for biological applications. *Mol Membr Biol.* 2010;27:364–81.
- Ye J, Liu E, Yu Z, Pei X, Chen S, Zhang P, Shin MC, Gong J, He H, Yang VC. CPP-assisted intracellular drug delivery, what is next? *Int J Mol Sci.* 2016;17(11):1892.
- Zhang J-a A, Pawelchak J. Effect of pH, ionic strength and oxygen burden on the chemical stability of EPC/cholesterol liposomes under accelerated conditions: part 1: lipid hydrolysis. *Eur J Pharm Biopharm.* 2000;50:357–64.
- Zhang ZH, Abbad S, Pan RR, Waddad AY, Hou LL, Lv HX, Zhou JP. N-octyl-N-Arginine chitosan micelles as an oral delivery system of insulin. *J Biomed Nanotechnol.* 2013;9:601–9.
- Zhang Q, Lu L, Zhang L, Shi K, Cun X, Yang Y, Liu Y, Gao H, He Q. Dual-functionalized liposomal delivery system for solid tumors based on RGD and a pH-responsive antimicrobial peptide. *Sci Rep.* 2016;6:19800.
- Zhao Y, Ren W, Zhong T, Zhang S, Huang D, Guo Y, Yao X, Wang C, Zhang WQ, Zhang X, Zhang Q. Tumor-specific pH-responsive peptide-modified pH-sensitive liposomes containing doxorubicin for enhancing glioma targeting and anti-tumor activity. *J Control Release.* 2016;222:56–66.
- Zhou XH, Po ALW. Peptide and protein drugs: II. Non-parenteral routes of delivery. *Int J Pharm.* 1991;75:117–30.

Chapter 7

Peptides and Their Delivery to the Brain



Waleed Elballa, Kelly Schwinghamer, Eric Ebert, and Teruna J. Siahaan

Contents

7.1	Introduction.....	238
7.2	Structure of the BBB and In Vitro and In Vivo Models for Brain Delivery.....	240
7.2.1	Method for Assessing BBB Permeability.....	241
7.3	Passive Diffusion Across the BBB via Transcellular Pathway.....	242
7.4	Receptor-Mediated Transcytosis of Peptides Through the BBB.....	245
7.5	Peptide Conjugates for Delivering Drugs Across the BBB.....	248
7.6	Brain Drug Delivery Using Nanoparticles.....	250
7.7	Modulation of the BBB to Improve Delivery via Paracellular Pathways.....	252
7.7.1	Osmotic Blood-Brain Barrier Disruption (BBBD) Method.....	253
7.7.2	Blood-Brain Barrier Modulators (BBBMs) of the Intercellular Junction Proteins.....	254
7.8	Nasal Delivery of Peptides.....	259
7.9	Conclusions.....	262
	References.....	262

Abstract There has been great progress in utilizing peptides as therapeutic agents in clinical settings. More than 500 peptides are being investigated in preclinical studies; however, there are only a small number of peptides being used for brain diseases. This is due to the difficulty of delivering peptides to the brain. Therefore, this chapter describes the progress of developing methods to deliver peptides to the brain. Several different pathways that are used by peptides to enter the brain from the bloodstream are discussed. Various methods and factors that have been explored for improving the delivery of peptides into the brain are reviewed here.

W. Elballa · K. Schwinghamer · E. Ebert · T. J. Siahaan (✉)
Department of Pharmaceutical Chemistry, School of Pharmacy, The University of Kansas,
Kansas, USA
e-mail: welballa@ku.edu; kschwin2@ku.edu; eric.ebert@ku.edu; siahaan@ku.edu

Keywords Blood-brain barrier · Cyclization · Receptor-mediated transcytosis
· Paracellular pathway · Intranasal delivery · Intercellular junction proteins
· Peptide conjugates

7.1 Introduction

In the past decades, there have been many advancements in developing peptides as therapeutic and diagnostic agents (Fosgerau and Hoffmann 2015; Kaspar and Reichert 2013; Uhlig et al. 2014) (Fosgerau and Hoffmann 2015). With increasing knowledge to stabilize, formulate, and deliver peptides, the number of approved peptide drugs is expected to increase in the future. Examples of peptide drugs on the market include octreotide, exenatide, integrilin, calcitonin, oxytocin, insulin, and many others. Some of these peptide drugs were derived from natural substances or designed from the active region(s) of proteins.

Many peptides have been developed to treat patients with brain diseases such as Alzheimer's disease (AD), Parkinson's disease (PD), multiple sclerosis (MS), and brain tumors (e.g., glioblastoma, medulloblastoma) (Oller-Salvia et al. 2016). Unfortunately, the progress in developing peptides for treating brain diseases is still very limited. One of the potential reasons for the lack of progress is the difficulty in effectively delivering peptides into the brain. One of the major challenges to deliver peptides to the brain is their inability to cross the blood-brain barrier (BBB) (Fig. 7.1). In general, the physicochemical properties of peptides (i.e., size, hydrogen bonding potential, cLogP) prevent their passage across the BBB in therapeutically relevant amounts. In this case, most hydrophilic and charged peptides cannot passively diffuse through the membranes of the BBB endothelial cells into the brain (Fig. 7.1; Path A). Some peptides can cross the BBB due to transporters that carry them across the endothelial cells from the systemic circulation stream into the brain (Fig. 7.1; Path B). Many methods have been investigated to improve the delivery of peptide drugs into the brain, including intranasal brain delivery, modulation of the BBB (e.g., osmotic, ultrasound, and adhesion peptides), receptor-mediated endocytosis, and cell-penetrating peptide methods.

Besides the selectivity and potency of peptide drugs to the target receptors, peptide plasma stability and clearance from the systemic circulation are important factors to consider when developing peptide drugs. Peptides are susceptible to enzymatic degradation by exo- and endo-peptidases in the blood. Peptides can also be cleared by the kidney and the liver from the bloodstream. It has been shown that cyclic peptides have higher stability against peptidases in the bloodstream than their respective linear peptides (McCully et al. 2018). The formation of cyclic peptides has been shown to improve peptide permeation through the biological barriers (e.g., intestinal mucosa barrier and the BBB). The incorporation of D-amino acid into the

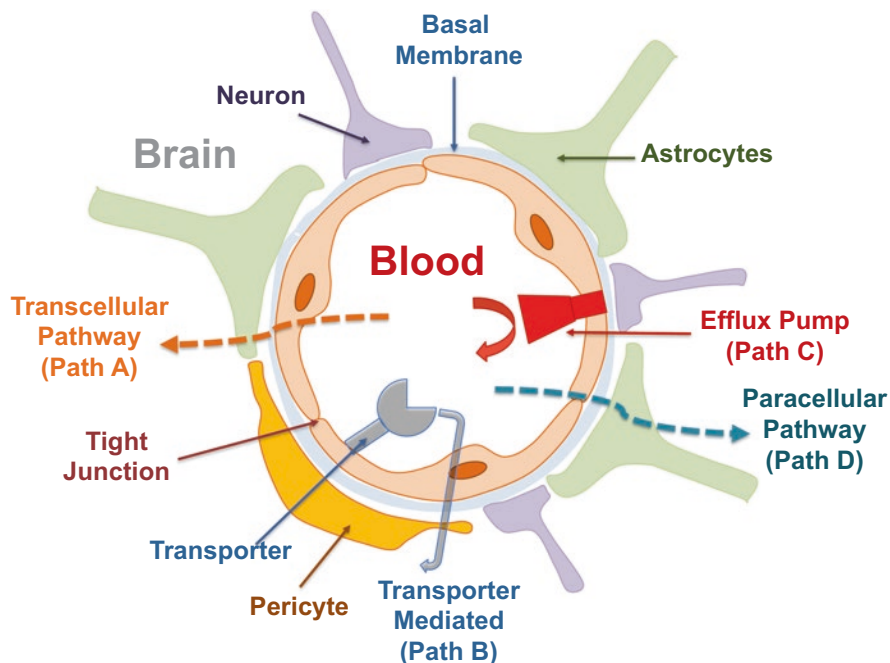


Fig. 7.1 The blood-brain barrier (BBB) is constructed by vascular endothelial cells on the basement membrane, and the cells are connected by the cell-cell adhesion proteins in the intercellular junction to form a tight junction, adherens junction, and desmosome. The vascular endothelial cells are surrounded by astrocytes, neurons, and pericytes. The BBB has transporters (e.g., glucose and amino acid transporter) and efflux pumps (e.g., P-glycoprotein (P-gp)). Peptide drugs can cross the BBB via passive diffusion through the transcellular pathway or paracellular pathway. The transcellular pathway is the diffusion pathway of the peptide through the cell membranes of the BBB, while the paracellular pathway is the diffusion pathway of the peptide between the cells or through the intercellular junctions

peptide could improve plasma stability; this is because proteolytic enzymes do not recognize D-amino acids (McCully et al. 2018). Another method to improve stability without eliminating biological activity is by forming the retro-inverso peptide in which the parent sequence is reversed and all the L-amino acids are replaced with D-amino acids (Chorev et al. 1979; McCully et al. 2018; Ghosh et al. 2010). The rationale is reversing the sequence, and changing the chirality of each amino acid will result in the same presentation of all amino acid side chains on the space for recognition by the target protein or receptor. In addition, the presence of D-amino acid will not be recognized by exo- and endo-peptidases for degradation in the bloodstream and tissues. However, forming the retro-inverso peptide may lower its biological activity because forming the retro-inverso may not mimic the secondary structure of the parent peptide (Li et al. 2010). Furthermore, the reverse peptide

bond presentation in the retro-inverso peptide may not mimic the backbone hydrogen bonding connections between the parent peptide backbone and the receptor; this results in a lower binding affinity of retro-inverso peptide to the receptor compared to the parent peptide (Li et al. 2010).

7.2 Structure of the BBB and In Vitro and In Vivo Models for Brain Delivery

To understand how peptides can enter the brain from the systemic circulation, it is necessary to discuss the structure of the BBB (Fig. 7.1). The BBB is made of endothelial microvessel cells that separate the blood stream and the extracellular fluid of the brain. The BBB endothelial cells wrap around forming tube-like structure with their intercellular cell membranes “glued” to each other by cell-cell adhesion proteins (i.e., occludins, claudins, cadherins, nectins). The abluminal side of the capillary basement membrane of the endothelial cells is populated by extracellular matrix proteins (e.g., collagen) and surrounded by pericytes, neurons, and astrocyte end-feet that anchor the BBB endothelial cells. The BBB has tight junctions in the intercellular space. The BBB endothelial cells are more restrictive compared to those vasculatures found in other parts of the body (Loscher and Potschka 2005; Sharif et al. 2018). The space between the adjacent endothelial cell plasma membranes is referred to as the intercellular junction where small molecules can penetrate through this space as the paracellular transport pathway (Fig. 7.1; Path D). The BBB is distinct from other peripheral capillaries because the BBB endothelial cells are continuous with lacking fenestration as well as having low pinocytosis activity (Loscher and Potschka 2005; Sharif et al. 2018).

The BBB endothelial surface is decorated with receptors, transporters, efflux pumps, and metabolic enzymes. The transporters have a role in carrying nutrient molecules (e.g., glucose, amino acid) into the brain, while the efflux pumps and metabolic enzymes prevent molecules from entering the brain. The efflux pumps include P-glycoprotein (Pgp), breast cancer resistance protein (BCRP), and multi-drug resistance-associated protein (MRP) (Hermann and Bassetti 2007; On and Miller 2014). The enzymes metabolize molecules on the surface as well as the ones that transcellularly cross the endothelial cells of the BBB. The metabolic products of the brain are also transported across the BBB into the systemic circulation for their clearance from the brain. The enzymes in the luminal and abluminal membranes of the BBB are gamma-glutamyl transpeptidase, amino acid transport system A 5'-nucleotidase, alkaline phosphatase, and Na PLUS_SPI /K PLUS_SPI ATPase (Sanchez del Pino et al. 1995). Drug molecules that cross the BBB endothelial cells are subjected to metabolism by cytochrome P450 (CYP 450) enzymes (i.e., CYP1B1, CYP3A4, CYP2U1), and this metabolism prevents the intact molecule to cross the BBB (Gherzi-Egea et al. 1995; Shawahna et al. 2011; Ghosh et al. 2011; Dauchy et al. 2008). CYP3A4 oxidizes various molecules while CYP1B1

metabolizes fatty acids (Shawahna et al. 2011; Ghosh et al. 2011). The activity of CYP 450 is also affected by different brain diseases (Kadry et al. 2020).

7.2.1 Method for Assessing BBB Permeability

To simplify transport study of molecules across the BBB, various in vitro cell culture models were developed. Initially, primary culture of bovine brain microvessel endothelial cells was used as a monolayer on a Transwell™ as a model of the BBB (Shah et al. 1989). Unfortunately, the primary culture of BBB was leaky paracellularly with low trans-endothelial electrical resistance (TEER) values with limited uses for studying paracellular transports of molecules. Thus, the in vitro culture model of BBB is not suitable to study transport properties of small molecules including small peptides with up to six amino acid residues; however, this model can be used to predict the transport properties of large peptides and proteins. To improve the BBB tightness, the primary cultures of BBB endothelial cells were co-cultured with pericytes, astrocytes, glial, and neuron cells to mimic in vivo conditions (Fletcher and Callanan 2012; Lippmann et al. 2013). Currently, many immortalized BBB endothelial cell lines have been developed as BBB in vitro models, including murine cells (e.g., bEnd.3.5, TM-BBB), rat cells (e.g., TR-BBB, RBE4), and human cells (e.g., hCMEC/D3, HMEC-1). Recently, 3D models of the BBB have been developed to replicate the in vivo conditions with the BBB endothelial cells surrounded by pericytes, astrocytes, and extracellular matrix proteins. These 3D models have very tight intercellular junctions with TEER values as high as 1650 ohm/cm², similar to values found in vivo (Lippmann et al. 2013; Naik and Cucullo 2012; Weksler et al. 2013).

Initially, in situ rat brain perfusion developed by Takasato et al. was used to study the transport of radioactive-labeled (i.e., ³H and ¹⁴C) peptides and proteins across the BBB (Takasato et al. 1984; Kiptoo et al. 2011). In this study, the rats undergo surgery, under anesthesia, where a polyethylene catheter filled with heparinized saline is ligated to the left common carotid artery (LCCA). The radioactive-labeled peptide is infused through the LCCA using in perfusate solution delivered by syringe pump immediately after a heart vessel is cut to sacrifice the anesthetized animal. The level of radioactivity in the brain is counted using a scintillation counter to calculate the concentration of peptides delivered to the brain. Brain extraction and LC-MS/MS methods have been developed as alternative techniques to detect deposition of unlabeled peptide in the brain (Ulapane et al. 2017).

Several in vivo methods are also being used to study transport of peptides across the BBB. Intravenous (i.v.) administration via the tail vein followed by detection of peptide deposition in the brain is normally used to study the delivery of molecules across the BBB in animal models. Alternatively, intracarotid artery administration has been used to deliver the peptides, because it has immediate access to the brain vasculature. Recently, IRdye-800 CW-labeled molecules (i.e., peptides and proteins) have been used to determine peptide brain delivery that can be detected using

near IR fluorescence (NIRF) imaging (Ulapane et al. 2017, 2019a, b). The advantage of using NIRF for imaging the deposition of molecules in the brain is that NIRF imaging has a low background interference from the tissues to the emission of light from the analyte molecules. This method can quantitatively determine the amounts of peptides deposited in the brain as well as be sensitive and convenient to study delivery of peptides and proteins into the brain (Ulapane et al. 2017, 2019a, b). Peptides or proteins conjugated to gadopentetic acid (Gd-DTPA) have also been used to study peptide and protein depositions in the brain of living animal in qualitative and quantitative manners (Ulapane et al. 2017; On et al. 2014; Tabanor et al. 2016). The advantage of MRI is that the deposition of the delivered molecule in different sections of the brain can be determined in living animal. In contrast, quantitative determinations of depositions in different brain sections using NIRF have to be done using the dissected and homogenized brain sections. In addition, the access to an MRI instrument may be more limited to majority of researchers compared to that of an NIRF imaging instrument.

7.3 Passive Diffusion Across the BBB via Transcellular Pathway

One way that peptides could cross the BBB endothelium is via the transcellular pathway in which the peptides from the blood partition into luminal cell membranes on the blood side followed by entering the cytoplasm (Fig. 7.1; Path A). From the cytoplasm, the peptide needs to cross the abluminal cell membranes on the brain side to enter the brain extracellular fluid. Normally, molecules with physicochemical properties that follow Lipinski's Rule of Five can diffuse passively via the transcellular pathway of the BBB. To follow the Rule of Five, a peptide molecule should have (a) cLogP lower than 5, (b) MW less than 500 g/mol, (c) less than 5 H-bond donors (e.g., NH, OH), and (d) less than 10 H-bond acceptors (e.g., O and N) (Lipinski et al. 2001; Lipinski 2004, 2016). Most peptides (i.e., hexapeptide or larger) have high hydrophilicity and MW higher than 500 g/mol with more than 5 H-bond donors and 10 H-bond acceptors as well. Thus, these peptides have difficulty in effectively crossing the BBB via passive diffusion. Some peptides and other hydrophilic molecules that have their own transporters can cross the BBB through transcellular pathway into the brain (Fig. 7.1; Path B). For example, the surface of the BBB has transporters for glucose, amino acid, and di-/tri-peptides to carry them from the blood into the brain. These transporters have also been exploited to carry drugs into the brain.

There are molecules with physicochemical properties that follow Lipinski's Rule of Five but still cannot cross the BBB. Molecules belonging to this category are recognized by the efflux pumps such as P-glycoprotein (Pgp) and multidrug resistance (MDR) that expel them from the luminal membranes to prevent them from crossing the BBB (Fig. 7.1; Path C). For example, the anticancer drug daunomycin

has physicochemical properties that are conducive for passive diffusion across the BBB. However, ^3H -daunomycin could not effectively cross the BBB when studied using the *in situ* rat brain perfusion method (Kiptoo et al. 2011). This is because ^3H -daunomycin is a substrate for efflux pumps such as P-glycoprotein (Pgp). The recognition of ^3H -daunomycin by Pgp can be determined by competition studies using verapamil as an inhibitor of Pgp. Delivering ^3H -daunomycin in the presence of verapamil enhanced the brain deposition of ^3H -daunomycin (Kiptoo et al. 2011). Thus, ^3H -daunomycin is also substrate for Pgp in the BBB (Kiptoo et al. 2011). Some efforts have been made to utilize Pgp inhibitors to improve brain delivery of drugs.

It has been shown that some peptides are substrates for Pgp or MDR on the BBB. Using cell culture models of biological barriers (i.e., intestinal mucosa and BBB), the efflux substrate activity for peptides can be evaluated (Fig. 7.1; Path C) (Ouyang et al. 2002, 2009a, b). If a peptide is a substrate for efflux pumps, the apparent permeability of the substrate of apical (AP) side-to-basolateral (BL) side (AP-to-BL) will be lower than that of BL-to-AP (Ouyang et al. 2009a; On and Miller 2014). This is due to the high presence of efflux pumps on the AP surface of BBB endothelial cells. In other words, the apparent permeability of BL-to-AP is normally larger than that of AP-to-BL. Theoretically, without the effects of the efflux pumps, the passive diffusion of BL-to-AP is the same as AP-to-BL (the ratio of BL-to-AP/AP-to-BL is equal to one). As mentioned previously, efflux pump inhibitors such as cyclosporine A and GF120918 can be used to determine whether the peptide is recognized by efflux pumps. In this case, in the presence of cyclosporine A or GF 120918, the AP-to-BL transport across the BBB of the peptide under study increases (On et al. 2014; On and Miller 2014).

The formation of cyclic peptides from the parent linear peptides improves the peptide passive transport by two- to threefold across Caco-2 cell monolayers via the transcellular pathway in the absence of efflux pump activity (Fig. 7.1; Path A) (Okumu et al. 1997). To enhance delivery of peptides, a cyclic peptide prodrug using acyloxyalkoxy promoity was synthesized to improve the delivery of delta-sleep-inducing peptide (DSIP: H-Trp-Ala-Gly-Gly-Asp-Ala-OH; Fig. 7.2a) (Pauletti et al. 1996). The acyloxyalkoxy cyclic peptide prodrug of DSIP (AOA-DSIP; Fig. 7.2b) has significantly better transport across the Caco-2 cell monolayers than that of the parent linear peptide (Fig. 7.2a) (Pauletti et al. 1996). This indicates that the formation of the cyclic prodrug can enhance the passive diffusion across the cell membranes of Caco-2 cell monolayers. The cyclic prodrug can also be converted to the parent linear peptide by esterase.

The same cyclic peptide prodrug method was also applied to the delta opioid peptide called (D-Ala 2, D-Leu 5) enkephalin (DADLE: H-Tyr-D-Ala-Gly-Phe-D-Leu-OH, Fig. 7.2c). As observed using a model hexapeptide, it was hypothesized that cyclization of DADLE as prodrugs would improve the passive diffusion across the BBB compared to parent linear DADLE peptide. The formation of cyclic peptide prodrugs would lower the hydrogen bonding potential and enhance partitioning to cell membranes. Thus, several different cyclic prodrugs of DADLE were synthesized, including acyloxyalkoxy-based cyclic prodrug of DADLE (AOA-DADLE;

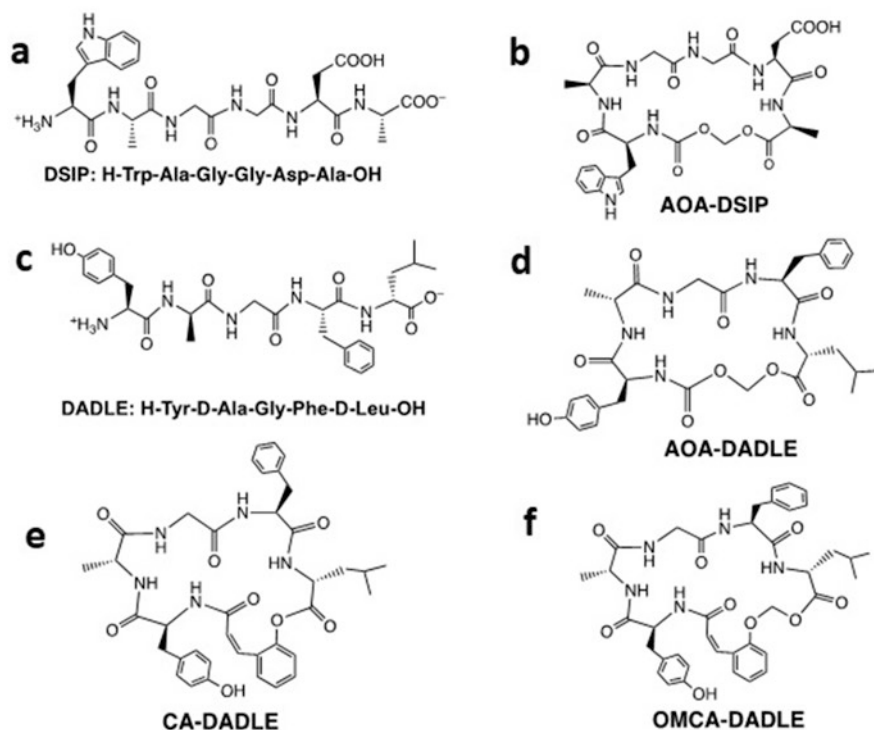


Fig. 7.2 (a–f) Structure of DSIP and DADLE peptide and their cyclic prodrugs. The cyclic peptide prodrugs (**b** and **d–f**) were synthesized to improve permeation of the peptide across the biological barriers such as the BBB and intestinal mucosa barrier by changing their physicochemical properties to become conducive the partition to cell membranes of the biological barriers. The cyclic peptide prodrugs were constructed using various esterase-sensitive promoieties such as acyloxyalkoxy (AOA, **b** and **d**), coumarinic acid (CA, **e**), and oxymethyl-modified coumarinic acid (OMCA, **f**) promoiety. The peptide prodrugs can be converted to parent peptides by esterase in the blood and brain

Fig. 7.2d), coumarinic acid-based cyclic prodrug of DADLE (CA-DADLE; Fig. 7.2e), and oxymethyl-modified coumarinic acid-based cyclic prodrug of DADLE (OMCA-DADLE; Fig. 7.2f) (Ouyang et al. 2002, 2009a). Although the membrane partition and passive diffusion of these cyclic peptide prodrugs were improved compared to parent linear peptides, the cyclic peptide prodrugs of DADLE are substrates of Pgp efflux pump. Therefore, the overall penetration across biological barriers was impeded by Pgp (Fig. 7.1; Path C). The recognition of cyclic prodrug by efflux pump was determined using various cell monolayers such as Caco-2, MDCK-WT, MDCK-MDR1 (Pgp), and MDCK-MRP2 cells. MDCK-WT, MDCK-MDR1, and MDCK-MRP2 cell monolayers have been used as alternative models for the BBB. MDCK-MDR1 and MDCK-MRP2 cells have high expression of MDR1 and MRP2 efflux pumps on their surface, respectively, to determine cyclic peptide prodrug recognition by the efflux pumps. In general, the data showed that

AOA-DADLE, CA-DADLE, and OMCA-DADLE have higher BL-to-AP transport than AP-to-BL transport, indicating that all three cyclic peptides are substrates for efflux pumps (Ouyang et al. 2002, 2009a). The ratios of BL-to-AP/AP-to-BL for AOA-DADLE, CA-DADLE, and OMCA-DADLE were 16, 35, and 35, respectively (Ouyang et al. 2002, 2009a). Transport studies of AOA-DADLE, CA-DADLE, and OMCA-DADLE were also done in the presence of efflux pump substrate (Pgp inhibitors) such as GF 120918, cyclosporin A, and PSC-833. The data indicated that the BL-to-AP/AP-to-BL ratios for AOA-DADLE, CA-DADLE, and OMCA-DADLE when delivered with GF 120918 were lowered to 0.5, 1.02, and 1.48, respectively (Ouyang et al. 2002, 2009a). This is an indication that all three cyclic prodrugs are substrates for efflux pumps. Using *in situ* rat brain perfusion method, the GF-120918 enhanced the delivery of CA-DADLE into the brain by 460-fold compared to that of CA-DADLE alone (Ouyang et al. 2009a). This is another confirmation that CA-DADLE is a substrate for Pgp efflux pump on the rat BBB. Unlike small molecules in which structural modification can avoid the efflux pump recognition to improve passive permeability, the structural change in peptides may not significantly improve their passive diffusion because the changes cannot overcome the efflux pump activity.

7.4 Receptor-Mediated Transcytosis of Peptides Through the BBB

The BBB is decorated with many transporters and receptors that can carry molecules from the blood stream into the brain (Fig. 7.1; Path B). These transporters include glucose, amino acid, and di-/tripeptide transporters. Glucose transporter, GLUT-1, is a saturable and efficient transporter to provide a glucose for the brain to function properly. This transporter facilitates glucose diffusion to create balance from the higher to the lower concentration between the blood and the brain; thus, any excess and unutilized glucose in the brain is transported back into the blood (Banks et al. 2012). The brain metabolic rate can be measured by determining the rate of glucose uptake (Banks et al. 2012). GLUT-1 has also been exploited for drug delivery into the brain; however, the success of this method has not yet been realized.

Similarly, there are many protein transporters on the surface of the BBB endothelial cells, including transporters for insulin (Frank and Pardridge 1981; Frank et al. 1986), insulin-like growth factor 1 (IGF-1) (Frank et al. 1986), and transferrin (Tf) (Fishman et al. 1987; Visser et al. 2004) melanotransferrin (p97) (Demeule et al. 2002), apolipoproteins (Apo) A and E (Herz and Marschang 2003), leptin (Banks and Farrell 2003), immunoglobulin G (Zlokovic et al. 1990), tumor necrosis factor alpha (TNF- α) (Pan and Kastin 2002), epidermal growth factor (Pan and Kastin 1999), and interleukin (Banks et al. 2001). As an example, insulin is transported by insulin receptor across the BBB into the brain via a saturable mechanism (Banks 2004). The insulin transporter regulates the balance between insulin in the blood

and the brain. The balance is also affected by glucose in which during a hyperglycemic condition, the transport of insulin across the BBB is increased.

The use of receptor-mediated transport system normally relies on a higher expression of receptors on the luminal side of the BBB compared to those in other organs. Therefore, the ligand is targeted to the BBB endothelial cell rather than other organs (Oller-Salvia et al. 2016; Broadwell et al. 1988; Frank et al. 1986). Thus, protein or peptide ligands can be conjugated with drugs to make protein-drug or peptide-drug conjugates for targeting drugs into the brain. In some cases, the receptor-mediated uptake of the conjugate can be inhibited by the endogenous ligand (e.g., peptide, protein), which lowers transport effectiveness of the conjugate to cross the BBB (Oller-Salvia et al. 2016). To avoid competition between the conjugate and endogenous ligand for the receptor, monoclonal antibodies (mAbs) or peptide ligands that bind to the transport receptor at a different site from that of endogenous ligand were developed. In this case, the mAb or the targeting peptide can avoid interrupting the receptor function to bind its endogenous ligand.

Several mAbs have been developed as antibody-drug conjugates (ADCs) to carry drugs across the BBB into the brain. MAbs to transferrin and insulin receptors have been developed as ADCs to deliver drugs into the brain. Transferrin receptor (TfR), found abundantly on the BBB, transfers iron into the cells by transporting iron-bound transferrin (Tf) proteins. Many clinical developments of TfR mAbs as ADC therapeutics have been carried out for brain diseases, where the major challenge was to increase transcytosis efficiency of ADC into the brain (Paterson and Webster 2016). Normally, the uptake of TfR ADC from the luminal side into the BBB endothelial cells is effective; however, the translocation of TfR ADC from the endothelial cells into the brain can be inefficient. This is due to trapping of TfR ADC in the endosomes of endothelial cells; furthermore, tight binding of TfR mAb to TfR causes a high degradation of TfR ADC in lysosomes of endothelial cells. It has been shown that the higher the affinity of the TfR mAb to TfR, the higher the degradation of ADC in lysosomes. This limits the transcytosis of ADC into the abluminal region. Due to the mAb tight binding to TfR, the release of TfR ADC from TfR receptor at the abluminal side of endothelial cells into the brain could also be inefficient. Overall, the TfR ADC released into the brain fluid to diffuse throughout the brain is lower than the uptake from the blood. There have been some efforts to lower the binding affinity of TfR mAb to TfR to improve the transcytosis efficiency of TfR ADC from the blood into the brain. Recently, a phage display method was used to discover new peptide ligands for TfR that can be used for drug delivery across the BBB (Oller-Salvia et al. 2016).

The brain transport mechanisms of opioid peptides such as D-Phe-Cys-Tyr-D-Trp-Arg-Thr-Pen-Thr-NH₂ (CTAP), [D-penicillamin^{2,5}] enkephalin (DPDPE), and biphalin (Tyr-D-Ala-Gly-Phe-NH₂) were compared using in situ rat brain perfusion method (Egleton et al. 1998). It was found that CTAP penetrated the BBB via passive diffusion, while DPDPE crossed the BBB through a combination of passive diffusion and saturable transport mechanism. Finally, biphalin crossed the BBB via passive diffusion as well as utilizing large neutral amino acid carrier (Egleton et al. 1998). Both DPDPE and biphalin showed time-dependent linear uptake, and both

had higher brain uptake than sucrose, while CTAP had higher uptake than inulin (Egleton et al. 1998). To test whether all three peptides were transported by receptor-mediated process, the uptake of radioactive-labeled peptide was inhibited by unlabeled peptide. Addition of 100 μ M non-radioactive DPDPE and biphalin peptides inhibited the uptake of both 3 H-DPDPE and I^{125} -biphalin, respectively, suggesting both peptides were transported by receptor-mediated process. In contrast, the uptake of 3 H-CTAP was not inhibited by non-radioactive CTAP, indicating that the BBB penetration of CTAP was not mediated by receptor (Egleton et al. 1998).

To ensure that the measured peptide radioactivity was due to deposition in the brain and not from the trapping of the peptide in the BBB endothelial cells, a capillary depletion experiment was carried out. The capillary depletion was done to remove the BBB vascular endothelial cells from the brain homogenates. The capillary depletion study showed that only 1% of biphalin and 10% of DPDPE were in the BBB vascular endothelial cells, while the rest of peptide resided in the brain as measured by radioactivity. These data indicated that majority of brain radioactivity came from biphalin and DPDPE that were deposited in the brain parenchyma not in the vasculature (Egleton et al. 1998). In contrast, 50% of CTAP resided in the vasculature, and 50% was distributed in the brain, indicating a high amount of the CTAP was trapped in the BBB vascular endothelium (Egleton et al. 1998). The uptake of both 3 H-DPDPE and 125 I-biphalin in the rat brain was partially saturable, and they followed Michaelis-Menten kinetics (Egleton et al. 1998). Therefore, it is critical to perform capillary depletion studies to ensure that the detected molecules are in the brain not in the BBB endothelial vasculatures.

The brain uptake of CTAP was studied using 3 H-CTAP and compared to 14 C-inulin and 3 H-morphine using the in situ brain perfusion method (Abbruscato et al. 1997). 14 C-inulin is a paracellular marker, and 3 H-morphine is a μ -opioid receptor agonist. The radioactivity of 3 H-CTAP was six times higher than that of 14 C-inulin in the brain and cerebral spinal fluid (CSF) at different time points. However, the brain deposition of 3 H-CTAP was lower than 3 H-morphine, due to morphine's higher hydrophobicity than CTAP leading to more passive diffusion across the BBB (Abbruscato et al. 1997). CTAP has reasonable stability in the serum and blood with $t_{1/2} > 500$ min. Presumably, CTAP's high binding affinity to albumin contributes to CTAP's plasma stability.

Ghrelin, a 28-mer peptide, can cross the BBB with unelucidated mechanism (van der Lely et al. 2004; Rhea et al. 2018). The BBB transport of ghrelin was proposed to be mediated by the growth hormone secretagogue receptor (GHSR); however, GHSR was not the only receptor that can transport ghrelin into the brain (Rhea et al. 2018). Human and mouse desacyl ghrelins (DAGs) were transported across the BBB faster than acyl ghrelin (AG), indicating that acylation levels of ghrelin influence the brain uptake. Injection of ghrelin peptides via i.v. administration led to a measurable influx into the brain from the blood stream. Human ghrelin peptides are usually transported more efficiently across the BBB than mouse ghrelin; their transport rates are in the following order hDAG > mDAG > hAG > mAG (Rhea et al. 2018). The highest level of ghrelin was found in the olfactory bulb, regardless of their structures (i.e., acylated or non-acylated). Human DAG peptide has the fastest

influx rate into the brain, while mouse AG peptide has the slowest influx rate in both wild-type (WT) and GHSR null mice (Rhea et al. 2018). There was no significant difference of ghrelin influx rate between WT and GHSR null mice, indicating that GHSR was not the main transport receptor (Rhea et al. 2018). Using ^{125}I -hDAG and acid precipitation, hDAG was found to be intact in serum and the brain after crossing the BBB. The brain uptake of ^{125}I -hDAG was not saturable, suggesting the transport was not a receptor-mediated process. After the uptake, ghrelin was distributed throughout the brain with high levels in the olfactory bulb and pons-medulla (Rhea et al. 2018).

Cell-penetrating peptides (CPPs), such as transactivator of transcription (TAT) peptides from fragments of HIV TAT protein, have been shown to undergo adsorptive-mediated transcytosis (AMT) across the BBB (Oller-Salvia et al. 2016; Herve et al. 2008; Green and Loewenstein 1988; Frankel and Pabo 1988). There are over 100 CPPs, from 5-mer to 40-mer peptide, that have been investigated over several decades (Lindgren and Langel 2011). CPPs have been investigated to carry drugs across the BBB into the brain (Oller-Salvia et al. 2016). TAT peptide was used to delivery β -galactosidase into the brain by aiding the transcellular passage of the enzyme across the BBB (Oller-Salvia et al. 2016; Herve et al. 2008). Unlike receptor-mediated process, the cellular uptake of CPP is unsaturable; cells from organs other than BBB endothelial cells also can engulf CPP. In general, cellular uptake of CPPs was not selective to a particular group of cell type, and their exocytosis from the BBB endothelium into the brain needs further investigation. There is a possibility that CPPs and their cargo could be trapped in the BBB endothelium without entering the brain (Oller-Salvia et al. 2016). In summary, BBB selectivity and transcytosis effectiveness of CPPs across the BBB need further investigation.

7.5 Peptide Conjugates for Delivering Drugs Across the BBB

Several peptides have been found to deliver drugs across the BBB via receptor-mediated process. Angiopep-2 peptide (ANG, Table 7.1) was found to undergo receptor-mediated transcytosis across the BBB endothelial microvessel cells (Li et al. 2016). This peptide was transported by low-density lipoprotein-1 (LRP1) receptors on the surface of the BBB endothelial cells (Li et al. 2016). Conjugation of ANG peptide with β -secretase inhibitor (SI) produced ANG-SI peptide (Table 7.1). ANG-SI inhibited the production of amyloid-beta ($\text{A}\beta$) in neuronal cells, demonstrating that ANG peptide was responsible for the internalization of SI peptide that led to SI peptide activity (Kim et al. 2016).

A combination of ANG and TAT (Table 7.1) peptides were used to deliver paclitaxel (PTX) by forming ANG-TAT-PTX conjugate that treats U87 glioblastoma brain tumors in mice (Li et al. 2016). Mice treated with ANG-TAT-PTX conjugate had higher survival rate than those treated with ANG-PTX conjugate and PTX alone (Li et al. 2016). The results signify that ANG-TAT combination improved targeting

Table 7.1 Peptide sequences for drug delivery

Peptide name	Sequence
ANG	TFFYGGSRGKRNNFKTEEY
ANG-SI	TFFYGGSRGKRNNFK-EVN-sta-VAEF
ANG-TAT	TFFYGGSRGKRNNFK- TEEYGRKKRRQRRRPPQQ
gHo	NHQQQNPHQPPM-NH ₂
pVEC	LLILRRRIRKQAHAAHSK-NH ₂
GKRK	GKRK
RVG29	YTIWMPENPRPGTPCDIFTNSRGKRASNGC
C2	CDIFTNSRGKRA
C2-9r	CDIFTNSRGKRAGGGGrrrrrrrrr
RI-C2	Arkgrsntfide
RI-C2-9r	arkgrsntfideGGGGrrrrrrrrr

of PTX into the brain and brain tumor. Presumably, the brain uptake of ANG-TAT-PTX conjugate was due to recognition of ANG peptide by LRP-1 receptors as well as adsorptive-mediated transcytosis (AMT) of TAT peptide by the BBB endothelial cells. In summary, the combination of ANG and TAT peptide is better than ANG peptide alone in delivering PTX to the brain and brain tumor cells.

Glioma-homing (gHo) peptide (Table 7.1) was discovered using phage display method, because it binds to U251 glioma cells (Eriste et al. 2013). gHo peptide was conjugated to a CPP called pVEC peptide (Table 7.1) to make pVEC-gHo peptide. The combined peptide delivered 5(6)-carboxy fluorescein (FAM) and doxorubicin (Dox) by conjugating the drug to its N-terminus to produce FAM-pVEC-gHo and Dox-pVEC-gHo conjugates, respectively. FAM-pVEC-gHo selectively bound to U87 tumor cells but not HeLa and HEK even at very high concentration as determined by confocal microscopy; this signifies the selectivity of the peptide combination (Eriste et al. 2013). The gHo peptide directs the conjugate (i.e., FAM-pVEC-gHo or Dox-pVEC-gHo) to glioma cells, while pVEC has a role to improve uptake via AMT for transcytosis across the BBB (Eriste et al. 2013). Administration of FAM-pVEC-gHo conjugate to animal with subcutaneous U87 tumors showed deposition of the conjugate in tumor but not in other tissues such as the brain, kidney, and liver. Dox-pVEC-gHo conjugate was used to treat animals with subcutaneous U87 tumor; the tumors in treated mice were significantly smaller than those in untreated group. Unfortunately, both free doxorubicin and Dox-pVEC-gHo derivative did not prolong the survival of animal with intracranial gliomas, indicating that the conjugate did not cross the BBB effectively (Eriste et al. 2013). Further studies are still required to assess whether Dox-pVEC-gHo can cross the BBB to improve the outcome of treating intracranial tumors in mice (Eriste et al. 2013).

7.6 Brain Drug Delivery Using Nanoparticles

Nanoparticles have been investigated to deliver small molecule drugs, peptides, proteins, and oligonucleotides into the brain. A VCR-GKRK-APO nanocage delivery system was developed using a dual targeting system with GKRK peptide and apoferritin (APO) that encapsulates vincristine (VCR) as anticancer drug (Zhai et al. 2018). This nanocage was used to treat glioma brain tumor in the animal model. In this nanocage, GKRK peptide targets heparan sulfate proteoglycan (*HSPG*), while APO is used to target transferrin receptor-1 (TfR1) in the BBB endothelial and tumor cells. HUVEC and U87MG cells internalized fluorescence-labeled GKRK-APO significantly better than that of APO alone, suggesting that a dual targeting system was better than the single targeting system (Zhai et al. 2018). It was also shown that drug-free GKRK-APO and APO were not toxic to HUVEC and U87MG cells as determined by MTT assay, indicating the nontoxic nature of the GKRK-APO drug carrier (Zhai et al. 2018).

The uptake of GKRK-APO nanocages by tumor cells was followed using noninvasive NIR fluorescence (NIRF) imaging and immunofluorescence methods. After administration, a higher accumulation of GKRK-APO nanocages found in glioma brain tumor cells implanted in animals compared to APO nanocages. This demonstrates that GKRK-APO nanocage can cross the BBB to target the tumor cells. Compared to free vincristine alone, vincristine-loaded GKRK-APO has enhanced cytotoxicity against U87MG tumor cells. In addition, it can cross the bEnd3 cell monolayer BBB in vitro model (Zhai et al. 2018). In vivo, the brain deposition of vincristine was 6.5-fold higher when delivered using the GKRK nanocage compared to that of free vincristine (Zhai et al. 2018). The blood clearance of VCR-GKRK-APO was also slower than that of vincristine alone. The animals treated with the VCR-GKRK-APO nanocage have smaller glioma tumor diameter compared to controls as determined by MRI (Zhai et al. 2018). Histology studies indicated no significant damage to the liver, kidney, brain, lung, and heart cells after administration of VCR-GKRK-APO nanocage. The histology studies implied that vincristine delivery with nanocage was not toxic for other non-targeted organs. In summary, nanocages with dual targeting moieties enhanced vincristine brain delivery and selectivity in the animal models.

Another example of improved nanoparticle BBB delivery through peptide targeting involves the rabies virus glycoprotein (RVG) (Kumar et al. 2007; Liu et al. 2009; Javed et al. 2016). A 29-amino acid peptide from rabies virus glycoprotein (RVG) was conjugated via a polyethylene glycol (PEG) linker to polyamidoamine dendrimers (PAMAM) to produce PAMAM-PEG-RVG29 nanoparticles (NPs) for delivering genes into the brain (Liu et al. 2009). Recently, RVG peptide was proposed to bind GABA_B receptor where it was previously proposed to bind nicotinic acetylcholine receptor (nAChR) (Liu et al. 2009). Fluorescently labeled NPs were engulfed by brain capillary endothelial cells (BCEC). To test whether the cellular uptake was via a receptor-mediated process, the NPs were first incubated on BCECs at 4 °C and 37 °C. Lower cellular uptake of NPs was observed at 4 °C compared to

37 °C incubation. Suppression of NP internalization activity at 4 °C implies that the internalization is due to receptor-mediated uptake process. Preincubation of cells with free RVG29 peptide inhibited the uptake of PAMAM-PEG-RVG29 NPs, supporting the idea that the uptake of NPs by BCEC was mediated by receptors of RVG29 peptide. The uptake of NPs was also inhibited by GABA, verifying that the internalization of RVG2-studded NPs was mediated by GABA_B receptors through clathrin-mediated endocytosis (Liu et al. 2009). In contrast, acetylcholine, nicotine, or mecamylamine did not inhibit the uptake of PAMAM-PEG-RVG29 NPs, showing that the nicotinic acetylcholine receptor (nAChR) was not the internalization receptor of the NPs (Liu et al. 2009).

DNA was incorporated into PAMAM-PEG-RVG29 NPs via charge-charge interactions. PAMAM modification with RVG29 peptide did not influence the DNA encapsulation properties compared to unmodified PAMAM. PAMAM-PEG-RVG29 NPs were effective in delivering DNA into cell cytoplasm after incubation for 15 and 60 min, and more DNA was found in the cytoplasm than in the nucleus (Liu et al. 2009). After i.v. administration of fluorescent-labeled PAMAM-PEG-RVG29/DNA NPs, the RVG29 containing NPs accumulated in the brain with a significantly higher amount than that of fluorescent-labeled PAMAM/DNA NPs. The results confirmed the important role of RVG29 peptide in the cellular uptake of NPs (Liu et al. 2009). In addition, delivery of PAMAM-PEG-RVG29/pGL2 increased the transfection efficiency of luciferase in the brains of mice compared to those injected with control PAMAM/pGL2 (Liu et al. 2009).

RVG peptide was also used to deliver siRNA across the BBB for silencing neuronal gene activity (Kumar et al. 2007). RVG peptide bound to green fluorescent protein (GFP)-encoding lentiviral vector could specifically transfect neuronal Neuro-2a cells but not HeLa cells *in vitro*, indicating cell selective delivery of GFP gene by RVG peptide. RVG peptide bound to anti-GFP siRNA can silence the GFP gene in Neuro-2a cells (Kumar et al. 2007). Similarly, GFP transgenic mice dosed with RVG peptide linked to anti-GFP-siRNA significantly suppressed GFP expression compared to control. There were no significant differences in GFP expression in all other organs compared to control, confirming that RVG directed anti-GFP-siRNA into the brain (Kumar et al. 2007). RVG peptide was also used to target antiviral FvE hairpin RNA in mice infected with Japanese encephalitis virus (JEV). The results showed an increase in survival of infected mice by 80% compared to untreated control mice; the untreated mice all died within 10 days (Javed et al. 2016). These results indicated that RVG peptide transported siRNA across the BBB to silence the gene for JEV replication (Javed et al. 2016).

C2 decapeptide (Table 7.1) was derived from the RVG29 peptide, where it can bind to neuronal cells as well as be transported through the BBB into the brain (Javed et al. 2016). C2 peptide and retro-inverso C2 peptide (RI-C2, Table 7.1) have been utilized to deliver siRNA into the brain. RI-C2 has a reversed sequence of C2 with D-amino acids. M17 cells can uptake complexes of C2-9r and RI-C2-9r peptides with rhodamine-labeled siRNA into the cytoplasm. C2-9r and RI-C2-9r peptides have nine D-arginine residues added to C2 and RI-C2 spaced by four Gly residues to bind siRNA via opposite charge interactions to make C2-9r-siRNA and

RI-C2-9r-siRNA complexes (Javed et al. 2016). Delivery of C2-9r-siRNA complex could knockdown 60–90% of the α -synuclein protein expression. RI-C2-9r-siRNA behaved the same way as the parent C2-9r-siRNA; however, siRNA in the RI-C2-9r-siRNA had longer plasma stability than that in the parent C2-9r-siRNA formulation (Javed et al. 2016). Delivery of C2-9r-siRNA and RI-C2-9r-siRNA via i.v. administration into mice could knockdown α -synuclein levels for up to 72 h in different parts of the brain without eliciting an immune response (Javed et al. 2016). The downregulation of α -synuclein level is protected against neurodegeneration and pathological symptoms in the 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) mouse model of Parkinson's disease (Javed et al. 2016). The treated mice also showed improvement in rotarod behavioral tests with decreased dopaminergic neuronal loss compared to untreated control mice (Javed et al. 2016). The results demonstrated that C2-9r and RI-C2-9r peptides can effectively deliver siRNA into the brain.

Although RVG29 peptide and its derivatives can cross the BBB, further studies will be needed to improve the transport capacity and selectivity to carry molecules across the BBB. In addition, the mechanism of transport of RVG29 and its derivatives is still unclear because some data suggest that the uptake of RVG29 is due to nicotinic acetylcholine or GABA_B receptors rather than RVG receptor. Thus, more studies are needed to elucidate the transport mechanism of RVG29 and its derivatives.

7.7 Modulation of the BBB to Improve Delivery via Paracellular Pathways

Ions and small hydrophilic molecules can penetrate through the BBB via the paracellular pathways (Fig. 7.1; Path D). The paracellular pathways only allow molecules with hydrodynamic radius of <11 Å to cross the BBB because the tight junctions have small pores that limit passive diffusion of many medium (e.g., peptides) or large molecules (e.g., proteins). The BBB intercellular junctions consist of three different sections. First, the tight junction section is found in the luminal side and forms the closest membrane-membrane contact between opposing cells. The tight junctions, normally referred to as the “Kiss” region, provide the most restricted passage way. The glue between the opposing membranes is constructed by protein-protein interactions of occludins, claudins, and junctional adhesion molecules (JAMs). The second section underneath the tight junction is *adherens* junction connected by cell-cell adhesion proteins such as VE-cadherins and nectins. The final section is the desmosome that is constructed by desmocollin and desmoglein interactions where desmocollins and desmogleins are in the classic cadherin family.

Many efforts have been investigated to improve delivery of peptides and proteins into the brain via modulation of paracellular pathways. Most of these methods were aimed at increasing the porosity of the intercellular junctions to enhance passive

diffusion of peptides and proteins from blood to brain via the paracellular pathway (Fig. 7.1; Path D). The most successful method to deliver drugs to brain tumor patients is the blood-brain barrier disruption (BBBD) method utilizing hyperosmotic solution. Several additional methods have been developed to selectively modulate the protein-protein interactions in the intercellular junctions of the BBB using cell adhesion peptides.

7.7.1 Osmotic Blood-Brain Barrier Disruption (BBBD) Method

The BBBD, or osmotic brain delivery method, has been successfully used to deliver anticancer drugs to brain tumor patients. This method utilizes hypertonic mannitol solution to disrupt the BBB to increase the porosities of the BBB intercellular junctions (Neuwelt et al. 1984a, b; c, 1985; Doolittle et al. 2014). In this case, administration of hypertonic solution via internal carotid artery (ICA) shrinks the vascular endothelial cells of the BBB to create large pores in the BBB paracellular pathways to allow passive diffusion of drug molecules (e.g., small drugs and proteins) through the paracellular pathways into the brain (Fig. 7.1; Path D). One caution is that prolonged opening of the BBB by osmotic method may cause brain inflammation and epilepsy (Luo and Shusta 2020; Marchi et al. 2007). Besides modulation of the intercellular junctions, osmotic brain delivery method has also been suggested to induce vesicular transport as well as the presence of fenestrations.

In preclinical studies, the BBBD method was used to infuse radioactive-labeled methotrexate (MTX) into rat with tumor on the right hemisphere at a dose of 4000 ng/g body weight (Neuwelt et al. 1984a). The brain depositions of MTX were compared after infusions of MTX with the same dose without hypertonic mannitol and with hypertonic mannitol into the right ICA of rats. The data showed that infusion of MTX with hypertonic mannitol into the right ICA has higher depositions of MTX in tumor (5×), tissue surrounding tumor in the right hemisphere (2.9×), and brain tissue distant to tumor (10×) compared to infusion of MTX alone (Neuwelt et al. 1984a). This suggests that mannitol disrupts the BBB to allow permeation of MTX into the brain and tumor. However, there was no difference in MTX depositions in contralateral left hemisphere of the brain when MTX was delivered with and without mannitol. The results indicate that infusing MTX with mannitol to the right ICA did not effectively enhance delivery to the left hemisphere of the brain. In other words, the MTX did not effectively diffuse from the right to the left hemisphere of the brain. When MTX was infused to the left ICA with hypertonic mannitol, the deposition of MTX on the left hemisphere was higher than MTX found in tumor on the right hemisphere, area surrounding the tumor, or right brain distant from the tumor. These results indicate that hypertonic solution can enhance the delivery of anticancer drug MTX into the brain and tumor. In addition, the ICA site

chosen for administering the drug can influence the deposition of the drug in the targeted hemisphere of the brain.

7.7.2 *Blood-Brain Barrier Modulators (BBBMs) of the Intercellular Junction Proteins*

Recently, focused ultrasound (FUS) has been developed to modulate the intercellular junctions of the BBB transiently to improve drug permeation from the blood into the brain. The effects of FUS in delivering microbubbles containing drugs across the BBB have been observed using MRI (Burgess et al. 2015). In preclinical study, FUS delivery of etoposide to glioblastoma in mice increased brain tumor-to-blood ratio by 3.5-fold, and it prolonged animal survival and decreased tumor growth by 45% (Wei et al. 2021). Drugs with various sizes and doses in different sizes of microbubbles can be effectively delivered to the brain using FUS with different frequency and repetition of ultrasound pulses in preclinical studies (Burgess et al. 2015). It is envisioned that the noninvasive nature of FUS in combination with MRI monitoring is a clear advantage of this method over invasive brain delivery methods (e.g., intracerebroventricular injection). However, the repeatability of FUS in the clinical setting is still unclear, and the side effects from off-target delivery of FUS to brain tissues require full and extensive investigations.

Inspired by the BBBD method using hypertonic mannitol, a new method to modulate the BBB to increase porosity of the paracellular pathway was developed by modulating the cell-cell adhesion molecules connecting the two opposing cell membranes (Lutz and Siahaan 1997b). The idea is that inhibiting cell-cell adhesion molecules can increase the BBB paracellular pathway porosity and enhance the paracellular permeation of molecules across the BBB. Inhibition of

Table 7.2 Sequences of BBBM: HAV and ADT peptides

Peptide name	Sequence
<i>HAV peptides</i>	
HAV6	Ac-SHAVSS-NH ₂
HAV4	Ac-SHAVAS-NH ₂
cHAVc1	Cyclo(1,8)Ac-CSHAVASC-NH ₂
cHAVc3	Cyclo(1,6)Ac-CSHAVC-NH ₂
HAVscr	Ac-HSVSAS-NH ₂
<i>ADT peptides</i>	
ADT6	Ac-ADTPPV-NH ₂
ADTC1	Cycloid(1,7)Ac-CADTPPVC-NH ₂
ADTC5	Cyclo(1,7)Ac-CDTPPVC-NH ₂

cadherin-cadherin interactions by cadherin peptides was first investigated to increase porosity of the BBB paracellular pathway in an equilibrium and reversible fashion. In this case, cadherin peptides derived from the homophilic contact regions of the protein that are responsible for the cadherin-cadherin interactions were designed as blood-brain barrier modulators (BBBMs). HAV and ADT peptides (Table 7.2) derived from the contact regions of the extracellular-1 (EC1) domain of E-cadherin have been shown to modulate the BBB intercellular junctions *in vitro* and *in vivo*.

Initially, cadherin peptides (HAV and ADT peptides; Table 7.1) were evaluated to modulate *in vitro* intercellular junctions of cell monolayers of bovine brain microvessel endothelial cell (BBMEC) (Lutz and Siahaan 1997a), MDCK (Makagiansar et al. 2001; Sinaga et al. 2002), and Caco-2 cells (Kiptoo et al. 2011; Calcagno et al. 2004). It was found that HAV and ADT peptides (Table 7.2) modulate the intercellular junctions of cell monolayers as indicated by lowering *trans*-epithelial electrical resistance (TEER) values of the MDCK and Caco-2 monolayers upon incubation with cadherin peptides. HAV and ADT peptides also enhanced the transport of ^{14}C -mannitol paracellular marker molecules across the cell monolayers from the apical side (AP) to basolateral side (BL) (Fig. 7.3a) (Makagiansar et al. 2001; Laksitorini et al. 2015). As an example, HAV6 peptide produced an eightfold enhancement in mannitol permeability compared to control, while the derivative

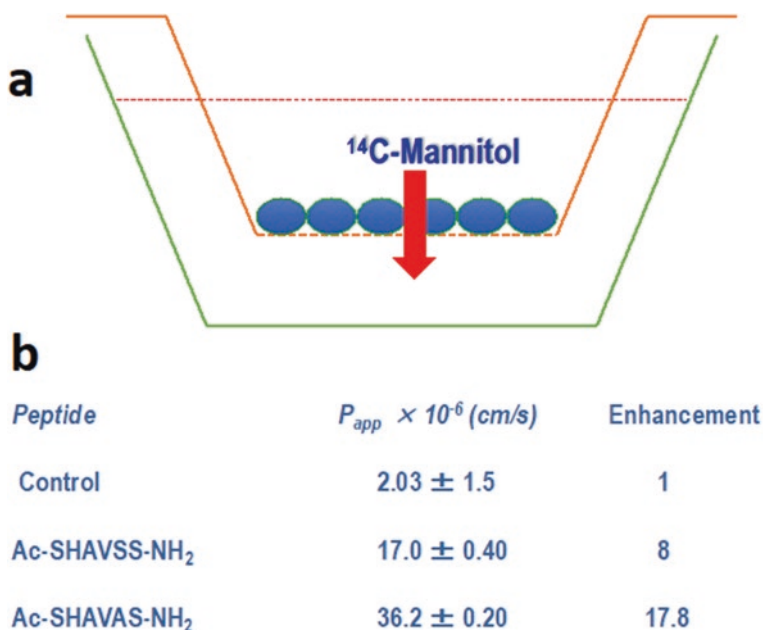


Fig. 7.3 (a) Diagram of biological cell culture monolayers as an *in vitro* model of the BBB. ^{14}C -mannitol is used as a paracellular marker to evaluate the modulation of the BBB paracellular pathways by the BBBMs such as HAV peptides. (b) HAV peptides enhanced the permeation of ^{14}C -mannitol from the apical (ap) side to basolateral (bl) side of MDCK cells as an *in vitro* model of the biological barrier

HAV4 enhanced mannitol 17 times more than control in the MDCK cell culture model (Fig. 7.3b).

The first animal study carried out using HAV6 peptide (Table 7.2) to enhance brain delivery of ^{14}C -mannitol into the brain used the in situ rat brain perfusion method, which is a well-established method to study drug delivery across the BBB (Takasato et al. 1984; Neuwelt et al. 1984a). Infusions of ^{14}C -mannitol alone and ^{14}C -mannitol PLUS_SPI LABL6 peptide (EATDSG) were used as negative controls (Kiptoo et al. 2011). The data showed that HAV6 peptide significantly enhanced brain deposition of ^{14}C -mannitol compared to that of ^{14}C -mannitol alone or ^{14}C -mannitol PLUS_SPI LABL6 peptide.¹⁴ This indicates that HAV6 modulates the BBB to increase the paracellular porosity that allows mannitol to pass through the BBB (Kiptoo et al. 2011).

To test whether HAV6 can enhance brain delivery of drugs that are Pgp substrates, the delivery of ^3H -daunomycin into the brain was evaluated using the in situ rat brain perfusion method (Kiptoo et al. 2011). Because ^3H -daunomycin is a substrate of Pgp, its infusion showed low deposition in the brain because it was

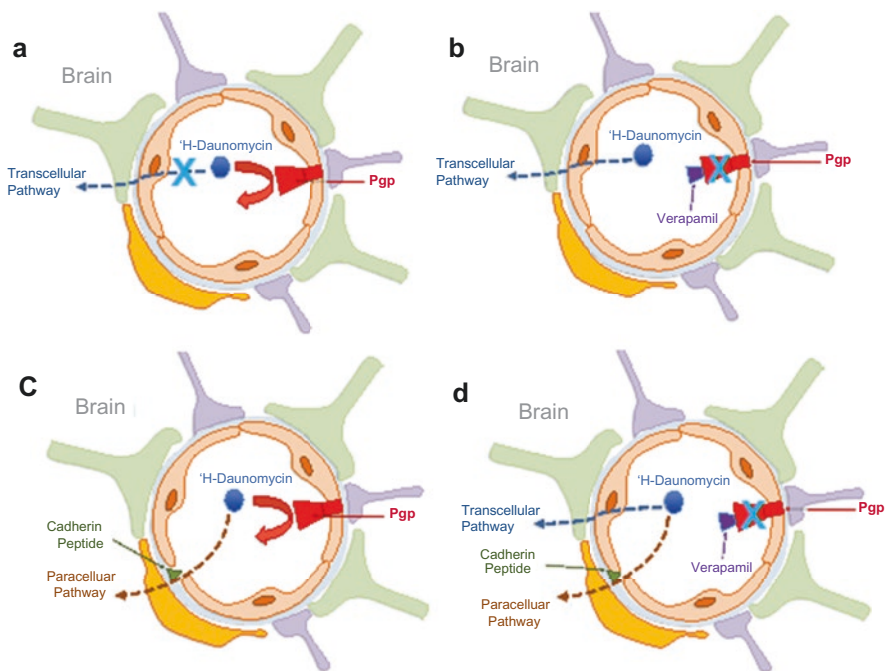


Fig. 7.4 Mechanism of ^3H -daunomycin transport across the BBB as Pgp substrate. (a) ^3H -daunomycin is a substrate of Pgp, and it was prevented by Pgp to diffuse through the BBB via transcellular pathway. (b) Inhibition of Pgp by verapamil enhanced transcellular diffusion ^3H -daunomycin across the BBB. (c) Modulation of the BBB intercellular junctions by cadherin peptides allows the penetration of ^3H -daunomycin through the BBB via the paracellular pathway. (d) Co-administration of ^3H -daunomycin with verapamil and cadherin peptide enhanced the BBB diffusion of ^3H -daunomycin via both trans- and paracellular pathways

prevented by Pgp to diffuse through the transcellular pathway (Fig. 7.4a). However, co-delivery of verapamil as a Pgp inhibitor with ^3H -daunomycin significantly enhanced passive diffusion of ^3H -daunomycin via transcellular pathway (Fig. 7.4b). A combination of ^3H -daunomycin and HAV6 significantly enhanced the brain deposition of ^3H -daunomycin compared to delivering ^3H -daunomycin alone. The results indicate that HAV6 improved the BBB penetration via paracellular pathway (Fig. 7.4c). Also, it was found that scramble HAV6 (HAV6scr; Table 7.2) did not improve the BBB permeation of ^3H -daunomycin, suggesting that the sequence specificity of HAV6 was necessary for BBBM activity (Kiptoo et al. 2011). A combination of HAV6 and verapamil led to higher delivery of ^3H -daunomycin compared to HAV6 or verapamil alone, suggesting improved BBB penetration of ^3H -daunomycin was via both transcellular and paracellular pathways (Fig. 7.4d).

Subsequently, HAV6 peptide has been shown to enhance brain delivery of camptothecin-glutamic acid (CPT-Glu) conjugate and gadopentetic acid (Gd-DTPA) as detected using LC-MS/MS and magnetic resonance imaging (MRI), respectively (Tabanor et al. 2016). A significantly higher amount of CPT-Glu was found in the rat brain after administration of CPT-Glu along with HAV6 via the rat's left carotid artery compared to CPT-Glu administration alone in *in situ* rat brain perfusion method (Tabanor et al. 2016). In the same study, Gd-DTPA, as an MRI enhancing agent, was administered via *i.v.* injections, and after 9 min, peptide or vehicle was administered via *i.v.* followed by monitoring the amount of Gd-DTPA in the brain by MRI (Tabanor et al. 2016). Immediately after the injection of HAV6 peptide, significantly higher depositions of Gd-DTPA in the rat brains were observed compared to those injected with vehicle (Tabanor et al. 2016). The significant increases were observed in the hippocampus, cerebellum, brain ventral, deep rostral, and deep caudal of the brain (Tabanor et al. 2016). These results support the idea that HAV6 peptide immediately increases the porosity of the BBB paracellular pathway to allow paracellular penetration of Gd-DTPA.

To evaluate whether BBBMs can also work to modulate the BBB in mice, rhodamine 800 (R800) and 25 kDa IRdye800CW-polyethylene glycol (IRdye800CW-PEG) were delivered via *i.v.* injections with and without HAV6 peptide (On et al. 2014). As in the daunomycin study, R800 was selected because it is a substrate for Pgp efflux pump (Fig. 7.4a), while IRdye800CW-PEG (25 kDa) was selected as a large molecule marker (On et al. 2014). Brain depositions of both molecules can be quantified using near IR fluorescence (NIRF) imaging. It was found that administration of R800 along with HAV6 significantly increased the brain deposition of R800 compared to administration of R800 alone. As in daunomycin, although R800 is a substrate for Pgp, HAV6 peptide increased the BBB permeation of R800 via the paracellular pathway (Fig. 7.4c). Administration of R800 along with Pgp inhibitor GF120918 enhanced the penetration of R800 via the transcellular pathway of the BBB compared to administration of R800 alone, confirming that R800 is a substrate of Pgp (Fig. 7.4b) (On et al. 2014). Furthermore, brain delivery of IRdye800CW-PEG was enhanced significantly by HAV6 peptide but not by GF120918 when compared to administration of IRdye800CW-PEG alone. GF120918 did not enhance the transport of IRdye800CW-PEG because IRdye800CW-PEG cannot cross through

the transcellular pathway due to its hydrophilicity. Overall, HAV6 can increase the delivery of efflux pump substrate R800 and a large molecule PEG via the paracellular pathway (On et al. 2014).

The activity of HAV6 peptide to deliver Gd-DTPA into the brain of Balb/c mice was studied in time-dependent manner by monitoring its brain deposition every 3 min for 42 min using MRI after i.v. administration (On et al. 2014). The results indicated that the brain depositions of Gd-DTPA were significantly higher (two–fourfold) when delivered with HAV6 compared to those administered with Gd-DTPA alone (On et al. 2014). In addition, enhancement of Gd-DTPA brain deposition was observed in the first 3 min after i.v. administration. The duration of the BBB pore opening generated by HAV6 was determined using a pretreatment experiment in which the mice were first treated with HAV6 peptide. After a certain time delay, marker molecules such as Gd-DTPA were administered. It is interesting to find that after 60 min delay, no enhancement of Gd-DTPA brain deposition was observed, indicating that the BBB modulation by HAV6 peptide was over in after 60 min. The results suggested that modulation of the BBB by HAV6 peptide is temporary and reversible.

Another BBBM called ADTC5 peptide (Table 7.2) was found to enhance brain delivery of ^{14}C -mannitol and Gd-DTPA in mice (Laksitorini et al. 2015). Similar to HAV6 peptide, the duration of BBB modulation by ADTC5 was also determined by pretreatment with ADTC5 followed by delivery of Gd-DTPA. Using Gd-DTPA, it was observed that the BBB paracellular pathway pore opening by ADTC5 was closed between 2 and 4 h, indicating longer paracellular opening compared to HAV6 peptide. ADTC5 peptide can also enhance the brain deposition of cIBR7 peptide (Cyclo(1,8)CPRGG SVC) in rats and IRdye-800cw-labeled cLABEL peptide (IRdye-800cw-Cyclo(1,12)PenITDGEATDSGC) in mice as quantitatively determined by mass spectrometry and NIRF imaging, respectively (Ulapane et al. 2017).

It is interesting to find that ADTC5 peptide enhanced brain delivery of various size proteins, including 15 kDa lysozyme, 65 kDa albumin, and 150 kDa IgG monoclonal antibody (mAb) (Ulapane et al. 2019b). However, ADTC5 cannot enhance the delivery of 220 kDa fibronectin, suggesting that there is a size limit of protein that can be delivered to the brain by ADTC5. The data showed that ADTC5 peptide was better than HAV6 peptide in delivering lysozyme and albumin across the BBB (Ulapane et al. 2019b). The duration of BBB opening depended on the size of the delivered molecule. For example, during BBB pretreatment using BBBM peptide before delivering the marker molecule, it showed that the paracellular pathway pore opening by ADTC5 lasted about 40 min when delivering a small molecule such as Gd-DTPA. However, the BBB paracellular opening was less than 10 min for the larger 65 kDa albumin permeability marker. It hypothesized that the BBBMs disrupt the intercellular junctions to generate large, medium, and small pores. The large pores collapse rapidly to medium followed by the conversion of medium pores to small pores in time-dependent fashion. Thus, the BBB paracellular opening for large- and medium-size molecules is shorter than for small molecules (Ulapane et al. 2017).

7.8 Nasal Delivery of Peptides

An intranasal delivery method is one way to enhance delivery of peptides into the brain by bypassing the BBB in a noninvasive manner. Some peptides and proteins such as oxytocin and insulin have delivered intranasally to reach phase IV clinical trials for neurodegenerative diseases such as Alzheimer's disease and related brain diseases (Samaridou and Alonso 2018). Thus, many peptides and proteins have been extensively investigated for their brain delivery via nasal administration to potentially treat neurodegenerative diseases (Meredith et al. 2015). One of the advantages of nasal delivery method is that it avoids peptide degradation in the blood as well as it can target a specific brain region such as an olfactory bulb and its surrounding brain regions (Meredith et al. 2015). Thus, nasal delivery could be more favorable than a systemic delivery via i.v. administration because during i.v. administration the peptide could be degraded and cleared in the systemic circulation.

To improve nasal delivery of peptides, many strategies have been developed. In general, peptides had to overcome nasal mucosa, olfactory epithelium barrier, and the cribriform plate before entering the olfactory bulb of the brain (Fig. 7.5) (Meredith et al. 2015; Samaridou and Alonso 2018). The peptides also had to survive enzymatic degradation at the olfactory epithelium. There are three ways that peptides can enter the brain. First, peptides enter the nasal cavity followed by crossing transcellular or paracellular pathway of the olfactory epithelium as well as crossing the cribriform plate into the olfactory bulb. Second, peptides enter the nasal cavity and are absorbed into the blood vessels of the systemic circulation followed by crossing the BBB into the brain. Third, the peptide could diffuse through

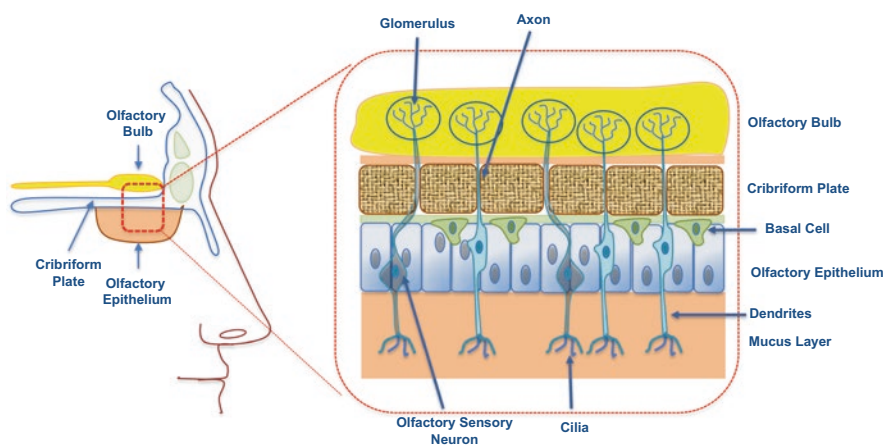


Fig. 7.5 The structure of nasal cavity as route for the delivery of peptides via intranasal delivery. The peptides need to cross the olfactory epithelium and cribriform plate to enter the olfactory bulb of the brain. The structure of the epithelium is decorated with sensory neurons with dendrites and cilia along with axons penetrating cribriform plates into the olfactory bulb

neuronal axons of the trigeminal nerve (Fig. 7.5) (Samaridou and Alonso 2018). Although the trigeminal nerve is one of the proposed transport pathways for peptides into the brain, this pathway has not been fully investigated. It has been proposed that the paracellular pathway is the major route for peptides to cross the epithelial olfactory barrier and through the cribriform plate (Fig. 7.5). It has been shown that tight junction disruptors such as carnitines and ultrasound improved peptide intranasal delivery. Cell-penetrating peptides (CPPs), receptor-mediated endocytosis process, and nanoparticles have also been explored to enhance transcellular transport of peptides through the nasal olfactory epithelium.

Oxytocin has been delivered intranasally for treating autism spectrum disorder (ASD) in clinical trials. The pharmacokinetics and brain distribution of oxytocin have been fully investigated (Tanaka et al. 2018). The nasally administered oxytocin can effectively reach the brain compared to those administered intravenously and intraperitoneally (Tanaka et al. 2018). The majority of oxytocin was found in the olfactory bulb followed by the frontal and occipital halves of the brain. Because the site of action for treatment of ASD is near the olfactory bulb, the intranasal delivery is suitable for this purpose (Tanaka et al. 2018). During nasal delivery, oxytocin was chemically stable and can be absorbed across the nasal epithelium (Tanaka et al. 2018). The plasma concentrations of the intranasally delivered oxytocin were much lower than intravenous (Tanaka et al. 2018). In a stress mouse model that is similar to ASD, intranasal delivery of oxytocin lowered the plasma levels of a stress hormone corticosteroid compared to that of i.v. administered oxytocin, implying the nasal delivery was more effective than that of i.v. administration (Tanaka et al. 2018).

Vasoactive intestinal peptide (VIP) has been formulated for nasal delivery with optimum stability in 1% bovine serum albumin concentration (BSA) and 0.1% lauroylcarnitine (LC) (Dufes et al. 2003). Nasal delivery of VIP was also more effective than i.v. administration with high deposition in the cerebellum, central gray, amygdaloid nuclei, and thalamic, hypothalamic, and olfactory bulbs where VIP receptors are located (Dufes et al. 2003). The nasal delivery of VIP was pH dependent, and the best brain uptake of ^{125}I -VIP was at pH 9 as detected by radioactivity levels in the brain (Dufes et al. 2003). The presence of LC in the formulation increased brain uptake of VIP at pH 4 and 7; this was due to the effect of LC in disrupting the BBB tight junctions by lowering the expression levels of tight junction proteins such as claudins (e.g., claudin 1, 4, and 5) (Doi et al. 2011). During i.v. administration, a low amount of intact VIP was found in the brain, and most radioactivity found distributed throughout the whole brain emanated from ^{125}I -VIP degradation products because VIP was degraded rapidly in the blood (Dufes et al. 2003). In contrast, intact ^{125}I -VIP was found in the brain when delivered via nasal route. These results suggest that nasal peptide delivery was better than i.v. delivery and the penetration of VIP was via the paracellular pathways of the olfactory epithelium.

Pituitary adenylate cyclase-activating polypeptide (PCAP) alone can cross the BBB when administered via i.v. route; however, the transport across the BBB was not effective due to its recognition by an efflux pump and degradation in the blood. Alternatively, PACP in formulation with six monosaccharide cyclodextrin delivered

intranasally enhanced brain deposition at the olfactory bulb but not the other brain regions (Meredith et al. 2015). On the other hand, cyclodextrin with seven monosaccharides increased PCAP levels in the whole brain except the olfactory bulb and striatum (Meredith et al. 2015). Although the mechanism of action of these cyclodextrins is still not clear, it has been proposed that they interact with cholesterol of cell membranes to disrupt cell membrane integrity of the olfactory epithelium to improve peptide permeation through para- and transcellular pathways (Kiss et al. 2010; Hussain et al. 2003). Similarly, intranasal co-administration of galanin-like peptide (GALP) with cyclodextrin resulted in its deposition on all brain regions, while administration of GALP alone resulted in a high deposition in the olfactory bulb only.

Other peptides such as exendin, insulin, and leptin-like peptide have been found to enter the brain when administered via intranasal route with higher brain deposition in the olfactory bulb compared to those delivered via i.v. administrations. In general, intranasal delivery increased the peptide deposition at the olfactory bulb better than other brain regions. As with other biological barriers, peptides and proteins have physicochemical properties that are not conducive to cross via the transcellular pathway of the olfactory epithelium; thus, the transport pathway through epithelium is most likely via the paracellular route. For effective brain delivery, the peptide should penetrate the mucosa epithelium and withstand enzymatic degradation in the mucosa epithelium. Because of the limited delivery space of the olfactory epithelium, peptides had to be delivered in a high concentration with less than 400 μ L of delivery volume using intranasal liquid delivery systems. The role of permeation enhancers such as cyclodextrins, CPP, and others became important for intranasal peptide brain delivery.

Nanotechnology has also been utilized for intranasal delivery of peptides to the brain because nanoparticles can encapsulate a high dose of peptide as well as provide peptide protection from enzymatic degradation in the mucosa epithelium. Nanoparticles have been constructed from various materials, including polylactic/glycolic acid, chitosan, gelatin, or cationic liposomes. Liposomes, with size around 100 nm, have extensively been used to deliver peptides intranasally to olfactory bulb through the axons (Samaridou and Alonso 2018). Because mucus has negative charges, it is preferable that the nanoparticles have positive charges for their diffusion through nasal mucus layers and olfactory epithelium (Fig. 7.5). The presence of surfactants (e.g., Tween 80, polysorbate 80, poloxamer 188) can improve the penetration of peptide- or protein-loaded nanoparticles; they presumably behave as membrane disruptors that allow nanoparticle penetration through paracellular pathways of nasal epithelium. Pegylated liposomes have been developed to deliver H102 peptide for treatment of Alzheimer's disease in a rat model. In addition, substance P peptide was successfully delivered to the brain using gelatin-lipid nanoparticles via nasal delivery route. In summary, the physicochemical properties of the nanoparticles (i.e., size, charge, composition of surface) influence their penetration across the nasal epithelium into the olfactory bulb.

For receptor-mediated transport, the surface of nanoparticles was decorated with peptides or proteins for improving selectivity and uptake by nasal epithelium.

Lactoferrin has been used to decorate nanoparticle surfaces for nasal delivery of neuropeptide (Neuwelt et al. 1985). The function of lactoferrin is to target lactoferrin receptors, which are highly expressed in the brain (Samaridou and Alonso 2018). CPP-decorated chitosan nanoparticles have also been used to improve brain delivery of neurotrophic factors via intranasal administration.

7.9 Conclusions

Delivery of drug and diagnostic molecules to the brain is still challenging; however, many advancements have been made to deliver small up to large molecules across the BBB and into the brain. Osmotic BBB disruption method has been successfully used to deliver anticancer drugs to brain tumor patients. Preclinically, BBBM peptides have been used to deliver various proteins to the brain in animal models of MS, Alzheimer's disease, and brain tumors; however, more work is needed to evaluate the side effects of this method. Intranasal delivery of peptides to the brain has shown some success in the preclinical and clinical studies with the hope that the method can be used to effectively deliver various peptides to the brain. Thus, there is a hope that some of these methods can help advance the development of therapeutic and diagnostic agents for brain disease.

References

- Abbruscato TJ, Thomas SA, Hruby VJ, Davis TP. Blood-brain barrier permeability and bioavailability of a highly potent and mu-selective opioid receptor antagonist, CTAP: comparison with morphine. *J Pharmacol Exp Ther.* 1997;280:402–9.
- Banks WA. The source of cerebral insulin. *Eur J Pharmacol.* 2004;490:5–12.
- Banks WA, Farrell CL. Impaired transport of leptin across the blood-brain barrier in obesity is acquired and reversible. *Am J Physiol Endocrinol Metab.* 2003;285:E10–5.
- Banks WA, Farr SA, La Scola ME, Morley JE. Intravenous human interleukin-1 α impairs memory processing in mice: dependence on blood-brain barrier transport into posterior division of the septum. *J Pharmacol Exp Ther.* 2001;299:536–41.
- Banks WA, Owen JB, Erickson MA. Insulin in the brain: there and back again. *Pharmacol Ther.* 2012;136:82–93.
- Broadwell RD, Balin BJ, Salzman M. Transcytotic pathway for blood-borne protein through the blood-brain barrier. *Proc Natl Acad Sci U S A.* 1988;85:632–6.
- Burgess A, Shah K, Hough O, Hynynen K. Focused ultrasound-mediated drug delivery through the blood-brain barrier. *Expert Rev Neurother.* 2015;15:477–91.
- Calcagno AM, Fostel JM, Reyner EL, Sinaga E, Alston JT, Mattes WB, Siahaan TJ, Ware JA. Effects of an E-cadherin-derived peptide on the gene expression of Caco-2 cells. *Pharm Res.* 2004;21:2085–94.
- Chorev M, Shavitz R, Goodman M, Minick S, Guillemin R. Partially modified retro-inverso-enkephalinamides: topochemical long-acting analogs in vitro and in vivo. *Science.* 1979;204:1210–2.

- Dauchy S, Duthiel F, Weaver RJ, Chassoux F, Dumas-Duport C, Couraud PO, Scherrmann JM, De Waziers I, Declèves X. ABC transporters, cytochromes P450 and their main transcription factors: expression at the human blood-brain barrier. *J Neurochem*. 2008;107:1518–28.
- Demeule M, Poirier J, Jodoin J, Bertrand Y, Desrosiers RR, Dagenais C, Nguyen T, Lanthier J, Gabathuler R, Kennard M, Jefferies WA, Karkan D, Tsai S, Fenart L, Cecchelli R, Beliveau R. High transcytosis of melanotransferrin (P97) across the blood-brain barrier. *J Neurochem*. 2002;83:924–33.
- Doi N, Tomita M, Hayashi M. Absorption enhancement effect of acylcarnitines through changes in tight junction protein in Caco-2 cell monolayers. *Drug Metab Pharmacokinet*. 2011;26:162–70.
- Doolittle ND, Muldoon LL, Culp AY, Neuwelt EA. Delivery of chemotherapeutics across the blood-brain barrier: challenges and advances. *Adv Pharmacol*. 2014;71:203–43.
- Dufes C, Olivier JC, Gaillard F, Gaillard A, Couet W, Muller JM. Brain delivery of vasoactive intestinal peptide (VIP) following nasal administration to rats. *Int J Pharm*. 2003;255:87–97.
- Egleton RD, Abbruscato TJ, Thomas SA, Davis TP. Transport of opioid peptides into the central nervous system. *J Pharm Sci*. 1998;87:1433–9.
- Eriste E, Kurrikoff K, Suhorutsenko J, Oskolkov N, Copolovici DM, Jones S, Laakkonen P, Howl J, Langel U. Peptide-based glioma-targeted drug delivery vector gHoPe2. *Bioconjug Chem*. 2013;24:305–13.
- Fishman JB, Rubin JB, Handrahan JV, Connor JR, Fine RE. Receptor-mediated transcytosis of transferrin across the blood-brain barrier. *J Neurosci Res*. 1987;18:299–304.
- Fletcher NF, Callanan JJ. Cell culture models of the blood-brain barrier: new research. In: Montenegro PA, Juarez SM, editors. *The blood-brain barrier: new research*. Hauppauge: Nova Science Publisher; 2012.
- Fosgerau K, Hoffmann T. Peptide therapeutics: current status and future directions. *Drug Discov Today*. 2015;20:122–8.
- Frank HJ, Pardridge WM. A direct in vitro demonstration of insulin binding to isolated brain microvessels. *Diabetes*. 1981;30:757–61.
- Frank HJ, Pardridge WM, Morris WL, Rosenfeld RG, Choi TB. Binding and internalization of insulin and insulin-like growth factors by isolated brain microvessels. *Diabetes*. 1986;35:654–61.
- Frankel AD, Pabo CO. Cellular uptake of the tat protein from human immunodeficiency virus. *Cell*. 1988;55:1189–93.
- Ghersi-Egea JF, Leininger-Muller B, Cecchelli R, Fenstermacher JD. Blood-brain interfaces: relevance to cerebral drug metabolism. *Toxicol Lett*. 1995;82–83:645–53.
- Ghosh C, Gonzalez-Martinez J, Hossain M, Cucullo L, Fazio V, Janigro D, Marchi N. Pattern of P450 expression at the human blood-brain barrier: roles of epileptic condition and laminar flow. *Epilepsia*. 2010;51:1408–17.
- Ghosh C, Puvenna V, Gonzalez-Martinez J, Janigro D, Marchi N. Blood-brain barrier P450 enzymes and multidrug transporters in drug resistance: a synergistic role in neurological diseases. *Curr Drug Metab*. 2011;12:742–9.
- Green M, Loewenstein PM. Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. *Cell*. 1988;55:1179–88.
- Hermann DM, Bassetti CL. Implications of ATP-binding cassette transporters for brain pharmacotherapies. *Trends Pharmacol Sci*. 2007;28:128–34.
- Herve F, Ghinea N, Scherrmann JM. CNS delivery via adsorptive transcytosis. *AAPS J*. 2008;10:455–72.
- Herz J, Marschang P. Coaxing the LDL receptor family into the fold. *Cell*. 2003;112:289–92.
- Hussain A, Yang T, Zaghoul AA, Ahsan F. Pulmonary absorption of insulin mediated by tetradecyl-beta-maltoside and dimethyl-beta-cyclodextrin. *Pharm Res*. 2003;20:1551–7.
- Javed H, Menon SA, Al-Mansoori KM, Al-Wandi A, Majbour NK, Ardah MT, Varghese S, Vaikath NN, Haque ME, Azzouz M, El-Agnaf OM. Development of nonviral vectors targeting the brain as a therapeutic approach for Parkinson's disease and other brain disorders. *Mol Ther*. 2016;24:746–58.

- Kadry H, Noorani B, Cucullo L. A blood-brain barrier overview on structure, function, impairment, and biomarkers of integrity. *Fluids Barriers CNS*. 2020;17:69.
- Kaspar AA, Reichert JM. Future directions for peptide therapeutics development. *Drug Discov Today*. 2013;18:807–17.
- Kim JA, Casalini T, Brambilla D, Leroux JC. Presumed LRP1-targeting transport peptide delivers beta-secretase inhibitor to neurons in vitro with limited efficiency. *Sci Rep*. 2016;6:34297.
- Kiptoo P, Sinaga E, Calcagno AM, Zhao H, Kobayashi N, Tambunan US, Siahaan TJ. Enhancement of drug absorption through the blood-brain barrier and inhibition of intercellular tight junction resealing by E-cadherin peptides. *Mol Pharm*. 2011;8:239–49.
- Kiss T, Fenyvesi F, Bacskay I, Varadi J, Fenyvesi E, Ivanyi R, Szente L, Tosaki A, Vecsernyes M. Evaluation of the cytotoxicity of beta-cyclodextrin derivatives: evidence for the role of cholesterol extraction. *Eur J Pharm Sci*. 2010;40:376–80.
- Kumar P, Wu H, McBride JL, Jung KE, Kim MH, Davidson BL, Lee SK, Shankar P, Manjunath N. Transvascular delivery of small interfering RNA to the central nervous system. *Nature*. 2007;448:39–43.
- Laksitorini MD, Kiptoo PK, On NH, Thliveris JA, Miller DW, Siahaan TJ. Modulation of intercellular junctions by cyclic-ADT peptides as a method to reversibly increase blood-brain barrier permeability. *J Pharm Sci*. 2015;104:1065–75.
- Li C, Pazgier M, Li J, Li C, Liu M, Zou G, Li Z, Chen J, Tarasov SG, Lu WY, Lu W. Limitations of peptide retro-inverso isomerization in molecular mimicry. *J Biol Chem*. 2010;285:19572–81.
- Li Y, Zheng X, Gong M, Zhang J. Delivery of a peptide-drug conjugate targeting the blood brain barrier improved the efficacy of paclitaxel against glioma. *Oncotarget*. 2016;7:79401–7.
- Lindgren M, Langel U. Classes and prediction of cell-penetrating peptides. *Methods Mol Biol*. 2011;683:3–19.
- Lipinski CA. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov Today Technol*. 2004;1:337–41.
- Lipinski CA. Rule of five in 2015 and beyond: target and ligand structural limitations, ligand chemistry structure and drug discovery project decisions. *Adv Drug Deliv Rev*. 2016;101:34–41.
- Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*. 2001;46:3–26.
- Lippmann ES, Al-Ahmad A, Palecek SP, Shusta EV. Modeling the blood-brain barrier using stem cell sources. *Fluids Barriers CNS*. 2013;10:2.
- Liu Y, Huang R, Han L, Ke W, Shao K, Ye L, Lou J, Jiang C. Brain-targeting gene delivery and cellular internalization mechanisms for modified rabies virus glycoprotein RVG29 nanoparticles. *Biomaterials*. 2009;30:4195–202.
- Loscher W, Potschka H. Role of drug efflux transporters in the brain for drug disposition and treatment of brain diseases. *Prog Neurobiol*. 2005;76:22–76.
- Luo H, Shusta EV. Blood-brain barrier modulation to improve glioma drug delivery. *Pharmaceutics*. 2020;12:1085.
- Lutz KL, Siahaan TJ. Modulation of the cellular junction protein E-cadherin in bovine brain microvessel endothelial cells by cadherin peptides. *Drug Deliv*. 1997a;4:187–93.
- Lutz KL, Siahaan TJ. Molecular structure of the apical junction complex and its contribution to the paracellular barrier. *J Pharm Sci*. 1997b;86:977–84.
- Makagiansar IT, Avery M, Hu Y, Audus KL, Siahaan TJ. Improving the selectivity of HAV-peptides in modulating E-cadherin-E-cadherin interactions in the intercellular junction of MDCK cell monolayers. *Pharm Res*. 2001;18:446–53.
- Marchi N, Angelov L, Masaryk T, Fazio V, Granata T, Hernandez N, Hallene K, Diglaw T, Franic L, Najm I, Janigro D. Seizure-promoting effect of blood-brain barrier disruption. *Epilepsia*. 2007;48:732–42.
- McCully M, Sanchez-Navarro M, Teixido M, Giralt E. Peptide mediated brain delivery of nano- and submicroparticles: a synergistic approach. *Curr Pharm Des*. 2018;24:1366–76.

- Meredith ME, Salameh TS, Banks WA. Intranasal delivery of proteins and peptides in the treatment of neurodegenerative diseases. *AAPS J.* 2015;17:780–7.
- Naik P, Cucullo L. In vitro blood-brain barrier models: current and perspective technologies. *J Pharm Sci.* 2012;101:1337–54.
- Neuwelt EA, Barnett PA, Frenkel EP. Chemotherapeutic agent permeability to normal brain and delivery to avian sarcoma virus-induced brain tumors in the rodent: observations on problems of drug delivery. *Neurosurgery.* 1984a;14:154–60.
- Neuwelt EA, Hill SA, Frenkel EP. Osmotic blood-brain barrier modification and combination chemotherapy: concurrent tumor regression in areas of barrier opening and progression in brain regions distant to barrier opening. *Neurosurgery.* 1984b;15:362–6.
- Neuwelt EA, Lawrence MS, Blank NK. Effect of gentamicin and dexamethasone on the natural history of the rat *Escherichia coli* brain abscess model with histopathological correlation. *Neurosurgery.* 1984c;15:475–83.
- Neuwelt EA, Barnett PA, McCormick CI, Frenkel EP, Minna JD. Osmotic blood-brain barrier modification: monoclonal antibody, albumin, and methotrexate delivery to cerebrospinal fluid and brain. *Neurosurgery.* 1985;17:419–23.
- Okumu FW, Pauletti GM, Vander Velde DG, Siahaan TJ, Borchardt RT. Effect of restricted conformational flexibility on the permeation of model hexapeptides across Caco-2 cell monolayers. *Pharm Res.* 1997;14:169–75.
- Oller-Salvia B, Sanchez-Navarro M, Giralte E, Teixido M. Blood-brain barrier shuttle peptides: an emerging paradigm for brain delivery. *Chem Soc Rev.* 2016;45:4690–707.
- On NH, Miller DW. Transporter-based delivery of anticancer drugs to the brain: improving brain penetration by minimizing drug efflux at the blood-brain barrier. *Curr Pharm Des.* 2014;20:1499–509.
- On NH, Kiptoo P, Siahaan TJ, Miller DW. Modulation of blood-brain barrier permeability in mice using synthetic E-cadherin peptide. *Mol Pharm.* 2014;11:974–81.
- Ouyang H, Tang F, Siahaan TJ, Borchardt RT. A modified coumarinic acid-based cyclic prodrug of an opioid peptide: its enzymatic and chemical stability and cell permeation characteristics. *Pharm Res.* 2002;19:794–801.
- Ouyang H, Andersen TE, Chen W, Nofsinger R, Steffansen B, Borchardt RT. A comparison of the effects of p-glycoprotein inhibitors on the blood-brain barrier permeation of cyclic prodrugs of an opioid peptide (DADLE). *J Pharm Sci.* 2009a;98:2227–36.
- Ouyang H, Chen W, Andersen TE, Steffansen B, Borchardt RT. Factors that restrict the intestinal cell permeation of cyclic prodrugs of an opioid peptide (DADLE): part I. role of efflux transporters in the intestinal mucosa. *J Pharm Sci.* 2009b;98:337–48.
- Pan W, Kastin AJ. Entry of EGF into brain is rapid and saturable. *Peptides.* 1999;20:1091–8.
- Pan W, Kastin AJ. TNF α transport across the blood-brain barrier is abolished in receptor knockout mice. *Exp Neurol.* 2002;174:193–200.
- Paterson J, Webster CI. Exploiting transferrin receptor for delivering drugs across the blood-brain barrier. *Drug Discov Today Technol.* 2016;20:49–52.
- Pauletti GM, Gangwar S, Okumu FW, Siahaan TJ, Stella VJ, Borchardt RT. Esterase-sensitive cyclic prodrugs of peptides: evaluation of an acyloxyalkoxy moiety in a model hexapeptide. *Pharm Res.* 1996;13:1615–23.
- Rhea EM, Salameh TS, Gray S, Niu J, Banks WA, Tong J. Ghrelin transport across the blood-brain barrier can occur independently of the growth hormone secretagogue receptor. *Mol Metab.* 2018;18:88–96.
- Samaridou E, Alonso MJ. Nose-to-brain peptide delivery – the potential of nanotechnology. *Bioorg Med Chem.* 2018;26:2888–905.
- Sanchez del Pino MM, Hawkins RA, Peterson DR. Biochemical discrimination between luminal and abluminal enzyme and transport activities of the blood-brain barrier. *J Biol Chem.* 1995;270:14907–12.

- Shah MV, Audus KL, Borchardt RT. The application of bovine brain microvessel endothelial-cell monolayers grown onto polycarbonate membranes in vitro to estimate the potential permeability of solutes through the blood-brain barrier. *Pharm Res.* 1989;6:624–7.
- Sharif Y, Jumah F, Coplan L, Krosser A, Sharif K, Tubbs RS. Blood brain barrier: a review of its anatomy and physiology in health and disease. *Clin Anat.* 2018;31:812–23.
- Shawahna R, Uchida Y, Declèves X, Ohtsuki S, Yousif S, Dauchy S, Jacob A, Chassoux F, Daumas-Duport C, Couraud PO, Terasaki T, Scherrmann JM. Transcriptomic and quantitative proteomic analysis of transporters and drug metabolizing enzymes in freshly isolated human brain microvessels. *Mol Pharm.* 2011;8:1332–41.
- Sinaga E, Jois SD, Avery M, Makagiarsar IT, Tambunan US, Audus KL, Siahaan TJ. Increasing paracellular porosity by E-cadherin peptides: discovery of bulge and groove regions in the EC1-domain of E-cadherin. *Pharm Res.* 2002;19:1170–9.
- Tabanor K, Lee P, Kiptoo P, Choi IY, Sherry EB, Eagle CS, Williams TD, Siahaan TJ. Brain delivery of drug and MRI contrast agent: detection and quantitative determination of brain deposition of CPT-Glu using LC-MS/MS and Gd-DTPA using magnetic resonance imaging. *Mol Pharm.* 2016;13:379–90.
- Takasato Y, Rapoport SI, Smith QR. An in situ brain perfusion technique to study cerebrovascular transport in the rat. *Am J Phys.* 1984;247:H484–93.
- Tanaka A, Furubayashi T, Arai M, Inoue D, Kimura S, Kiriyama A, Kusamori K, Katsumi H, Yutani R, Sakane T, Yamamoto A. Delivery of oxytocin to the brain for the treatment of autism spectrum disorder by nasal application. *Mol Pharm.* 2018;15:1105–11.
- Uhlig T, Kyprianou T, Martinelli FG, Oppici CA, Heiligers D, Hills D, Calvo XR, Verhaert P. The emergence of peptides in the pharmaceutical business: from exploration to exploitation. *EuPA Open Proteom.* 2014;4:58–69.
- Ulapane KR, On N, Kiptoo P, Williams TD, Miller DW, Siahaan TJ. Improving brain delivery of biomolecules via BBB modulation in mouse and rat: detection using MRI, NIRF, and mass spectrometry. *Nanotheranostics.* 2017;1:217–31.
- Ulapane KR, Kopec BM, Siahaan TJ. Improving in vivo brain delivery of monoclonal antibody using novel cyclic peptides. *Pharmaceutics.* 2019a;11:568.
- Ulapane KR, Kopec BM, Siahaan TJ. In vivo brain delivery and brain deposition of proteins with various sizes. *Mol Pharm.* 2019b;16:4878–89.
- van der Lely AJ, Tschop M, Heiman ML, Ghigo E. Biological, physiological, pathophysiological, and pharmacological aspects of ghrelin. *Endocr Rev.* 2004;25:426–57.
- Visser CC, Stevanovic S, Heleen Voorwinden L, Gaillard PJ, Crommelin DJ, Danhof M, De Boer AG. Validation of the transferrin receptor for drug targeting to brain capillary endothelial cells in vitro. *J Drug Target.* 2004;12:145–50.
- Wei HJ, Upadhyayula PS, Poulipoulos AN, Englander ZK, Zhang X, Jan CI, Guo J, Mela A, Zhang Z, Wang TJC, Bruce JN, Canoll PD, Feldstein NA, Zacharoulis S, Konofagou EE, Wu CC. Focused ultrasound-mediated blood-brain barrier opening increases delivery and efficacy of etoposide for glioblastoma treatment. *Int J Radiat Oncol Biol Phys.* 2021;110:539–50.
- Weksler B, Romero IA, Couraud PO. The hCMEC/D3 cell line as a model of the human blood brain barrier. *Fluids Barriers CNS.* 2013;10:16.
- Zhai M, Wang Y, Zhang L, Liang M, Fu S, Cui L, Yang M, Gong W, Li Z, Yu L, Xie X, Yang C, Yang Y, Gao C. Glioma targeting peptide modified apoferritin nanocage. *Drug Deliv.* 2018;25:1013–24.
- Zlokovic BV, Skundric DS, Segal MB, Lipovac MN, Mackic JB, Davson H. A saturable mechanism for transport of immunoglobulin G across the blood-brain barrier of the guinea pig. *Exp Neurol.* 1990;107:263–70.

Chapter 8

Emerging Peptide Drug Modalities for Intracellular Target Space



Tomi K. Sawyer

Contents

8.1 Introduction.....	268
8.2 Intracellularly Targeted Peptides: Some Historical Milestones.....	269
8.3 Expanding Intracellular Target Space: Emerging Peptide Modalities.....	274
8.4 Unlocking the Secrets of Cell Permeability: Exploiting Innovative Tools.....	277
8.5 Future Intracellularly Targeted Peptide Drugs: Clinical Trials and Beyond.....	280
References.....	281

Abstract Historically, peptide drug discovery has been very successful in the development of receptor-targeted medicines such as related to insulin and glucagon-like peptide 1 agonists as well as many other examples with respect to many receptors and other extracellular targets. In recent years, there has been significant progress to advance new peptide modalities focused especially on macrocyclic design and leveraging lead molecules from biological and/or chemical diversity approaches, including mRNA-display libraries, phage-display libraries, DNA-encoded synthetic libraries, and one-bead-one-compound synthetic libraries. Such work builds upon existing peptidomimetic and peptide analog optimization strategies involving a native cognate peptide (or protein fragment) and iterative structure-based design. Likewise, there has been incredible progress in structural biology and computational modeling that is contributing to peptide drug modalities, including linear and macrocyclic peptides as well as peptidomimetic analogs thereof. Collectively, this armamentarium of peptide modalities has contributed to the acceleration of breakthrough preclinical molecules. A greater understanding of drug-like properties to tackle an increasing number of intracellular targets (e.g., enzymes and protein–protein interactions) as well as deeper insights related to cell uptake mechanisms, including passive transport and both cationic and lipophilic partitioning models, is being achieved. This chapter exemplifies a few specific cases of intracel-

T. K. Sawyer (✉)
Maestro Therapeutics, Southborough, MA, USA

lular targets and varying peptide drug modalities which illustrate success toward a new wave of novel peptide therapeutics.

Keywords Cell-penetrating peptides · Macrocyclic peptides · Peptidomimetics · Intracellular targets · Protease · Proteasome · Phosphatase · Farnesyl transferase · GTPase · Src SH2 · XIAP · MCL-1 · HIF-1 α · NEMO · CFTR · MDM2 · MDM4 · Protein · protein interactions

8.1 Introduction

One of the most extraordinary adventures for peptide drug discovery since the beginning of the twenty-first century has been the pioneering efforts across academia, biotech, and pharma to advance the generation, optimization, and development of breakthrough peptide therapeutics for intracellular targets. In this chapter, it is my intention to reflect upon key scientific concepts and innovative technologies that have contributed to some very hopeful emerging peptide modalities for intracellular target space. I have been blessed to have seen this phenomenal story take place from the time of the discovery of cyclosporine A (CsA) and my graduate studies in the field of peptide science at the University of Arizona which began in the mid-1970s. Although my earliest foray into what may now be defined as peptide drug discovery was actually focused on G protein-coupled receptors, I acquired a multidisciplinary mindset from my mentor, Professor Victor Hruby, which enabled me to do “the deep dive” into the abyss of intracellular space to tackle many different types of therapeutic targets over my career, including proteases (e.g., HIV-1 protease, interleukin-converting enzyme), phosphatases (e.g., PTP1b), transferases (e.g., Ras farnesyl transferase), kinases (e.g., Src), GTPases (e.g., K-Ras), apoptotic modulatory proteins (e.g., Mcl-1), and transcription/translation factors (e.g., p53, β -catenin, eIF4E, and Myc). This work has integrated and leveraged chemistry, biology, structural biology, biophysical chemistry, computational chemistry, cell biology, and pharmacology in rather fascinating ways to develop both tools and rules to design novel peptides having cell permeability, stability, and in vivo efficacy that may be further advanced as clinical candidates.

Relative to the varying intracellular targets which I have abovementioned, the diversity of peptide modalities that have been advanced as key tools or preclinical and/or clinical development candidates include peptidomimetics (including de novo designed nonpeptides), macrocyclic peptides (incorporating α -helical, reverse-turn, and β -strand motifs), and both classic and highly modified linear peptides (incorporating unusual amino acids, backbone surrogates, and/or non-amino acid building blocks) and, with increasing focus, conjugates thereof with other therapeutic modalities. Several examples of such diversity will be shared in this chapter. Of course, CsA well-illustrates a macrocyclic peptide incorporating unusual amino acids (e.g., non-canonical side chains and a D-isomer) and multiple N-methylation of the peptide backbone. In fact, CsA has inspired many academic, biotech, and pharma

efforts to leverage macrocyclic peptides for intracellular targets with respect to seeking CsA-like passive permeability properties. Such endeavors have led to a deeper understanding of biophysical, conformation, and structural principles correlating with cellular uptake for such macrocyclic peptides. It will be the overarching goal of this chapter to highlight key learnings from peptide drug discovery efforts that have been especially exploited by innovative macrocyclization design concepts and platforms.

8.2 Intracellularly Targeted Peptides: Some Historical Milestones

In retrospect, drug discovery efforts to advance intracellularly targeted peptides have gone through many different pathways, both conceptually and experimentally, in terms of how to traverse the cell membrane (*vide infra*). Consequently, the design of peptides having cell permeability properties and the ability to modulate intracellular therapeutic targets has required in most cases a *tour de force* to successfully advance peptides, including peptidomimetics and de novo designed nonpeptides.

Cyclosporine A (CsA) The macrocyclic peptide, natural product CsA (Fig. 8.1) has several chemical attributes to understand its structure-activity and structure-permeability relationships. The N-to-C backbone cyclized structure of CsA includes several N-methylated amino acids and one D-amino acid (Stahelin 1996). It exhibits passive permeability, despite having >500 molecular weight, and such has been attributed to its conformational flexibility and ability to exhibit intramolecular

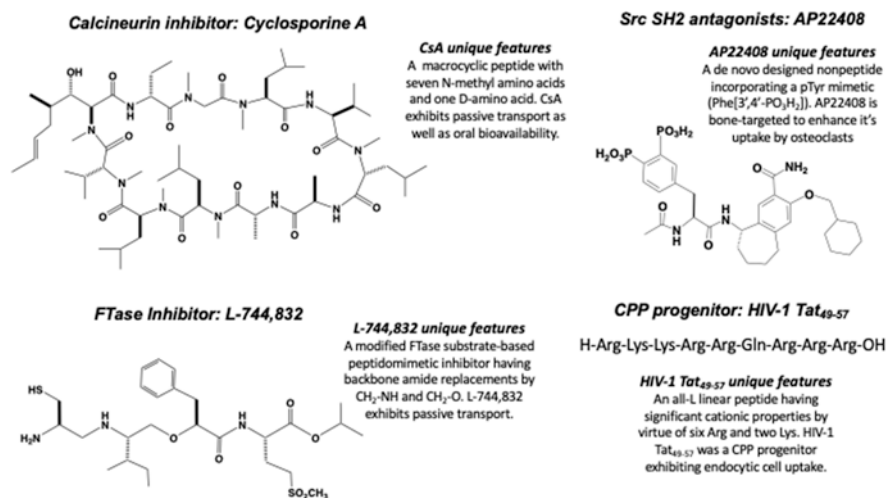


Fig. 8.1 Chemical structures of some historical intracellularly targeted peptides and peptidomimetics including CsA, AP22408, HIV-1 TAT, U-81749, bortezomib, and grazoprevir

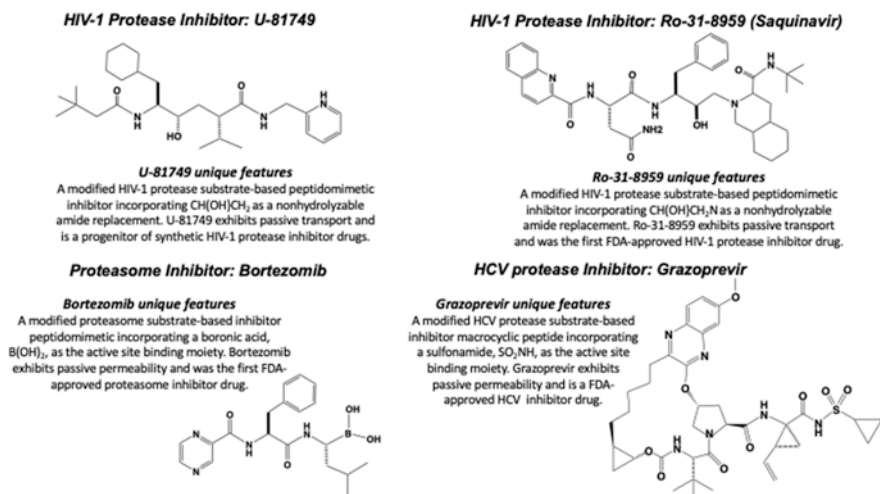


Fig. 8.1 (continued)

H-bonding favorable to lipid membrane interaction as well as intermolecular H-bonding with water. The impact of CsA on the field of intracellularly targeted peptide drug discovery has been extraordinary, especially with respect to its intrinsic cell permeability properties. In fact, CsA is a key benchmark macrocyclic peptide for passive transport (e.g., cell permeability and/or oral bioavailability) for macrocyclic peptide drug discovery efforts (Nielsen et al. 2017; Naylor et al. 2017; Pye et al. 2017; Chatterjee et al. 2012; Kelly et al. 2021; Naylor et al. 2018). Mechanistically, CsA binds to the cytosolic protein cyclophilin (also known as immunophilin) within lymphocytes (e.g., T cells). The CsA–cyclophilin complex then inhibits calcineurin and subsequent calcineurin-dependent production of interleukin-2 (Azzi et al. 2013).

Src SH2 Antagonist, AP22408 In the early 1990s, unique noncatalytic domains were identified with the first being the tyrosine kinase and Src and then for many intracellular proteins, including kinases, phosphatases, and adapter proteins such as Grb-2 (Sawyer et al. 2002). Historically, the biotech company Ariad Pharmaceuticals was founded to explore the signal transduction role of such Src homology (SH) domains, including both SH2 and SH3. In the specific case of SH2 domains, the cognate ligand was determined to be phosphotyrosine (pTyr) containing proteins and with sequence specificity about the pTyr residue, particularly those immediately C-terminal to it (Sawyer et al. 2002). Noteworthy was a series of novel peptidomimetic and de novo nonpeptide designs that ultimately led to the first potent, in vivo active Src SH2 antagonist AP22408 (Fig. 8.1) (Bohacek et al. 2001; Shakespeare et al. 2000). Based on a co-crystal structure of the noncatalytic Src homology 2 (SH2) domain of Src complexed with citrate in the phosphotyrosine (pTyr) binding pocket, the design of a novel 3',4'-diphosphonophenylalanine (Dpp) as a pTyr

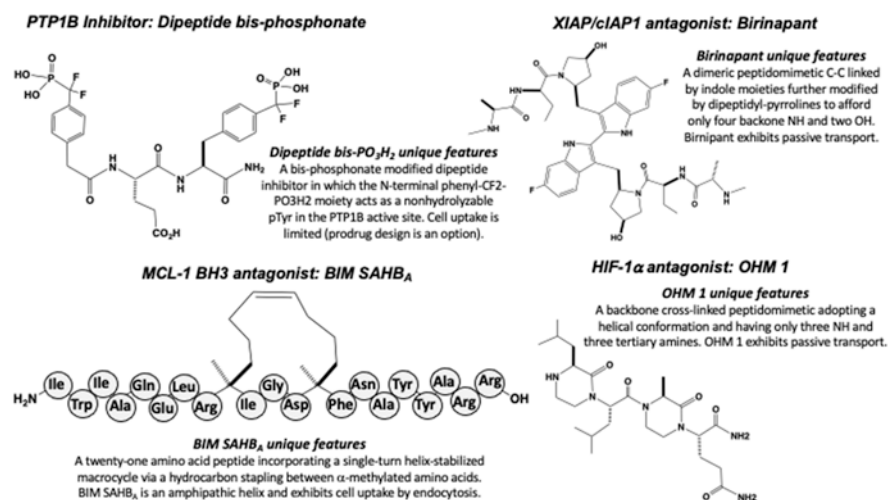


Fig. 8.2 Chemical structures of some intracellularly targeted peptide modalities focused on PTP1B, XIAP, MCL-1, MDM2/4, K-Ras, NEMO, CTFR-CAL, and HIF/p300

mimic was achieved. AP22408 also incorporates a bicyclic nonpeptide template to replace the tripeptide sequence C-terminal to the pTyr.

Ras Farnesyl Transferase Inhibitor, L-744,832 The highly coveted cancer target family of Ras proteins requires lipid modification by farnesyl isoprenoid by the farnesyltransferase (FTase) as a primary pathway and by an alternative process involving geranylgeranyltransferase (GGTase). Both enzymes are capable of effecting prenylation of Ras proteins as a so-called CAAX motif in an irreversible manner at the tetrapeptide's cysteine sulfhydryl group (Appels et al. 2005). Albeit the rational design of FTase inhibitors has successfully generated many potent molecules, including clinical candidates, this strategy has not shown efficacy against KRAS-driven cancers in humans. Hence, the pursuit of dual-specific inhibitors of both FTase and GGTase became a new strategy for next-generation clinical candidates, assuming they may overcome toxicity limitations (Appels et al. 2005). Exemplary of designed peptidomimetic CAAX-based FTase inhibitors is L-744,832 (Fig. 8.2) which was shown to be effective in combination with taxane-induced mitotic arrest and apoptosis in breast cancer cell lines (Lobell et al. 2002).

Cell-Penetrating Peptide (CPP) Progenitor, HIV-1 TAT₄₉₋₅₇ Investigations on HIV-1 showed that the TAT protein contained a cell membrane transduction motif enabling permeability which might be exploited as a carrier modality if conjugated to drug payloads (Vives 2003). The HIV-1 Tat₄₉₋₅₇ peptide RKKRRQRRR (Fig. 8.1) was the progenitor of what is now a superfamily (>100) of CPPs and the first of major subclass that is structurally characterized as having arginine-rich sequences. Other noteworthy CPPs discovered subsequently included antennapedia homeodo-

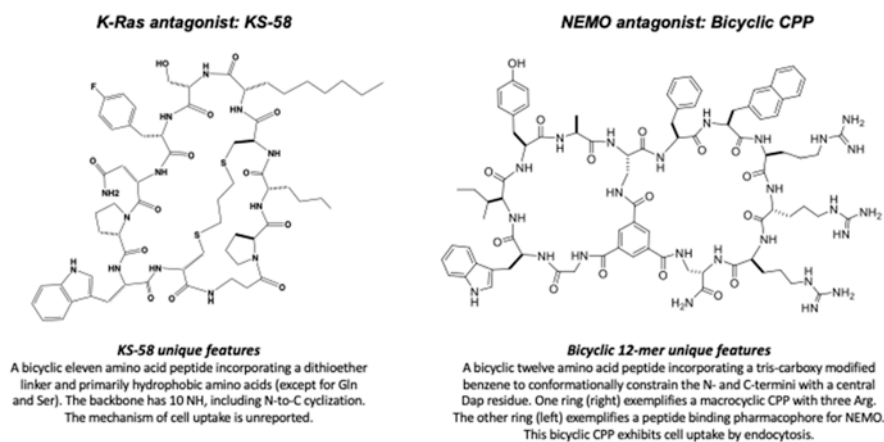


Fig. 8.2 (continued)

main protein_{43–58}, viral protein VP22_{267–300}, and nuclear localization signal sequences (Pooga and Langel 2015; Milletti 2012; Koren and Torchilin 2012). Furthermore, beyond the more widely used term CPPs, other names are found in the literature protein transduction domains (PTDs) and membrane translocating sequences (MTSs). All may be classified into three groups: (i) basic peptides such as Tat peptide, (ii) basic/amphiphilic peptides such as Antp, and (iii) hydrophobic peptides such as MTS (Futaki et al. 2003). More recently, a new class of hybrid macrocyclic CPPs has been developed (Appiah Kubi and Pei 2020) to further expand the potential therapeutic utility of this modality (vide infra).

HIV-1 Protease Inhibitors, U-81749 and Saquinavir One of the most promising targets that were first identified as critical to HIV-1 cellular infection and processing to enable replication of the retrovirus was an aspartyl protease, namely HIV-1 protease (Debouck et al. 1987). Noteworthy, HIV-1 protease was unique in that it was a C-2 symmetric homodimer with its two catalytic aspartyl residues being part of an active site created upon homodimerization of the two relatively small-sized (99 amino acids) monomers. HIV-1 protease inhibitor drug discovery became a worldwide effort throughout the 1990s (Debouck 1992). The first reported synthetic peptidomimetic inhibitor of HIV-1 protease exhibiting cellular activity was U-81749 (Fig. 8.1) (McQuade et al. 1990). It was essentially a tripeptide template incorporating a nonhydrolyzable amide isostere (i.e., CH(OH)CH₂) that was exemplary of diverse amide bond surrogates that were designed to advance highly potent peptidomimetic as well as novel nonpeptide inhibitors of HIV-1 protease (Ghosh et al. 2016; Roberts et al. 1990). The peptidomimetic saquinavir was the first HIV-1 protease inhibitor that was FDA-approved for the treatment of HIV infection in AIDS patients.

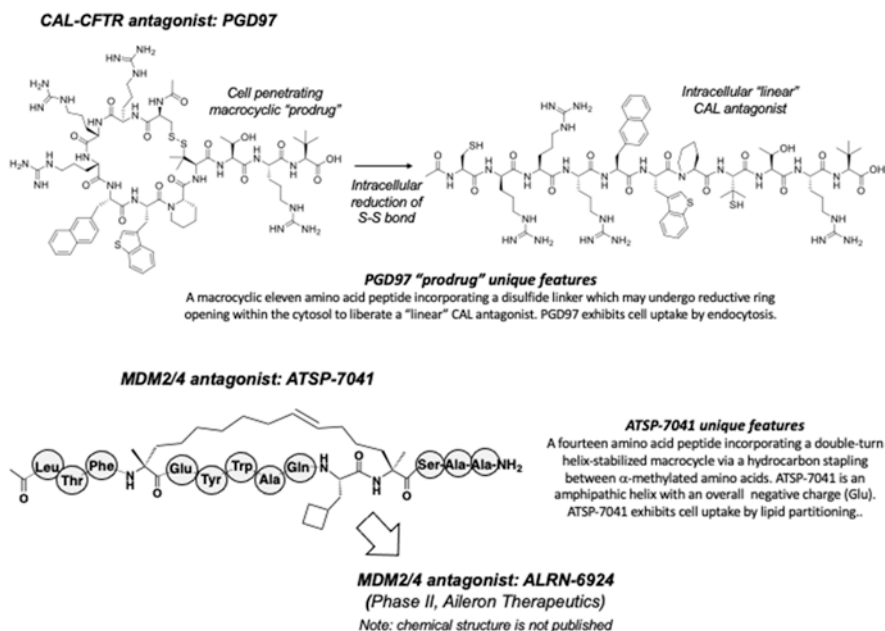


Fig. 8.2 (continued)

Bortezomib, the Proteasome Inhibitor A rather unique intracellular protease is the proteasome, and by way of the well-known ubiquitin-proteasome pathway, there exists targeted destruction of cellular proteins. With respect to cell cycle regulation and both cell proliferation and survival, especially in cancer cells, the proteasome was first recognized as a compelling therapeutic target for cancer cell therapy (Fogli et al. 2021). The first proteasome inhibitor that was advanced into clinical trials was bortezomib (Fig. 8.1), a boronic acid-containing peptidomimetic that was designed to effectively inhibit the serine protease active site of the proteasome (Adams 2002, 2004). Bortezomib was highly potent ($K_i < 1$ nM) and effective across a broad range of cancer cell lines.

HCV Protease Inhibitor, Grazoprevir A breakthrough for the treatment of chronic hepatitis C was achieved by treatment by a combination of NS3/4A protease inhibitors and NS5A inhibitors such as exemplified by grazoprevir (Fig. 8.1) and elbasvir, respectively (Matthew et al. 2020). The structure-based design of grazoprevir illustrates drug design focus on the substrate-binding active sites of NS3/4A to achieve optimal molecular recognition for both increased potency and decreased resistance (Harper et al. 2012).

8.3 Expanding Intracellular Target Space: Emerging Peptide Modalities

In the beginning, the most compelling therapeutic targets for peptide drug discovery were receptors (e.g., G protein-coupled receptors). However, as advancements in molecular and cell biology were being achieved during the latter half of the twentieth century, it was increasingly obvious that a universe of intracellular targets existed for the discovery of peptide modalities. Varying approaches, including the generation of synthetic or biological peptide libraries, target reporter cellular screening, and structural biology (X-ray, NMR, HDX-MS, and, more recently, cryo-EM) are each significantly contributing to an expanding treasury of intracellular therapeutic targets. Several peptide modalities have been advanced to interrogate intracellular target space. Such efforts generally first identify high-affinity binding leads that may then be tested in cellular assays to initiate the early lead optimization process. Unquestionably, this is where the proverbial “rubber hits the road” in terms of translating intracellular target druggability from binding to cellular efficacy as well as cellular permeability and cellular metabolic stability.

The most powerful approaches that are enabling such lead identification efforts include super-diverse macrocyclic peptide library screening derived from synthetic (e.g., one-bead, one-compound) or biologic (e.g., phage-, mRNA-, or DNA-display) technologies (Qian et al. 2015; Chen and Heinis 2015; Bashiruddin and Suga 2015; Zhu et al. 2018; Appiah Kubi et al. 2019). Exemplary peptide and peptidomimetic modalities, including macrocyclic and CPP-conjugate peptides, are described (vide infra) relative to several intracellular targets (e.g., PTP1B, XIAP-BIR3, Mcl-1 BH3, p53-MDM2/4, Ras GTPase–Raf, NEMO, CTFR-CAL, and HIF-1 α).

PTP1B Phosphatase The tyrosine phosphatase PTP1B is ubiquitously expressed, including in tissues such as the liver, muscles, and fat that are responsive to insulin (Zhang and Zhang 2007; Zinker et al. 2002). There exists a great deal of biochemical, genetic, and pharmacological evidence implicating PTP1B as a negative regulator in insulin signaling. Specifically, PTP1B can interact with and dephosphorylate activated insulin receptor or insulin receptor substrates. Aberrant expression of PTP1B can contribute to diabetes and obesity in humans. Antisense-based oligonucleotides targeting PTP1B have advanced to clinical trials and have shown efficacy in type 2 diabetes (Elchebly et al. 1999). A highly potent peptidomimetic inhibitor (Fig. 8.2) of PTP1B has been discovered (Shen et al. 2001) which leveraged a nonhydrolyzable pTyr moiety and optimized interactions at adjacent binding pockets to the active site of PTP1B. Noteworthy is the potency and high selectivity of this peptidomimetic for PTP1B versus other phosphatases. Such PTP1B inhibitors may enhance insulin signaling and augment insulin-stimulated glucose uptake.

XIAP-BIR3 The X-chromosome-linked inhibitor of apoptosis (XIAP) and cytosolic inhibitor of apoptosis (cIAP) represent a family of proteins that act as inhibitors of caspases by direct interaction through their baculoviral IAP repeat (BIR)

domains with caspases. Peptide, peptidomimetic, and small-molecule drug discovery has been focused on the BIR domains to identify inhibitors that would displace caspase-bound XIAP proteins (Abbas and Larisch 2020). TL32711 (Birinapant) is a novel bivalent peptidomimetic inhibitor (Fig. 8.2) that shows preferential binding to cIAP1 versus cIAP2 and XIAP, and it has advanced to clinical trials (Seigal et al. 2015).

MCL-1 The BH3 domain of proapoptotic intracellular protein BIM can bind to the BH3 hydrophobic groove of BCL-2 antiapoptotic proteins and directly activate the apoptotic effector proteins BAK and BAX. The hydrocarbon-stapled α -helical peptide BIM-SAHB_A (Fig. 8.2) was designed relative to replacing a key salt bridge in an i to $i + 4$ manner to incorporate alpha-methylated amino acids having terminal alkene moieties on each of the two amino acid side chains that would undergo ring-closing metathesis (Edwards et al. 2015). BIM-SAHB_A was found to primarily bind the intracellular antiapoptotic BCL-2 family protein MCL-1. Specificity studies showed MCL-1 knockout mouse embryonic fibroblasts were resistant to apoptosis induced by BIM-SAHB_A (Hadji et al. 2019).

HIF-1 α Cells express a family of hypoxia-inducible transcription factors (HIFs) when under a state of reduced oxygen levels. HIFs are heterodimeric basic helix–loop–helix proteins composed of a regulatory α (HIF1 α) and a constitutively expressed β (HIF1 β ; also termed ARNT) subunit. Furthermore, the C-terminal transactivation domain (CTAD) of HIF1 α interacts with coactivator protein p300 (or its ortholog CREB binding protein [CBP]) (Semenza 2012). It is the HIF/p300 complex which then mediates transactivation of hypoxia-inducible genes that play major roles in cancer (e.g., angiogenesis, invasion, and altered energy metabolism (REF)). Relative to the discovery of inhibitors of the protein–protein interaction between HIF and p300, the structure-based design of novel oxopiperazine helical mimetic (OHM) HIF inhibitor OHM-1 (Fig. 8.2) was found to exhibit high binding affinity to p300 and both cellular and in vivo efficacy to reduce tumor burden in a triple-negative breast cancer xenograft mouse model (Lao et al. 2014). Noteworthy is the chiral OHM peptidomimetic template which bridges peptide backbone NH groups via ethylene to conformationally constrain OHM-1 into α -helix.

K-Ras GTPase Another sought-after cancer target is that of Ras mutations, such as those occurring in K-Ras, N-Ras, and H-Ras (Khan et al. 2020). Recently, a small-molecule drug, sotorasib, which covalently forms a covalent bond to Cys-12 of mutated K-Ras, has been FDA approved for K-Ras G12C-mutated lung cancer (Skoulidis et al. 2021). A more desirable drug modality that may be able to target a greater range of mutated Ras in tumors remains a focus of intense research worldwide. A compelling macrocyclic peptide, KRpep-2d, that was first discovered from phage-display libraries was found to be the first selective inhibitor of K-Ras G12D, a predominant K-Ras mutation (Sakamoto et al. 2017). Subsequent lead optimization studies led to the bicyclic peptide, KS-58 (Fig. 8.2), that was shown to be active in vivo against K-Ras G12D-driven human pancreatic tumor xenografts (Sakamoto

et al. 2020). It was also shown that a combination of KS-58 with gemcitabine resulted in enhanced antitumor activity. KS-58 incorporates N-to-C cyclization and a dithioether bridge between two Cys residues to create the conformationally constrained bicyclic peptide inhibitor.

NEMO Nuclear factor kappa B (NF- κ B) represents a family of transcription factors involved in the regulation of immune response, inflammation, cell differentiation, and cell survival (Zhang et al. 2017). Two different signaling pathways, one canonical and one non-canonical, lead to the NF- κ B activation. With respect to canonical NF- κ B signaling, cellular receptor activation results in an active inhibitor of κ B (I κ B) kinase (IKK) complex which consists of IKK α , IKK β , and NEMO (*a.k.a.* IKK γ). Relative to IKK, inhibition of NEMO interaction is viewed compelling for anti-inflammatory and anticancer strategies as it may not modify basal NF- κ B activity required for normal B- and T-cell function (Baima et al. 2010). Exploiting the concept that macrocyclic peptides may bind challenging protein–protein interactions, the design of a bicyclic NEMO-targeted peptide (Fig. 8.2) simultaneously incorporating cell-permeability properties was shown to effectively inhibit NEMO–IKK β interaction as well as exhibit inhibition of canonical NF- κ B signaling in mammalian cells and the proliferation of cisplatin-resistant ovarian cancer cells (Rhodes et al. 2018). The NEMO inhibitor incorporated a 1,3,5-tricarboxy-benzene moiety to provide a cross-linking bicyclic structure, and its cell permeability properties correlated with its cationic substructure (i.e., Arg residues) as well as its hydrophobicity and conformational constraints to ultimately confer partitioning into lipid membranes and triggering of endocytosis to drive cellular uptake.

CFTR-CAL Mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene, which encodes for a chloride ion channel, is causative of cystic fibrosis (Dehecchi et al. 2018). Membrane expression of CFTR is negatively regulated by CFTR-associated ligand (CAL). Therefore, designing an inhibitor of CAL may rescue mutant CFTR function. Recently, the macrocyclic peptide PGD97 (Fig. 8.2) incorporating a disulfide capable of intracellular reduction was found (Dougherty et al. 2020) to have potent (low nM) and selective binding to CAL, good stability in serum, and efficacy in mutant F508del homozygous cells to increase short circuit currents as well as potential therapeutic effects of small-molecule correctors (e.g., VX-661). Therefore, PGD97 exemplifies a promising lead for the treatment of cystic fibrosis. This work provides incentive to further design strategies to create macrocyclic peptide “prodrugs” exploiting intracellular S-S ring opening to enable target binding properties.

p53–MDM2/4 The human transcription factor protein p53, and so-called guardian of the genome, is well known to induce cell cycle arrest and apoptosis in response to DNA damage and cellular stress and therefore has a critical role in protecting cells from malignant transformation (Eskandari et al. 2021; Ng et al. 2018; Carvajal et al. 2018). Inactivation of p53 by deletion or mutation or through overexpression of inhibitory proteins is most common in human cancers. Furthermore, cancers that

overexpress the suppressor proteins MDM2 and MDMX, but have wild-type p53, provide the opportunity to restore p53-dependent cell cycle arrest and apoptosis if the MDM2 and MDM4 may be effectively blocked by inhibitors. Both aberrant MDM2 overexpression and gene amplification as well as that of MDM4 exist in many tumors. The first potent and selective small-molecule inhibitors of MDM2, the so-called Nutlins (Vassilev 2005), provided proof of concept that restoration of p53 activity is a promising approach to cancer therapy. Nevertheless, these and other small-molecule efforts were limited to only MDM2 specificity, and essentially all were inactive against MDM4. In contrast, p53 mimicry via the design of the potent and *in vivo* effective stapled α -helical peptide ATSP-7041 (Chang et al. 2013) (Fig. 8.2) exemplified a key benchmark for this peptide modality as related to its collective structural features (e.g., α -methyl-amino acids and cyclization by hydrocarbon stapling), cell uptake properties (e.g., lipophilic partitioning and translocation from membrane to the cytosolic compartment), metabolic stability, and pharmacokinetic properties (Sawyer et al. 2018).

8.4 Unlocking the Secrets of Cell Permeability: Exploiting Innovative Tools

Of no surprise, over the past two decades there has been a profound focus on peptide modalities for intracellular targets in terms of drug design strategies that are becoming increasingly sophisticated in terms of exploiting innovative tools to tackle cell permeability and, with yet greater aspiration, oral bioavailability (Rosania and Thurber 2021; Hochman et al. 2021; Peier et al. 2021; Peraro et al. 2018; Furukawa et al. 2016, 2020; Sahni et al. 2020; Qian et al. 2014; Dougherty et al. 2019; Schwochert et al. 2016; Ahlbach et al. 2015; Bockus et al. 2015; Hewitt et al. 2015; Goetz et al. 2014; Wang et al. 2015; Aubry et al. 2010; Gordon et al. 2016). Such efforts include an increasing number of macrocyclic peptides being advanced with academia, biotech, and pharma focused on intracellularly targeted drug discovery and, along with it, oral bioavailability (Nielsen et al. 2015; Rafi et al. 2012; Guimaraes et al. 2012; Herce et al. 2014; Rezgui et al. 2016; LaRochelle et al. 2015). It is anticipated that accumulation of data from this work will enable QSAR models and predictive design in the future to exploit peptide modalities as novel therapeutics. Key physicochemical and biophysical properties of structurally diverse peptide and peptidomimetics can be systematically analyzed to support lead optimization (Table 8.1). Some properties that are being woven into QSAR models and predictive design strategies include molecular weight (e.g., 500–2000 dalton range), lipophilicity (experimental and/or calculated LogP and typically in the 2–5 range), H-bond donor and acceptors (typically seeking less H-bond donors to solvent via intramolecular H-bonding and/or masking by N-methylation), and polar surface area ([PSA], typically seeking lower PSA) (Holm et al. 2011). In this regard, benchmark peptides such as CsA and/or other well-characterized macrocyclic peptides

having cell permeability properties by varying mechanisms would be highly recommended.

As illustrated in this chapter, the design of peptides and peptidomimetics is one of the most intriguing opportunities nowadays because of the opportunistic availability of chemically diverse amino acid building blocks as well as novel conformational constraints by backbone modifications and/or macrocyclization. Such unnatural amino acids include D-enantiomers, N α -methylated amino acids, cyclic amino acids, C α -methylated amino acids, β -amino acids, and a host of novel side chain modifications for many of these building blocks. Indeed, conformational diversity is deemed especially critical to such multifaceted design strategies for both target binding, cellular permeability, and metabolic stability as related to macrocyclic peptides of varying ring size, bicyclization, and related innovative chemistries for cyclization that may impact their overall physicochemical and biophysical properties.

As a relatively simplistic model to map the likely multiple ways by which peptides and peptidomimetics may achieve cell uptake, it is apparent that three major mechanisms for cell permeability include passive transport, lipophilic partitioning, and cationic partitioning (Fig. 8.3). First, in the case of CsA and an emerging constellation of CsA-like macrocyclic peptides that are achieving cell permeability via passive transport (Furukawa et al. 2016, 2020; Schwochert et al. 2016; Ahlbach et al. 2015; Bockus et al. 2015; Hewitt et al. 2015), there is great promise for this peptide modality to tackle targets which have been deemed virtually undruggable with small molecules. Likewise, highly modified peptidomimetics such as previously exemplified by those targeting HIV-1 protease, proteasome, HCV protease, and HIF/p300 (vide supra) provide framework for designing such molecules. Second, in the case of ATSP-7041 (vide supra) and other stapled α -helical peptides targeting intracellular protein–protein interactions (Levin 2015; Guerlavais and Sawyer 2014; Sawyer et al. 2018; Chang et al. 2013; Peier et al. 2021), the design of amphipathic molecules having a high propensity to partition into cell membranes and then undergo translocation to the cytosol is being recognized as another mechanism of cell uptake that has similar attributes to passive transport albeit yet different. In contrast to what may be considered somewhat counter-intuitive in terms of its physicochemical and biophysical properties, ATSP-7041 is both a lipophilic and quite soluble peptide of which the latter may be attributed to a single Glu within its sequence. Third, in the case of Arg-rich linear and macrocyclic CPPs (Appiah Kubi and Pei 2020; Rhodes et al. 2018; Dougherty et al. 2019, 2020; Sahni et al. 2020; Qian et al. 2014; Herce et al. 2014; LaRochelle et al. 2015; Holm et al. 2011), this peptide modality is well recognized to leverage cationic partitioning into cell membranes and undergo delivery into the cytosolic compartment via endocytic mechanisms. Obviously, the requirement for endosomal escape is critical to the optimization of such CPPs, and as exemplified by the macrocyclic CPP inhibitors of NEMO and CFTR-CAL (vide supra), there is significant promise to both intracellular targeted design and to create novel conjugates with other modalities (e.g., peptide, protein, and oligonucleotide) as is being pursued by Entrada Therapeutics.

Table 8.1 Some computational, biophysical, and cell-based screening tools to explore cell permeability as well as enable peptide and peptidomimetic lead optimization

Permeability screening tool	Some key features and properties evaluated
Exposed polar surface area (EPSA)	Experimental EPSA values are determined using supercritical fluid chromatography Low EPSA values have been shown to correlate with high passive permeability and predicted oral bioavailability
ΔG (insertion)	ΔG (insertion) is a calculated value that refers to the solvent-free difference for transferring peptide in a low-dielectric conformation (LDC) from water to a low-dielectric environment (membrane-like) and is predictive of passive permeability
Parallel artificial membrane permeability assay (PAMPA)	PAMPA uses mixtures of phospholipids infused into lipophilic microfilters with a net negative charge (surrogate model system to correlate with experimental bioavailability)
Lipid: water phase partitioning	Both octanol: water phase partitioning and recent modification to incorporate fatty acid and pH gradient as shown for a guanidinium-rich peptide to be predictive of energy-independent translocation
Cell monolayer transcytosis	Caco-2 cells, or other cell types, are used to measure passive permeability from donor to acceptor compartments to correlate with in vivo oral bioavailability
Cell-based, target agnostic penetration assay (CAPA, NanoClick)	CAPA and nanoClick measure cytosolic delivery of specifically tagged peptides (chloroalkyl and azide, respectively) into cells that stably expresses a haloenzyme-reporter fusion protein (i.e., fusion with GFP and luciferase, respectively)
Radius of gyration (R-gyr)	R-gyr is calculated as the root-mean-square distance between a peptide's atoms and its center of gravity. R-gyr is an alternative property for MW for beyond-Rule-of-5 (bRo5) molecules
NMR analysis of intramolecular versus solvent H-bonding	Solution NMR studies using hydrogen/deuterium (H/D) exchange to determine peptide backbone amide temperature coefficients and intramolecular hydrogen bonding
Label-free mass spectrometric and fluorescently tagged cell uptake analysis	Direct measurement methods for cell uptake of peptides using MS methods and/or imaging studies (e.g., fluorescence correlation microscopy) to quantitate intracellular exposure

Adapted from Sawyer (2017) with permission from the Royal Society of Chemistry

Drug delivery remains integral to the future development of intracellularly targeted peptide and peptidomimetic therapeutics (Lemmer and Hamman 2013; Maher et al. 2016; Danielsen 2021; Brayden et al. 2020; Di 2015; Rader et al. 2018; Zizzari et al. 2021). Opportunities here include varying routes of administration, such as intravenous, subcutaneous, oral, and nasal. Collectively, biophysical, pharmacokinetic (PK), and absorption-distribution-metabolism-excretion properties remain extremely important for translating preclinical lead molecules to clinical candidates. Such properties may generally be “target agnostic” and have more to do with optimizing solubility, permeability, stability, and exposure levels in vivo to enable possible correlation between pharmacological efficacy and PK/ADME properties.

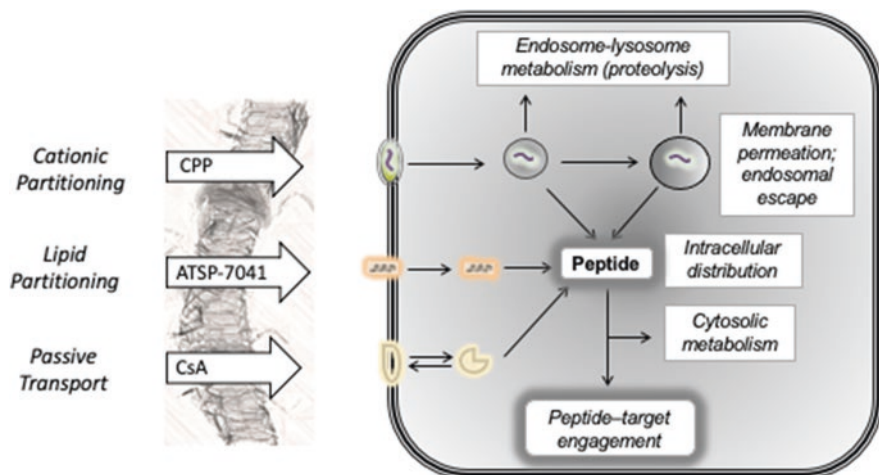


Fig. 8.3 Lipophilic partitioning, cationic partitioning, and passive transport models of cell permeability of intracellularly targeted peptides and peptidomimetics

Ultimately, the oral bioavailability potential to exploit to the high diversity of peptide modalities may also leverage formulations with permeability enhancers for either transcellular or paracellular transport.

8.5 Future Intracellularly Targeted Peptide Drugs: Clinical Trials and Beyond

Several intracellularly targeted peptidomimetics and peptides have been successfully advanced into clinical trials and/or FDA-approved. They include many FDA-approved peptidomimetics targeting HIV-1 protease (e.g., saquinavir by Hoffmann-La Roche, ritonavir by AbbVie, indinavir by Merck & Co., nelfinavir by Hoffman-La Roche, amprenavir by GlaxoSmithKline, lopinavir by AbbVie, atazanavir by Bristol-Myers Squibb, fosamprenavir by GlaxoSmithKline, tipranavir by Boehringer Ingelheim, and darunavir by Janssen Therapeutics), proteasome (e.g., bortezomib by Millennium Pharmaceuticals, carfilzomib by Onyx Pharmaceuticals, and ixazomib by Takeda), and HCV protease (e.g., grazoprevir by Merck, telaprevir by Vertex Pharmaceuticals and Johnson & Johnson, glecaprevir by AbbVie, and paritaprevir by Abbvie). Noteworthy is the Phase 2 clinical development of the stapled α -helical peptide ALRN-6924. Furthermore, numerous other macrocyclic peptides are undergoing intense preclinical development or are entering clinical trials for various intracellularly targets, including both enzymes and protein-protein interactions as described in this chapter and others (e.g., transcription factors β -catenin and Myc/Max). Lastly, albeit not intracellularly targeted in the classic sense, but cell membrane targeted in terms of known mechanisms of action, are the

FDA-approved antibiotic peptides (e.g., Orbactiv by the Med Company, Dalvance by Vicuron, and Cubicin by Cubist/MSD).

Beyond the realm of specific preclinical and clinical development of promising novel peptide and peptidomimetic therapeutics for intracellular targets, it is of the utmost importance to also highlight the fact that many biotech companies have contributed significantly in terms of innovative platforms to advance such peptide modality-inspired medicines. Examples of these companies include (A–Z listing) Aileron Therapeutics, Aplomex, Circle Pharma, Entrada Therapeutics, FAKnostics, Fog Pharma, IDP Pharma, Nimble Therapeutics, Orbit Discovery, PeptiDream, Polyphor, Promakhos Therapeutics, ProteXase Therapeutics, Ra Pharma (now UCB), SyntheX, Spotlight Therapeutics, and Unnatural Products.

References

- Abbas R, Larisch S. Targeting XIAP for promoting cancer cell death—the story of ARTS and SMAC. *Cells*. 2020;9:663.
- Adams J. Development of the proteasome inhibitor PS-341. *Oncologist*. 2002;7:9–16.
- Adams J. The development of proteasome inhibitors as anticancer drugs. *Cancer Cell*. 2004;5:417–21.
- Ahlbach CL, Lexa KW, Bockus AT, Chen V, Crews P, Jacobson MP, Lokey RS. Beyond cyclosporine A: conformation-dependent passive membrane permeabilities of cyclic peptide natural products. *Future Med Chem*. 2015;7:2121–30.
- Appels NM, Beijnen JH, Schellens JH. Development of farnesyl transferase inhibitors: a review. *Oncologist*. 2005;10:565–78.
- Appiah Kubi G, Pei D. Cell-penetrating and mitochondrion-targeting molecules. *Methods Enzymol*. 2020;641:311–28.
- Appiah Kubi G, Dougherty PG, Pei D. Designing cell-permeable macrocyclic peptides. *Methods Mol Biol*. 2019;2001:41–59.
- Aubry S, Aussedat B, Delaroche D, Jiao CY, Bolbach G, Lavielle S, Chassaing G, Sagan S, Burlina F. MALDI-TOF mass spectrometry: a powerful tool to study the internalization of cell-penetrating peptides. *Biochim Biophys Acta*. 2010;1798:2182–9.
- Azzi JR, Sayegh MH, Mallat SG. Calcineurin inhibitors: 40 years later, can't live without. *J Immunol*. 2013;191:5785–91.
- Baima ET, Guzova JA, Mathialagan S, Nagiec EE, Hardy MM, Song LR, Bonar SL, Weinberg RA, Selness SR, Woodard SS, Chrencik J, Hood WF, Schindler JF, Kishore N, Mbalaviele G. Novel insights into the cellular mechanisms of the anti-inflammatory effects of NF-kappaB essential modulator binding domain peptides. *J Biol Chem*. 2010;285:13498–506.
- Bashiruddin NK, Suga H. Construction and screening of vast libraries of natural product-like macrocyclic peptides using in vitro display technologies. *Curr Opin Chem Biol*. 2015;24:131–8.
- Bockus AT, Lexa KW, Pye CR, Kalgutkar AS, Gardner JW, Hund KC, Hewitt WM, Schwochert JA, Glassey E, Price DA, Mathiowetz AM, Liras S, Jacobson MP, Lokey RS. Probing the physicochemical boundaries of cell permeability and Oral bioavailability in lipophilic macrocycles inspired by natural products. *J Med Chem*. 2015;58:4581–9.
- Bohacek RS, Dalgarno DC, Hatada M, Jacobsen VA, Lynch BA, Macek KJ, Merry T, Metcalf CA 3rd, Narula SS, Sawyer TK, Shakespeare WC, Violette SM, Weigele M. X-ray structure of citrate bound to Src SH2 leads to a high-affinity, bone-targeted Src SH2 inhibitor. *J Med Chem*. 2001;44:660–3.

- Brayden DJ, Hill TA, Fairlie DP, Maher S, Mrsny RJ. Systemic delivery of peptides by the oral route: formulation and medicinal chemistry approaches. *Adv Drug Deliv Rev.* 2020;157:2–36.
- Carvajal LA, Neriah DB, Senecal A, Benard L, Thiruthuvanathan V, Yatsenko T, Narayanagari SR, Wheat JC, Todorova TI, Mitchell K, Kenworthy C, Guerlavais V, Annis DA, Bartholdy B, Will B, Anampa JD, Mantzaris I, Aivado M, Singer RH, Coleman RA, Verma A, Steidl U. Dual inhibition of MDMX and MDM2 as a therapeutic strategy in leukemia. *Sci Transl Med.* 2018;10:eaa03003.
- Chang YS, Graves B, Guerlavais V, Tovar C, Packman K, To KH, Olson KA, Kesavan K, Gangurde P, Mukherjee A, Baker T, Darlak K, Elkin C, Filipovic Z, Qureshi FZ, Cai H, Berry P, Feyfant E, Shi XE, Horstick J, Annis DA, Manning AM, Fotouhi N, Nash H, Vassilev LT, Sawyer TK. Stapled alpha-helical peptide drug development: a potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. *Proc Natl Acad Sci U S A.* 2013;110:E3445–54.
- Chatterjee J, Laufer B, Kessler H. Synthesis of N-methylated cyclic peptides. *Nat Protoc.* 2012;7:432–44.
- Chen S, Heinis C. Phage selection of bicyclic peptides based on two disulfide bridges. *Methods Mol Biol.* 2015;1248:119–37.
- Danielsen EM. Intestinal permeation enhancers: lessons learned from studies using an organ culture model. *Biochim Biophys Acta Biomembr.* 2021;1863:183474.
- Debouck C. The HIV-1 protease as a therapeutic target for AIDS. *AIDS Res Hum Retrovir.* 1992;8:153–64.
- Debouck C, Gorniak JG, Strickler JE, Meek TD, Metcalf BW, Rosenberg M. Human immunodeficiency virus protease expressed in *Escherichia coli* exhibits autoprocessing and specific maturation of the gag precursor. *Proc Natl Acad Sci U S A.* 1987;84:8903–6.
- Dechecchi MC, Tamanini A, Cabrini G. Molecular basis of cystic fibrosis: from bench to bedside. *Ann Transl Med.* 2018;6:334.
- Di L. Strategic approaches to optimizing peptide ADME properties. *AAPS J.* 2015;17:134–43.
- Dougherty PG, Sahni A, Pei D. Understanding cell penetration of cyclic peptides. *Chem Rev.* 2019;119:10241–87.
- Dougherty PG, Wellmerling JH, Koley A, Lukowski JK, Hummon AB, Cormet-Boyaka E, Pei D. Cyclic peptidyl inhibitors against CAL/CFTR interaction for treatment of cystic fibrosis. *J Med Chem.* 2020;63:15773–84.
- Edwards AL, Wachter F, Lammert M, Huhn AJ, Luccarelli J, Bird GH, Walensky LD. Cellular uptake and ultrastructural localization underlie the pro-apoptotic activity of a hydrocarbon-stapled BIM BH3 peptide. *ACS Chem Biol.* 2015;10:2149–57.
- Elchebly M, Payette P, Michalyszyn E, Cromlish W, Collins S, Loy AL, Normandin D, Cheng A, Himms-Hagen J, Chan CC, Ramachandran C, Gresser MJ, Tremblay ML, Kennedy BP. Increased insulin sensitivity and obesity resistance in mice lacking the protein tyrosine phosphatase-1B gene. *Science.* 1999;283:1544–8.
- Eskandari M, Shi Y, Liu J, Albanese J, Goel S, Verma A, Wang Y. The expression of MDM2, MDM4, p53 and p21 in myeloid neoplasms and the effect of MDM2/MDM4 dual inhibitor. *Leuk Lymphoma.* 2021;62:167–75.
- Fogli S, Galimberti S, Gori V, Del Re M, Danesi R. Pharmacology differences among proteasome inhibitors: implications for their use in clinical practice. *Pharmacol Res.* 2021;167:105537.
- Furukawa A, Townsend CE, Schwochert J, Pye CR, Bednarek MA, Lokey RS. Passive membrane permeability in cyclic peptomer scaffolds is robust to extensive variation in side chain functionality and backbone geometry. *J Med Chem.* 2016;59:9503–12.
- Furukawa A, Schwochert J, Pye CR, Asano D, Edmondson QD, Turmon AC, Klein VG, Ono S, Okada O, Lokey RS. Drug-like properties in macrocycles above MW 1000: backbone rigidity versus side-chain lipophilicity. *Angew Chem Int Ed Engl.* 2020;59:21571–7.
- Futaki S, Goto S, Suzuki T, Nakase I, Sugiura Y. Structural variety of membrane permeable peptides. *Curr Protein Pept Sci.* 2003;4:87–96.

- Ghosh AK, Osswald HL, Prato G. Recent progress in the development of HIV-1 protease inhibitors for the treatment of HIV/AIDS. *J Med Chem.* 2016;59:5172–208.
- Goetz GH, Philippe L, Shapiro MJ. EPSA: a novel supercritical fluid chromatography technique enabling the design of permeable cyclic peptides. *ACS Med Chem Lett.* 2014;5:1167–72.
- Gordon LJ, Allen M, Artursson P, Hann MM, Leavens BJ, Mateus A, Readshaw S, Valko K, Wayne GJ, West A. Direct measurement of intracellular compound concentration by RapidFire Mass Spectrometry offers insights into cell permeability. *J Biomol Screen.* 2016;21:156–64.
- Guerlavais V, Sawyer TK. Advancements in stapled peptide drug discovery & development. *Annu Rep Med Chem.* 2014;49:331–45.
- Guimaraes CR, Mathiowetz AM, Shalaeva M, Goetz G, Liras S. Use of 3D properties to characterize beyond rule-of-5 property space for passive permeation. *J Chem Inf Model.* 2012;52:882–90.
- Hadji A, Schmitt GK, Schnorenberg MR, Roach L, Hickey CM, Leak LB, Tirrell MV, LaBelle JL. Preferential targeting of MCL-1 by a hydrocarbon-stapled BIM BH3 peptide. *Oncotarget.* 2019;10:6219–33.
- Harper S, McCauley JA, Rudd MT, Ferrara M, DiFilippo M, Crescenzi B, Koch U, Petrocchi A, Holloway MK, Butcher JW, Romano JJ, Bush KJ, Gilbert KF, McIntyre CJ, Nguyen KT, Nizi E, Carroll SS, Ludmerer SW, Burlein C, DiMuzio JM, Graham DJ, McHale CM, Stahlhut MW, Olsen DB, Monteagudo E, Cianetti S, Giuliano C, Pucci V, Trainor N, Fandozzi CM, Rowley M, Coleman PJ, Vacca JP, Summa V, Liverton NJ. Discovery of MK-5172, a macrocyclic hepatitis C virus NS3/4a protease inhibitor. *ACS Med Chem Lett.* 2012;3:332–6.
- Herce HD, Garcia AE, Cardoso MC. Fundamental molecular mechanism for the cellular uptake of guanidinium-rich molecules. *J Am Chem Soc.* 2014;136:17459–67.
- Hewitt WM, Leung SS, Pye CR, Ponkey AR, Bednarek M, Jacobson MP, Lokey RS. Cell-permeable cyclic peptides from synthetic libraries inspired by natural products. *J Am Chem Soc.* 2015;137:715–21.
- Hochman J, Sawyer T, Duggal R. Overcoming cellular and systemic barriers to design the next wave of peptide therapeutics. In: *Quantitative analysis of cellular drug transport, disposition, and delivery.* New York: Springer; 2021.
- Holm T, Andaloussi SE, Langel U. Comparison of CPP uptake methods. *Methods Mol Biol.* 2011;683:207–17.
- Kelly CN, Townsend CE, Jain AN, Naylor MR, Pye CR, Schworchert J, Lokey RS. Geometrically diverse lariat peptide scaffolds reveal an untapped chemical space of high membrane permeability. *J Am Chem Soc.* 2021;143:705–14.
- Khan I, Rhett JM, O'Bryan JP. Therapeutic targeting of RAS: new hope for drugging the “undruggable”. *Biochim Biophys Acta, Mol Cell Res.* 2020;1867:118570.
- Koren E, Torchilin VP. Cell-penetrating peptides: breaking through to the other side. *Trends Mol Med.* 2012;18:385–93.
- Lao BB, Grishagin I, Mesallati H, Brewer TF, Olenyuk BZ, Arora PS. In vivo modulation of hypoxia-inducible signaling by topographical helix mimetics. *Proc Natl Acad Sci U S A.* 2014;111:7531–6.
- LaRochelle JR, Cobb GB, Steinauer A, Rhoades E, Schepartz A. Fluorescence correlation spectroscopy reveals highly efficient cytosolic delivery of certain penta-arg proteins and stapled peptides. *J Am Chem Soc.* 2015;137:2536–41.
- Lemmer HJ, Hamman JH. Paracellular drug absorption enhancement through tight junction modulation. *Expert Opin Drug Deliv.* 2013;10:103–14.
- Levin JJ. *Macrocycles in drug discovery.* Cambridge: Royal Society of Chemistry; 2015.
- Lobell RB, Liu D, Buser CA, Davide JP, DePuy E, Hamilton K, Koblan KS, Lee Y, Mosser S, Motzel SL, Abbruzzese JL, Fuchs CS, Rowinsky EK, Rubin EH, Sharma S, Deutsch PJ, Mazina KE, Morrison BW, Wildonger L, Yao SL, Kohl NE. Preclinical and clinical pharmacodynamic assessment of L-778,123, a dual inhibitor of farnesyl:protein transferase and geranylgeranyl:protein transferase type-I. *Mol Cancer Ther.* 2002;1:747–58.
- Maher S, Mrsny RJ, Brayden DJ. Intestinal permeation enhancers for oral peptide delivery. *Adv Drug Deliv Rev.* 2016;106:277–319.

- Matthew AN, Zephyr J, Nageswara Rao D, Henes M, Kamran W, Kosovrasti K, Hedger AK, Lockbaum GJ, Timm J, Ali A, Kurt Yilmaz N, Schiffer CA. Avoiding drug resistance by substrate envelope-guided design: toward potent and robust HCV NS3/4A protease inhibitors. *mBio*. 2020;11:e00172–20.
- McQuade TJ, Tomasselli AG, Liu L, Karacostas V, Moss B, Sawyer TK, Heinrikson RL, Tarpley WG. A synthetic HIV-1 protease inhibitor with antiviral activity arrests HIV-like particle maturation. *Science*. 1990;247:454–6.
- Milletti F. Cell-penetrating peptides: classes, origin, and current landscape. *Drug Discov Today*. 2012;17:850–60.
- Naylor MR, Bockus AT, Blanco MJ, Lokey RS. Cyclic peptide natural products chart the frontier of oral bioavailability in the pursuit of undruggable targets. *Curr Opin Chem Biol*. 2017;38:141–7.
- Naylor MR, Ly AM, Handford MJ, Ramos DP, Pye CR, Furukawa A, Klein VG, Noland RP, Edmondson Q, Turmon AC, Hewitt WM, Schwochert J, Townsend CE, Kelly CN, Blanco MJ, Lokey RS. Lipophilic permeability efficiency reconciles the opposing roles of lipophilicity in membrane permeability and aqueous solubility. *J Med Chem*. 2018;61:11169–82.
- Ng SY, Yoshida N, Christie AL, Ghandi M, Dharia NV, Dempster J, Murakami M, Shigemori K, Morrow SN, Van Scoyk A, Cordero NA, Stevenson KE, Puligandla M, Haas B, Lo C, Meyers R, Gao G, Cherniack A, Louissaint A Jr, Nardi V, Thorner AR, Long H, Qiu X, Morgan EA, Dorfman DM, Fiore D, Jang J, Epstein AL, Dogan A, Zhang Y, Horwitz SM, Jacobsen ED, Santiago S, Ren JG, Guerlavais V, Annis DA, Aivado M, Saleh MN, Mehta A, Tsherniak A, Root D, Vazquez F, Hahn WC, Inghirami G, Aster JC, Weinstock DM, Koch R. Targetable vulnerabilities in T- and NK-cell lymphomas identified through preclinical models. *Nat Commun*. 2018;9:2024.
- Nielsen DS, Lohman RJ, Hoang HN, Hill TA, Jones A, Lucke AJ, Fairlie DP. Flexibility versus rigidity for orally bioavailable cyclic hexapeptides. *Chembiochem*. 2015;16:2289–93.
- Nielsen DS, Shepherd NE, Xu W, Lucke AJ, Stoermer MJ, Fairlie DP. Orally absorbed cyclic peptides. *Chem Rev*. 2017;117:8094–128.
- Peier A, Ge L, Boyer N, Frost J, Duggal R, Biswas K, Edmondson S, Hermes JD, Yan L, Zimprich C, Sadruddin A, Kristal Kaan HY, Chandramohan A, Brown CJ, Thean D, Lee XE, Yuen TY, Ferrer-Gago FJ, Johannes CW, Lane DP, Sherborne B, Corona C, Robers MB, Sawyer TK, Partridge AW. NanoClick: a high throughput, target-agnostic peptide cell permeability assay. *ACS Chem Biol*. 2021;16:293–309.
- Peraro L, Deprey KL, Moser MK, Zou Z, Ball HL, Levine B, Kritzer JA. Cell penetration profiling using the chloroalkane penetration assay. *J Am Chem Soc*. 2018;140:11360–9.
- Pooga M, Langel U. Classes of cell-penetrating peptides. *Methods Mol Biol*. 2015;1324:3–28.
- Pye CR, Hewitt WM, Schwochert J, Haddad TD, Townsend CE, Etienne L, Lao Y, Limberakis C, Furukawa A, Mathiowetz AM, Price DA, Liras S, Lokey RS. Nonclassical size dependence of permeation defines bounds for passive adsorption of large drug molecules. *J Med Chem*. 2017;60:1665–72.
- Qian Z, LaRochelle JR, Jiang B, Lian W, Hard RL, Selner NG, Luechapanichkul R, Barrios AM, Pei D. Early endosomal escape of a cyclic cell-penetrating peptide allows effective cytosolic cargo delivery. *Biochemistry*. 2014;53:4034–46.
- Qian Z, Upadhyaya P, Pei D. Synthesis and screening of one-bead-one-compound cyclic peptide libraries. *Methods Mol Biol*. 2015;1248:39–53.
- Rader AFB, Weimuller M, Reichart F, Schumacher-Klinger A, Merzbach S, Gilon C, Hoffman A, Kessler H. Orally active peptides: is there a magic bullet? *Angew Chem Int Ed Engl*. 2018;57:14414–38.
- Rafi SB, Hearn BR, Vedantham P, Jacobson MP, Renslo AR. Predicting and improving the membrane permeability of peptidic small molecules. *J Med Chem*. 2012;55:3163–9.
- Rezgui R, Blumer K, Yeoh-Tan G, Trexler AJ, Magzoub M. Precise quantification of cellular uptake of cell-penetrating peptides using fluorescence-activated cell sorting and fluorescence correlation spectroscopy. *Biochim Biophys Acta*. 2016;1858:1499–506.

- Rhodes CA, Dougherty PG, Cooper JK, Qian Z, Lindert S, Wang QE, Pei D. Cell-permeable bicyclic peptidyl inhibitors against NEMO-IkappaB kinase interaction directly from a combinatorial library. *J Am Chem Soc.* 2018;140:12102–10.
- Roberts NA, Martin JA, Kinchington D, Broadhurst AV, Craig JC, Duncan IB, Galpin SA, Handa BK, Kay J, Krohn A, et al. Rational design of peptide-based HIV proteinase inhibitors. *Science.* 1990;248:358–61.
- Rosania GR, Thurber GM. Quantitative analysis of cellular drug transport, disposition, and delivery. New York: Springer; 2021.
- Sahni A, Qian Z, Pei D. Cell-penetrating peptides escape the endosome by inducing vesicle budding and collapse. *ACS Chem Biol.* 2020;15:2485–92.
- Sakamoto K, Kamada Y, Sameshima T, Yaguchi M, Niida A, Sasaki S, Miwa M, Ohkubo S, Sakamoto JI, Kamaura M, Cho N, Tani A. K-Ras(G12D)-selective inhibitory peptides generated by random peptide T7 phage display technology. *Biochem Biophys Res Commun.* 2017;484:605–11.
- Sakamoto K, Masutani T, Hirokawa T. Generation of KS-58 as the first K-Ras(G12D)-inhibitory peptide presenting anti-cancer activity in vivo. *Sci Rep.* 2020;10:21671.
- Sawyer TK. Renaissance in peptide drug discovery: the third wave. In: *Peptide-based drug discovery: challenges and new therapeutics.* Cambridge: Royal Society of Chemistry; 2017.
- Sawyer TK, Bohacek RS, Dalgarno DC, Eyermann CJ, Kawahata N, Metcalf CA 3rd, Shakespeare WC, Sundaramoorthi R, Wang Y, Yang MG. SRC homology-2 inhibitors: peptidomimetic and nonpeptide. *Mini Rev Med Chem.* 2002;2:475–88.
- Sawyer TK, Partridge AW, Kaan HYK, Juang YC, Lim S, Johannes C, Yuen TY, Verma C, Kannan S, Aronica P, Tan YS, Sherborne B, Ha S, Hochman J, Chen S, Surdi L, Peier A, Sauvagnat B, Dandliker PJ, Brown CJ, Ng S, Ferrer F, Lane DP. Macrocyclic alpha helical peptide therapeutic modality: a perspective of learnings and challenges. *Bioorg Med Chem.* 2018;26:2807–15.
- Schwochert J, Lao Y, Pye CR, Naylor MR, Desai PV, Gonzalez Valcarcel IC, Barrett JA, Sawada G, Blanco MJ, Lokey RS. Stereochemistry balances cell permeability and solubility in the naturally derived phepropeptin cyclic peptides. *ACS Med Chem Lett.* 2016;7:757–61.
- Seigal BA, Connors WH, Fraley A, Borzilleri RM, Carter PH, Emanuel SL, Fargnoli J, Kim K, Lei M, Naglich JG, Pokross ME, Posy SL, Shen H, Surti N, Talbott R, Zhang Y, Terrett NK. The discovery of macrocyclic XIAP antagonists from a DNA-programmed chemistry library, and their optimization to give lead compounds with in vivo antitumor activity. *J Med Chem.* 2015;58:2855–61.
- Semenza GL. Hypoxia-inducible factors: mediators of cancer progression and targets for cancer therapy. *Trends Pharmacol Sci.* 2012;33:207–14.
- Shakespeare W, Yang M, Bohacek R, Cerasoli F, Stebbins K, Sundaramoorthi R, Azimioara M, Chi V, Pradeepan S, Metcalf C. Structure-based design of an osteoclast-selective, nonpeptide src homology 2 inhibitor with in vivo antiresorptive activity. *Proc Natl Acad Sci.* 2000;97:9373–8.
- Shen K, Keng YF, Wu L, Guo XL, Lawrence DS, Zhang ZY. Acquisition of a specific and potent PTP1B inhibitor from a novel combinatorial library and screening procedure. *J Biol Chem.* 2001;276:47311–9.
- Skoulidis F, Li BT, Dy GK, Price TJ, Falchook GS, Wolf J, Italiano A, Schuler M, Borghaei H, Barlesi F, Kato T, Curioni-Fontecedro A, Sacher A, Spira A, Ramalingam SS, Takahashi T, Besse B, Anderson A, Ang A, Tran Q, Mather O, Henary H, Ngarmchamnanrith G, Friberg G, Velcheti V, Govindan R. Sotorasib for lung cancers with KRAS p.G12C mutation. *N Engl J Med.* 2021;384:2371–81.
- Stahelin HF. The history of cyclosporin A (Sandimmune) revisited: another point of view. *Experientia.* 1996;52:5–13.
- Vassilev LT. p53 activation by small molecules: application in oncology. *J Med Chem.* 2005;48:4491–9.
- Vives E. Cellular uptake [correction of utake] of the Tat peptide: an endocytosis mechanism following ionic interactions. *J Mol Recognit.* 2003;16:265–71.

- Wang CK, Northfield SE, Swedberg JE, Colless B, Chaousis S, Price DA, Liras S, Craik DJ. Exploring experimental and computational markers of cyclic peptides: charting islands of permeability. *Eur J Med Chem.* 2015;97:202–13.
- Zhang S, Zhang ZY. PTP1B as a drug target: recent developments in PTP1B inhibitor discovery. *Drug Discov Today.* 2007;12:373–81.
- Zhang Q, Lenardo MJ, Baltimore D. 30 years of NF-kappaB: a blossoming of relevance to human pathobiology. *Cell.* 2017;168:37–57.
- Zhu Z, Shaginian A, Grady LC, O’Keeffe T, Shi XE, Davie CP, Simpson GL, Messer JA, Evindar G, Bream RN, Thansandote PP, Prentice NR, Mason AM, Pal S. Design and application of a DNA-encoded macrocyclic peptide library. *ACS Chem Biol.* 2018;13:53–9.
- Zinker BA, Rondinone CM, Trevillyan JM, Gum RJ, Clampit JE, Waring JF, Xie N, Wilcox D, Jacobson P, Frost L, Kroeger PE, Reilly RM, Koterski S, Opgenorth TJ, Ulrich RG, Crosby S, Butler M, Murray SF, McKay RA, Bhanot S, Monia BP, Jirousek MR. PTP1B antisense oligonucleotide lowers PTP1B protein, normalizes blood glucose, and improves insulin sensitivity in diabetic mice. *Proc Natl Acad Sci U S A.* 2002;99:11357–62.
- Zizzari AT, Pliatsika D, Gall FM, Fischer T, Riedl R. New perspectives in oral peptide delivery. *Drug Discov Today.* 2021;26:1097–105.

Chapter 9

Regulatory Issues for Peptide Drugs



Seetharama D. Jois

Contents

9.1	Introduction.....	288
9.2	Definition of Biologics, Peptide, Polypeptide, and Protein Drugs.....	289
9.3	Investigational New Drug (IND) Application.....	290
9.4	New Drug Application (NDA).....	292
9.4.1	Chemistry.....	294
9.4.2	Raw Material Impurities.....	295
9.4.3	Process-Related Impurities.....	297
9.4.4	Stability.....	298
9.4.5	Nonclinical Pharmacology and Toxicology.....	299
9.4.6	Immunogenicity and Immunotoxicity.....	299
9.5	Abbreviated New Drug Application.....	300
9.6	Approval Process.....	302
9.7	Post-approval.....	302
	References.....	303

Abstract Peptides have characteristics of both proteins and small molecules, making them specific to bind to a target, similar to endogenous molecules. Peptide drugs fall in between small molecules and protein/antibody drugs. There is guidance developed for small molecules and proteins by regulatory agencies. However, for peptide drugs, guidance is limited, and FDA provides the guidelines for submission of data for chemistry and manufacturing of synthetic peptides. According to the FDA guidelines, “Peptide is a polymer composed of 40 or fewer amino acids.” In

S. D. Jois (✉)

School of Basic Pharmaceutical and Toxicological Sciences, College of Pharmacy, University of Louisiana at Monroe, Monroe, LA, USA

e-mail: jois@ulm.edu

2013, the US pharmacopeia convention (USP) formed a therapeutic peptide expert panel, and the panel recommended some guidelines for the quality of peptide synthesis and for considering peptides as an active pharmaceutical ingredient. With the anticipated demand for peptide therapeutics in metabolic diseases, autoimmune diseases, and cancer, there is an expanding market for peptide therapeutics in the next 10 years. Hence, a clear-cut regulatory requirement is essential for this class of therapeutics. This chapter covers some of the developments in the regulatory requirements for peptide-based therapeutics.

Keywords Peptides · Food and drug administration (FDA) · Investigational new drug (IND) · New drug application (NDA) · Peptide drug · Abbreviated new drug application (ANDA)

9.1 Introduction

The use of peptides as drugs was nearly 100 years old when the therapeutic use of insulin was approved in 1922 (Zaykov et al. 2016; Muttenthaler et al. 2021). Later, peptide hormones such as oxytocin and vasopressin were approved for clinical use, followed by somatostatin and gonadotropin-releasing hormone (GnRH)-related peptides and their analogs as agonists and antagonists. For the past 20 years, many peptide drugs have been approved for therapeutic use (Lau and Dunn 2018; Fosgerau and Hoffmann 2015; Muttenthaler et al. 2021). Peptide drugs fall into a different class of molecules that fall in between small molecules and protein/antibody drugs. They have characteristics of both proteins and small molecules, making them specific to bind to a target, similar to endogenous molecules and have potency in biological activity. The drug discovery, development, and approval by the Food and Drug Administration (FDA) process is a long road that involves several steps. The overall drug discovery and approval process are depicted in Fig. 9.1 (Moffat et al. 2017; Mohs and Greig 2017), and peptide drugs go through a similar pipeline for the approval process (Muttenthaler et al. 2021; Wu 2019).

Peptides can be extracted from natural resources and can be generated by molecular biology methods or synthetically produced (Itakura et al. 1992; Ladisch and Kohlmann 1992; Stahelin 1996; Coin et al. 2007). In terms of regulations to approve the peptide drugs for therapeutic purposes, there were no special FDA regulations. Larger peptides were treated as proteins, and smaller peptides were treated as small molecules (Uhlrig et al. 2014). With more than 50 peptide drugs approved within 20 years, regulations were adopted for peptide drug approval (Rastogi et al. 2019). With the anticipated demand for peptide therapeutics in metabolic diseases, autoimmune diseases, and cancer, there is an expanding market for peptide therapeutics in the next 10 years. Hence, a clear-cut regulatory requirement is essential for this class of therapeutics. In 2013, the US pharmacopeia convention (USP) formed a therapeutic peptide expert panel for guidance on peptide drugs, and in 2019, there was the American College of Toxicology (ACT) symposium on development and

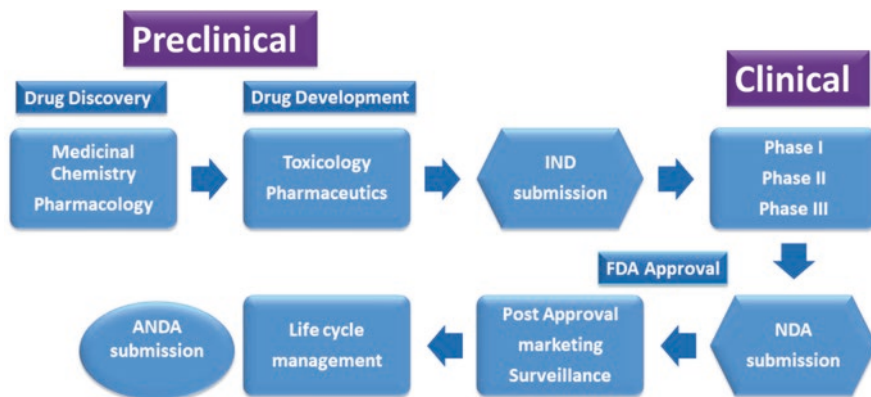


Fig. 9.1 Overview of drug discovery, development, and approval process

regulatory challenges for peptide therapeutics (Zane et al. 2021). FDA’s Center for Drug Evaluation and Research (CDER) released a draft guideline abbreviated new drug application (ANDAS) for certain highly purified synthetic peptide drug products that refer to listed drugs of recombinant DNA (rDNA) origin: Guidance for industry that specifically addresses peptides (Food and Administration 2017).

9.2 Definition of Biologics, Peptide, Polypeptide, and Protein Drugs

Peptide drugs are not classified under biologics, although many of the peptides approved by the FDA are modified versions of naturally occurring peptide ligands of human origin. Biologics are either isolated from natural sources and may be produced by biotechnology methods. The classical definition of peptides and proteins does not have a clear distinction between peptides, polypeptides, and proteins. FDA proposed a size-based cutoff for the definition of peptide, polypeptide, and protein drugs. The guidance, ANDAs for certain highly purified synthetic peptide drug products, provides a clear definition of peptide, polypeptide, and protein drugs. FDA defines some of these keywords in section 351(1) of the Public Health Service (PHS) Act. (Statements in the parenthesis are directly taken from FDA document).

A Peptide is “a polymer composed of 40 or fewer amino acids.” (<https://www.fda.gov/drugs/regulatory-science-action/impact-story-developing-tools-evaluate-complex-drug-products-peptides>)

Chemically Synthesized Polypeptide – “The term chemically synthesized polypeptide would mean any alpha amino acid polymer that (1) is made entirely by chemical synthesis and (2) is greater than 40 amino acids but less than 100 amino acids in size.”

Protein – “Any alpha amino acid polymer with a specific defined sequence that is greater than 40 amino acids in size”.

Biological Product – “a virus, therapeutic serum, toxin, antitoxin, vaccine, blood, blood component or derivative, allergenic product, protein (except any chemically synthesized polypeptide), or analogous product, or arsphenamine or derivative of arsphenamine (or any other trivalent organic arsenic compound), applicable to the prevention, treatment, or cure of a disease or condition of human beings”. (<https://www.federalregister.gov/documents/2020/02/21/2020-03505/definition-of-the-term-biological-product>)

With the definition of 40 amino acids or less regarded as peptides, we will restrict our descriptions to peptide drugs in this chapter. FDA also notes that ANDA is to demonstrate that the active ingredient in a proposed generic synthetic peptide drug product is the “same” as the active ingredient in a previously approved peptide of rDNA origin.

In terms of drug discovery and approval process, peptide drugs follow the same format as small molecules (Uhlig et al. 2014) (*Food and Drug Administration. Guidance for industry on biosimilars: Q & As regarding implementation of the BPCI. Act of 2009: questions and answers Part II; 2012* [<https://www.fda.gov/drugs/guidance-compliance-regulatory-information/guidances-drugs>]). A schematic diagram of the drug discovery and development process is provided in Fig. 9.1. The intent of this chapter is to provide some of the guidelines for peptide drug approval.

There is guidance developed for small molecules and protein by regulatory agencies, including International Council for Harmonization (ICH) S6 for proteins and ICHS1-S5, 7–8, ICH M3, and new chemical entities (NCEs) for small molecules.

However, for peptide drugs, guidance is limited as peptides fall between proteins and small molecules, and there are new guidelines developed. FDA provides the guidance for submission of data for chemistry and manufacturing of synthetic peptides (*FDA (2017) ANDAs for certain highly purified synthetic peptide drug products that refer to listed drugs of rDNA origin. FDA (1994) Guidance for industry for the submission of chemistry, manufacturing and controls information for synthetic peptide substances*).

The FDA has two independent specialized centers for the premarket review process, including (a) the Center for Biologics Evaluation and Research (CBER) and (b) the Center for Drug Evaluation and Research (CDER). Both biologics, small molecule drugs, and peptide drugs must first go through a rigorous process to determine their safety and efficacy in humans before they can be sold on the market. This involves data from basic research, preclinical data, and clinical trials in humans.

9.3 Investigational New Drug (IND) Application

There are two types of preclinical studies, *in vitro* and *in vivo*, that are required to prove that the drug does not have the potential to cause serious side effects in humans. Once the preclinical studies of a peptide drug are completed, before it is

tested in human clinical trials, the entity must apply for Investigational New Drug (IND) application. The term “Investigational New Drug (IND) refers to a drug developed by a pharmaceutical or biotech company or other organization that is ready for clinical trials on humans” (US FDA website). The intent of this application is to obtain the consent of the FDA to make sure that the drug is safe to administer in humans. According to FDA, the main objectives of IND are (1) to assure the safety and rights of subjects in all phases of an investigation and (2), in phases 2 and 3, to help assure that the quality of the scientific evaluation of the drug is adequate to permit an evaluation of the drug’s effectiveness and safety (21 CFR 312.22) (Guidance 2013). This application also provides the entity to transport the drugs to different sites within the country for the purpose of conducting clinical trials. IND applications generally contain the following information:

- (a) Preclinical pharmacology and toxicology studies – Data from in vivo studies in animal models to prove that the drug has efficacy and is safe to use in humans
- (b) Manufacturing details – Data from several batches of the manufactured drug to show that the proposed drug can be manufactured with proper control in place and the product quality attributes are reproducible between batches
- (c) Clinical protocols – Details of clinical protocol and qualification of investigators who oversee the protocol and the risk associated with exposing the drug to initial healthy volunteers

FDA recommends that before IND application is submitted, nonclinical studies related to the peptide drug should be conducted to make sure that the drug is safe in human clinical trials following the ICH guidelines (Table 9.1).

There are two types of INDs, commercial and research INDs. The commercial IND category is submitted by a commercial entity company interested in testing a drug to bring it to market. Research IND is usually from a non-profit organization or a research group from institutions. FDA reviews the IND applications and, based on the data, decides whether the new drug is safe enough to give it to humans in a clinical trial. At this stage, peptide synthesis or production should be scaled to the manufacturing level, and production should be sufficient for clinical trials. Quality controls should be established for active pharmaceutical ingredient (API) (Rasmussen 2018).

In the IND document for a peptide drug, the sponsors should provide detailed information about manufacturing. This information is evaluated to ensure that the chemistry and manufacturing data related to the peptide drug substance and peptide drug product does not pose any health risks to subjects enrolled in IND trials (Guidance 1995). Quality information for peptide drug products and drug substances should include a summary report containing physical, chemical, and biological characteristics relevant to the peptide. Some of the quality information needed are indicated in Table 9.2 (Wu et al. 2017a). In addition to this, genotoxicity and immunotoxicity data should be provided (Thybaud et al. 2016; Wang et al. 2020; Shelukhina et al. 2018). The sponsor should provide documentation to show that the peptide drug and the drug product both can be synthesized and produced with the required quality for clinical trials in large amounts, and batch-to-batch variation is minimal (Table 9.2).

Table 9.1 Guidance to documents for nonclinical study using ICH document

Study	Comments
Pharmacokinetic studies	Repeated dose tissue distribution Toxicokinetics, assessment of systemic exposure in toxicity studies
Chronic toxicity studies	Duration of chronic toxicity studies in rodent and non-rodent animals
Genotoxicity studies	Control of DNA reactive impurities in pharmaceuticals to limit potential carcinogenic risk
Carcinogenicity studies	Need for the study Dose selection Testing – to assess the risk of cancer induction by the chemical exposure to humans Type of tumors that develop in rats and mice when dosed for a long period (up to 2 years)
Safety pharmacology studies	QT interval prolongation
Reproductive toxicity studies	Toxicity of the pharmaceutical to reproduction
Immunotoxicity studies	Unintended immunosuppression or enhancement There are two groups in this class, according to ICH S8 guidance: 1. Drugs intended to modulate immune function for therapeutic purposes (e.g., to prevent organ transplant rejection) where adverse immunosuppression can be considered exaggerated pharmacodynamics 2. Drugs not intended to affect immune function but cause immunotoxicity due, for instance, to necrosis or apoptosis of immune cells or interaction with cellular receptors shared by both target tissues and nontarget immune system cells

Information obtained from ICH guidelines, <https://www.fda.gov/news-events/press-announcements/fda-brief-fda-finalizes-guidance-internationally-harmonized-recommendations-further-support-safe> (Niraghatam 2018)

9.4 New Drug Application (NDA)

FDA has two centers that review and approve drug marketing applications, namely the Center for Drug Evaluation and Research (CDER) and the Center for Biologics Evaluation and Research (CBER). Small molecule drugs are regulated by CDER; biological products can be regulated by either CDER or CBER (US Food and Drug Administration 2018a, b). FDA recommends that the sponsors follow the ICH guidelines for technical requirements for pharmaceuticals for human use. The overall process of drug approval is depicted in Fig. 9.2.

An NDA is an application to the marketing of a new drug in the United States. Typically, NDA consists of data for the new drug from nonclinical and clinical studies and a summary of formulation development and manufacturing process, including proposed labeling information. Drugs approved via NDA are regulated under section 505 of the Food, Drug, and Cosmetics Act (FD&C).

Table 9.2 IND requirements for peptide drugs

Labeling
Labeling should follow the format of structured product labeling (SPL) and physician's labeling rule (PLR) format
Summary
Overview of safety and efficacy, overall product quality for human clinical trials
Chemistry, manufacturing, and control information [21 CFR 312.23(a)(7)]
Drug substance [312.23 (a)(7)(iv)(a)]
A description of the drug substance, including its physical, chemical, or biological characteristics
The general method of preparation of the drug substance
The acceptable limits and analytical methods used to assure the identity, strength, quality, and purity of the drug substance
Information to support the stability of the drug substance during the toxicologic studies and the proposed clinical study(ies)
Drug product [21 CFR 312.23 (a)(7)(iv)(b)]:
A list of all components, which may include reasonable alternatives for inactive compounds, used in the manufacture of the investigational drug product, including both those components intended to appear in the drug product and those which may not appear, but which are used in the manufacturing process
Where applicable, the quantitative composition of the investigational new drug product, including any reasonable variations that may be expected during the investigational stage
A brief, general description of the method of manufacturing and packaging procedures as appropriate for the product
The acceptable limits and analytical methods used to assure the identity, strength, quality, and purity of the drug product
Information to support the stability of the drug substance during the toxicologic studies and the proposed clinical study(ies)
A brief general description of the composition, manufacture, and control of any placebo to be used in the proposed clinical trial(s) [21 CFR 312.23(a)(7)(iv)(c)]:
Pharmacology and toxicology information [21 CFR 312.23(a)(8)]
Pharmacology and drug distribution [21 CFR 312.23(a)(8)(i)
Toxicology: Integrated summary [21 CFR 312.23(a)(8)(ii)(a)]
Information obtained from Guidance for Industry, Content, and Format of Investigational New Drug Applications (INDs) for Phase 1 studies of drugs, including well-characterized, therapeutic, biotechnology-derived products (Guidance 1995)

According to FDA*, “The goals of the NDA are to provide enough information to permit FDA reviewer to reach the following key decisions. Whether the drug is safe and effective in its proposed use(s), and whether the benefits of the drug outweigh the risks.

- Whether the drug's proposed labeling (package insert) is appropriate and what it should contain.
- Whether the methods used in manufacturing the drug and the controls used to maintain the drug's quality are adequate to preserve the drug's identity, strength, quality, and purity.”

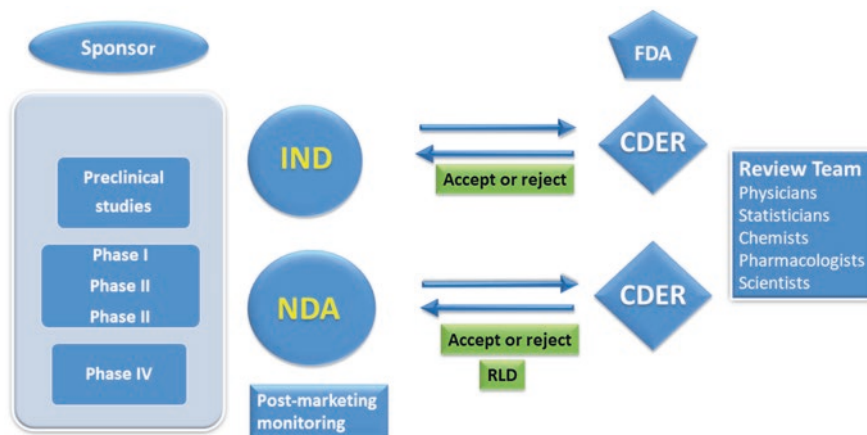


Fig. 9.2 Peptide drug review and approval process

*from the FDA website (<https://www.fda.gov/drugs/types-applications/new-drug-application-nda>).

The format and content for NDAs are specified in regulations (21 CFR 314.50). NDA contains the following information: (1) labeling, (2) summary, (3) chemistry, (4) nonclinical pharmacology and toxicology, (5) human pharmacokinetics and bio-availability, (6) clinical microbiology, (7) clinical data section, (8) safety update, and (9) statistics. Other reports include case reports and certifications and financial information.

We will describe some of the contents that need to be considered under chemistry, manufacturing, and controls under NDA.

9.4.1 Chemistry

Active and inactive ingredients in the drug product should be listed along with the dosage form and container closure system proposed for marketing. The development report should contain details of the development process for dosage form, formulation, and manufacturing process along with microbial attributes. The manufacturing details such as name and address of the manufacturer, flowcharts, details of equipment materials and in-process controls, testing of in-process for the finished peptide drug dosage form should be provided.

Most of the peptides are manufactured by solid-phase synthesis, and very few are by solution-phase synthesis (Kim and McAlpine 2013; Palomo 2014; Isidro-Llobet et al. 2019; Rasmussen 2018). Some cyclic peptides need a combination of solid-phase synthesis followed by a solution phase in the final step of the synthesis (Cheneval et al. 2014). Each step of the synthesis should be described in detail, including the amount of reactants, solvents, and reagents, reaction conditions, and coupling method; if coupling methods are repeated, there is need for such repeated

coupling, yield, and deprotection methods. If the reagents are obtained from other sources, for example, Fmoc protected amino acids, resins, the analytical data and quality of the products need to be evaluated. Typically, resin substitution level, swelling properties, particle size, and density should be provided. For amino acids, impurities present, chiral purity, and protecting groups, including side-chain protecting groups, should be provided. For coupling reactions, two independent testing methods should be provided (Kaiser, TNBS) (Kaiser et al. 1970; Hancock and Battersby 1976) as proof of a completed coupling reaction. For peptides with disulfide bonds, the Ellman test should be provided for the free sulfhydryl group present or not. Finally, based on the resin loading, efficiency yield of synthesis of desired peptide sequence should be reported. If the peptides have a large size, analytical data such as mass spectrometry data for molecular ions for intermediates should be reported (Wu 2019).

In 2013, the US pharmacopeia convention (USP) formed a therapeutic peptide expert panel, and the panel recommended some guidelines for the quality of peptide synthesis and for considering peptides as active pharmaceutical ingredients. Table 9.3 lists some of the quality controls that are recommended for peptide API required to obtain regulatory approval (Swietlow et al. 2014; Rastogi et al. 2019; Eggen et al. 2014; D'Addio et al. 2016; Zeng et al. 2019; D'Hondt et al. 2014a).

Furthermore, for the drug product manufacturing process, parameters such as dissolution, pH adjustment during compounding, order of addition of ingredients, holding time, and storage conditions before packaging must be considered. Detailed lyophilization conditions should be described.

During peptide manufacturing, lot-to-lot variation should be maintained within a range that ensures the therapeutic effects are realized. Heterogeneity in the peptide could be due to variability in the quality of the manufacturing process as well as the quality of the starting material (Rastogi et al. 2019).

9.4.2 Raw Material Impurities

Although the final API goes through quality control tests, it is important to start with some specifications for quality control tests for raw material used for peptide API manufacturing. ICH Q11, which has been effective since 2012, provides guidance on control of raw materials and impurities associated with raw materials and their fate during the manufacturing process. In the case of peptide drugs, raw materials include protected amino acids, derivatives of protected amino acids, and peptide fragments during long peptide synthesis. Along with this, reagents and solvents should be considered raw materials. The impurities in raw materials could be identified and qualified, identified, or unidentified (Eggen et al. 2014). During peptide manufacturing, the following raw materials should be considered: (1) Free amino acids with $N\alpha$ -protecting groups, (2) enantiomers, (3) other amino acid contaminants, (4) dipeptides or β -alanine containing Fmoc-amino acid derivatives (Eggen et al. 2014; Hlebowicz et al. 2005).

Table 9.3 Quality control parameters for peptide drugs

Characteristics	Comments
Appearance	Visual inspection
Description	US Adopted Name (USAN), chemical formula, amino acid sequence, pI value (isoelectric pH), solubility, molecular weight, and salt form (e.g., acetate or trifluoroacetate). Any modifications, such as esterification or amidation, should be described
Structure	The sequence of amino acids in three-letter code. For shorter peptides, (<15 AAs) 2D structure with chirality is preferred
<i>Identification</i>	
HPLC co-elution with a reference standard	HPLC methods should be established to separate reference standards and peptide drugs
Mass spectrometry	Monoisotopic mass \pm 1.0 mass units
Amino acid analysis	Hydrolysis protocol used
MS-MS sequencing	Complicated for longer sequences
NMR spectroscopy	Works for shorter peptides; for longer peptides, 1D NMR may be complicated to interpret
Peptide mapping	Used for longer peptide sequence (>20 AAs)
Enantiomeric purity	Chiral amino acid analysis
N-terminal sequencing	It may not be useful for cyclic or N-terminal modified peptides
Higher-order structure	Secondary structure can be determined for circular dichroism spectroscopy, which may not be useful for peptides that do not acquire a stable structure in solution
<i>Assay</i>	
HPLC assay	
Peptide content by amino acid analysis	Hydrolysis protocol
UV spectroscopy	May be useful for peptides containing amino acids Trp, Tyr, and Phe
Quantitative NMR	May be applicable for short peptides, stable internal standard is required
<i>Impurities</i>	
Peptide related substances	Process related and impurities of a degradation product. Limits for total and individual impurities must be specified
Residual solvents	May be limited to solvents used in the final step of the manufacturing process
Heavy metals	Contact with metal in manufacturing or metals used during synthesis
Residual trifluoroacetic acid (TFA)	Used for peptides that are not soluble in acid
Residual fluoride	Fmoc-chemistry does not use hydrofluoric acid, tested if Boc-chemistry is used
Other small molecule impurities	Nonpeptide impurities are required to follow ICHQ3A guidelines
<i>Specific tests</i>	
Counterion content	May be required for acetate, titration with AgNO ₃ may be used to determined chloride

(continued)

Table 9.3 (continued)

Characteristics	Comments
Water content	Coulometric titration method
Ellman test	Only for peptides containing a disulfide bond
Mass balance	Calculation
Bacterial endotoxin	Requirement for API
<i>Dosage form specific tests</i>	Injectables: test according to USP Nasal sprays: according to FDA guidelines

Information obtained from Swietlow et al. (2014), Rastogi et al. (2019), and Eggen et al. (2014)

Raw materials that are used should be of the highest grade available to avoid impurities that are generated during the manufacturing process. It is recommended to use US Pharmacopeia (USP), or national formulatory (NF) grade. In general, ICH Q7 does not require good manufacturing practice conditions (GMP) for raw materials. However, appropriate controls and impurity limits should be considered for the raw material used for manufacturing to obtain good quality API. Recommended guidelines for raw materials of peptide API are indicated in Table 9.4.

9.4.3 Process-Related Impurities

Peptides can undergo degradation and aggregation when they are subjected to stress related to manufacturing (D'Hondt et al. 2014a; Stalmans et al. 2016), processing, and shipping. These stress factors include temperature fluctuations, shaking, pH changes, and surfaces of the containers. Aggregation can happen as a result of denaturation of the peptides that may have secondary structure (Zapadka et al. 2017). Amino acids such as Asn can undergo deamination (Krogmeier et al. 2005; Li et al. 2003). Asp can undergo cyclization with the main chain (Grassi and Cabrele 2019). These changes may lead to the aggregation of peptides. Some hydrophobic amino acids aggregate depending on the conditions used during synthesis or process. Furthermore, aggregation leads to peptides being immunogenic. Thus, aggregation is an important factor to consider. Depending on the size of the peptide, variety of analytical methods can be used to monitor the aggregation, including size exclusion chromatography (SEC), dynamic light scattering (DLS), and analytical ultracentrifugation methods that provide information about the aggregation of peptides. Small aggregates may be very hard to detect in the case of peptides compared to proteins (Chen et al. 2021).

The final product of the formulation can also affect peptide aggregation. Excipients and preservatives such as polyethyl alcohol can change the conformation of peptides that have a particular secondary structure leading to aggregation. Polysorbate and peptide interaction can result in the generation of an immune response. Leachable container closure systems and syringes can affect peptides' immunogenicity (Lam et al. 1997; Bis et al. 2015; Costantino et al. 2009; Pang et al.

Table 9.4 Recommended guidelines for raw materials for therapeutic peptide synthesis

Quality attributes	Test method	Comments
Appearance	Visual inspection	
Identification	Mass spectrometry or HPLC	
Purity	HPLC	
Resin	On bead color test	Purchased from a qualified vendor Modification of the resin should be tested Stability of the linker tested under acidic or basic conditions. Test swelling properties under appropriate solvent used for the synthesis. Amino acid loading depends on swelling
Fmoc-amino acids	HPLC/TLC	Has to meet specific standards of quality control. Uncommon amino acids – description of synthesis, characterization by NMR, IR, MS, TLC, HPLC
<i>Related impurities</i>		
Free amino	TLC	
Other amino acids	Amino acid analysis	Refers to amino acids other than the one contained in the amino acid derivative, which is the subject of the specification
Specified	HPLC	Identity (if applicable) and limits set on the basis of batch history and the outcome of process characterization and a risk assessment
Unidentified	HPLC	Same as for specified
Enantiomer content	Chiral HPLC or Chiral GC-MS	Except for glycine derivatives
Assay	Titration	
Other components	As required	Includes potential reactive impurities, such as carboxylic acids determined during risk assessment
Reagents and solvents		In-house testing for vendor qualification to supply

Information obtained from Wu et al. (2017b), Food and Administration (2004), and Eggen et al. (2014). Obtained permission from International BioPharm, Control Strategies for Synthetic Therapeutic Peptide APIs – Part II: Raw Material Considerations By Ivo Eggen, Brian Gregg, Harold Rode, Aleksander Swietlow, Michael Verlander, and Anita Szajek

2007; Boven et al. 2005; Villalobos et al. 2005). Thus, leachable container closure systems should be studied for drug products. Details of container closure system that is used for marketing the drug product involve compatibility of construction materials with the peptide.

9.4.4 Stability

The stability of the peptide is more than the intact peptide and sequence. Peptides that have secondary structure and stability of the secondary structure should be included along with aggregation properties. Most of the peptides that are in powder

(lyophilized form) are stable under the low temperature and dry storage condition -20°C . However, in a solution state or under high humidity conditions, peptides can undergo degradation and/or aggregation. For example, methionine, tryptophan, and cysteine will be oxidized; asparagine and glutamine will be deamidated; the peptide backbone may be hydrolyzed (Grassi and Cabrele 2019; Furman et al. 2015; Zapadka et al. 2017). Stability testing protocols, accelerated stability testing conditions, storage conditions, and the proposed expiration date are essential (D'Hondt et al. 2014b; Blessy et al. 2014; Li et al. 2020). FDA requires stability studies for three batches of the product and long-term stability around up to 6 months and accelerated stability studies. Photostability and photosensitivity may also be needed.

9.4.5 Nonclinical Pharmacology and Toxicology

For pharmacology studies, the relevance of the model used for the study and clinical rationale should be provided along with routes of administration. Usually, primary pharmacology studies are done in rodents with a wide dose range and report the toxicities observed. Data from good laboratory practice (GLP) and acute toxicology in one animal model is provided. Potential carcinogenicity (Food and Administration 2002; Reynolds et al. 2020) and teratogenicity (Covington et al. 2004; Addis et al. 2000; Adam et al. 2011) are reported for FDA review (Reynolds et al. 2020; Niraghatam 2018). For details, one can look at ICH guidelines (ICH S3A, ICHS3 B, ICH S4, ICH S2 (R1), ICH M7, ICHS1A, S1B, S1C(R2), ICH S7A, S7B. ICH S5, S8).

9.4.6 Immunogenicity and Immunotoxicity

Therapeutic peptides generally do not produce immunogenicity. However, peptides used as vaccines generate an immune response (McGregor 2008). Most of the pre-clinical data is generated using animal models, and hence immune response and immunotoxicity data generated from animal models do not necessarily indicate immune response in humans. If a peptide drug indicates a high level of antibody production after a repeated dose of the drug during toxicity studies, then immunotoxicity should be evaluated. Depending on the size of the peptide, any linker used or conjugated to certain moieties, immunogenicity, immunotoxicity, as well as genotoxicity studies should be conducted (Rosenberg and Sauna 2018; Thybaud et al. 2016; Wang et al. 2020; Shankar et al. 2014).

Other reports must include human pharmacokinetics, bioavailability, and clinical microbiology. Along with this, a safety update is provided as an amendment to the NDA around 4 months after the original submission of the NDA.

9.5 Abbreviated New Drug Application

Abbreviated new drug application (ANDA) is an application to the FDA to manufacture and market a generic drug for an existing approved drug. Generally, the data is provided for the post-approval changes in accordance with section 506A of the Federal Food, Drug, and Cosmetic Act (the Act) and § 314.70 (21 CFR 314.70) (<https://www.fda.gov/drugs/types-applications/new-drug-application-nda>). The post-approval changes include (1) components and composition, (2) manufacturing sites, (3) manufacturing process, (4) specifications, (5) container closure system, and (6) labeling, as well as (7) miscellaneous changes and (8) multiple related changes. These changes can be categorized as major, moderate, and minor changes.

The major change is considered as changes in identity, strength, quality, purity, or potency of a drug product as these factors may relate to the safety or effectiveness of the drug product. Moderate and minimal changes are that it has a moderate to minimal change to the above characteristics.

In October 2017, FDA provided guidance for certain highly purified synthetic peptide drug products that refer to listed drugs of rDNA origin. This document was updated in May 2021 (Food and Administration 2017). For such products, the application should be submitted as an abbreviated new drug application (ANDA) under section 505(j) of the Federal Food, Drug, and Cosmetic Act (FD&C Act). The guidance is for application of a synthetic peptide drug product (synthetic peptide) that refers to a previously approved peptide drug product of recombinant deoxyribonucleic acid (rDNA) origin (peptide of rDNA origin). A peptide drug that was previously approved had a production procedure by using biotechnology methods is proposed to be produced by solid-state or solution-phase synthesis method then ANDA should be submitted.

ANDA should demonstrate that the active ingredient in a proposed synthetic drug product is the same as an active ingredient in a previously approved peptide of rDNA origin. The abbreviated new drug application includes the following:

Chemistry, Manufacturing, and Controls Section This section should describe composition, manufacture, and specification of the drug substance and the drug product.

Drug Substance This section should describe physical and chemical characteristics and stability, the name and address of its manufacturer, the method of synthesis (or isolation), and purification of the drug substance. Some of the required parameters are indicated in Table 9.5. Other descriptions include (a) the process controls used during manufacture and packaging; (b) specifications necessary to ensure the identity, strength, quality, and purity of the drug substance; (c) bioavailability of the drug products made from the substance; and (d) analytical procedures and acceptance criteria relating to stability, sterility, particle size, and crystalline form. The NDA may also provide additional information for the use of alternatives to meet any of these requirements (Food and Administration 2017).

Table 9.5 Data related to synthetic generic peptides for ANDA

Characteristics	Analytics or other methods
<i>1. Active ingredient sameness</i>	
Primary sequence and physicochemical properties	N- and C-terminal sequence, sequence by MS-MS
Secondary structure	Circular dichroism method
Oligomer/aggregation states	Depending on the size of the peptide, the following analytical methods can be used: size exclusion chromatography (SEC), dynamic light scattering (DLS), and analytical ultracentrifugation. Small aggregates of peptides are difficult to detect.
Biological activities	In vitro and animal model
<i>2. Impurities</i>	
RLD and generic synthetic have the same impurities	HPLC, mass spectrometry
Peptide-related impurities	Oxidation, glycosylation
Host cell-related impurities DNA, other proteins	Only for rDNA origin
Non-peptide-related impurities	Impurities that are 0.1% of the drug substance or greater should be identified (UHPLC-HRMS) Should not contain any new peptide-related impurity more than 0.5% of the drug substance. A new specified peptide-related impurity level higher than 0.5% of the drug substance raises concerns about the potential risk of immunogenicity. For each new specified peptide-related impurity that is not more than 0.5% of the drug substance, the ANDA should characterize the impurity. Further, the ANDA should provide justification for why such impurity does not affect the safety of the proposed generic synthetic peptide (including with respect to immunogenicity) and why it does not affect its effectiveness

Information obtained from ANDAs for certain highly purified synthetic peptide drug products that refer to listed drugs of rDNA origin guidance for industry (Food and Administration 2017)

In producing the above documents, ANDA requires a full description of the composition, manufacture, and specifications of the drug substance and product. Analytical data related to chemical characterization, stability, and purity needs to be submitted (Table 9.5). In addition, the bioavailability of the synthetic peptide should be presented. To assess whether the proposed generic synthetic peptide meets the approval standards, FDA can recommend additional studies related to in vitro data or animal pharmacokinetics/pharmacodynamic equivalence.

Drug Product ANDA requires whether the proposed generic uses the same inactive ingredient as the reference listed drug (RLD) and concentrations of the inactive ingredient used are within $\pm 5\%$ of that used in RLD. If new excipients are used, the

effect of the new excipient on the stability of peptide, peptide excipient interaction, immunoassays, and bioassays should be carried out. Furthermore, for the synthetic drug substance, how the other ingredients are added (order of addition), holding time in the storage, freeze-drying conditions, and process control need to be described.

9.6 Approval Process

Once the NDA is submitted to FDA with nonclinical and clinical data, CDER group members evaluate the documents, and if needed, a site visit is proposed for the manufacturing and product formulation and packing facility. Depending on the quality of the data and the efficacy, FDA can accept or reject the application. FDA strongly recommends that sponsor/applicant requests a meeting with them to make sure that the application is complete with all the required data for FDA to review and application is fillable. Once the application is filed within 45 days, the FDA committee will decide whether the application is ready for filing. If the application does not have the necessary information for review, then refuse to file letter will be sent. Hence, it is recommended to have a meeting with FDA before filing. If the application is acceptable, then FDA informs the sponsors of the review process. The review team will have a diverse background (Fig. 9.2) and, after the review, will write an assessment document. For new molecular entities, the review usually takes 10 months; the accelerated review may take place faster than that.

After the drug approval, the peptide drug can be marketed based on preapproval condition and post-marketing requirements (Fig. 9.2). In case the committee does not issue the approval, a complete response letter (CR) will be issued. The sponsor can meet the FDA for details of deficiencies and decide to resubmit or withdraw the application. In most cases, a mid-term review or mid-cycle review information request (IR) is communicated to the sponsor so that any response or remediation can be provided before the final decision. FDA can also seek advice from an advisory committee that is external to FDA with particular expertise.

9.7 Post-approval

Any changes in the approved peptide drugs or products manufacturing and components used for manufacturing, packaging, specifications, stability protocols, or expiration dates must be submitted as supplements. These changes are graded as major, moderate, or minor and submitted. Any change in the equipment also must be submitted as an annual report according to FDA guidelines (Wu et al. 2017a; FDA 2004).

Acknowledgments This work was supported by funding from the National Cancer Institute of the National Institutes of Health grant 1RO1CA255176 to SJ.

References

- Adam MP, Polifka JE, Friedman JM. Evolving knowledge of the teratogenicity of medications in human pregnancy. *Am J Med Genet C Semin Med Genet.* 2011;157(3):175–82. Wiley Online Library.
- Addis A, Sharabi S, Bonati M. Risk classification systems for drug use during pregnancy. *Drug Saf.* 2000;23:245–53.
- Bis RL, Singh SM, Cabello-Villegas J, Mallela KM. Role of benzyl alcohol in the unfolding and aggregation of interferon alpha-2a. *J Pharm Sci.* 2015;104:407–15.
- Blessy M, Patel RD, Prajapati PN, Agrawal YK. Development of forced degradation and stability indicating studies of drugs – a review. *J Pharm Anal.* 2014;4:159–65.
- Boven K, Stryker S, Knight J, Thomas A, van Regenmortel M, Kemeny DM, Power D, Rossert J, Casadevall N. The increased incidence of pure red cell aplasia with an Eprex formulation in uncoated rubber stopper syringes. *Kidney Int.* 2005;67:2346–53.
- Chen T, Tang S, Hecht ES, Yen CW, Andersen N, Chin S, Cadang L, Roper B, Estevez A, Rohou A, Chang D, Dai L, Liu P, Al-Sayah M, Nagapudi K, Lin F, Famili A, Hu C, Kuhn R, Stella C, Crittenden CM, Gruenhagen JA, Venkatramani C, Hannoush RN, Leung D, Vandlen R, Yehl P. Discovery of a dual pathway aggregation mechanism for a therapeutic constrained peptide. *J Pharm Sci.* 2021;110:2362–71.
- Cheneval O, Schroeder CI, Durek T, Walsh P, Huang YH, Liras S, Price DA, Craik DJ. Fmoc-based synthesis of disulfide-rich cyclic peptides. *J Org Chem.* 2014;79:5538–44.
- Coin I, Beyermann M, Bienert M. Solid-phase peptide synthesis: from standard procedures to the synthesis of difficult sequences. *Nat Protoc.* 2007;2:3247–56.
- Costantino HR, Culley H, Chen L, Morris D, Houston M, Roth S, Phoenix MJ, Foerder C, Philo JS, Arakawa T, Eidenschink L, Andersen NH, Brandt G, Quay SC. Development of Calcitonin Salmon Nasal Spray: similarity of peptide formulated in chlorobutanol compared to benzalkonium chloride as preservative. *J Pharm Sci.* 2009;98:3691–706.
- Covington DL, Tilson H, Elder J, Doi P. Assessing teratogenicity of antiretroviral drugs: monitoring and analysis plan of the Antiretroviral Pregnancy Registry. *Pharmacoepidemiol Drug Saf.* 2004;13:537–45.
- D’Addio SM, Bothe JR, Neri C, Walsh PL, Zhang J, Pierson E, Mao Y, Gindy M, Leone A, Templeton AC. New and evolving techniques for the characterization of peptide therapeutics. *J Pharm Sci.* 2016;105:2989–3006.
- D’Hondt M, Bracke N, Taevernier L, Gevaert B, Verbeke F, Wynendaele E, De Spiegeleer B. Related impurities in peptide medicines. *J Pharm Biomed Anal.* 2014a;101:2–30.
- D’Hondt M, Fedorova M, Peng CY, Gevaert B, Taevernier L, Hoffmann R, De Spiegeleer B. Dry heat forced degradation of busserelin peptide: kinetics and degradant profiling. *Int J Pharm.* 2014b;467:48–9.
- EGGEN I, GREGG B, RODE H, SWIETLOW A, VERLANDER M, SZAJEK A. Control strategies for synthetic therapeutic peptide APIs part II: raw material considerations. *Bio Pharm Int.* 2014;27:24.
- FDA, US. “Guidance for Industry for the submission of chemistry, manufacturing, and controls information for a therapeutic recombinant DNA-derived product or monoclonal antibody product for in vivo use.” US FDA, MD, USA (1996).
- Food, and Drug Administration. Guidance for industry: carcinogenicity study protocol submissions. Rockville: FDA; 2002.
- Food, and Drug Administration. Guidance for industry for the submission of chemistry. In: Manufacturing, and controls information for synthetic peptide substances; 2004 published by FDA.
- Fosgerau K, Hoffmann T. Peptide therapeutics: current status and future directions. *Drug Discov Today.* 2015;20:122–8.
- Furman JL, Chiu M, Hunter MJ. Early engineering approaches to improve peptide developability and manufacturability. *AAPS J.* 2015;17:111–20.

- Grassi L, Cabrele C. Susceptibility of protein therapeutics to spontaneous chemical modifications by oxidation, cyclization, and elimination reactions. *Amino Acids*. 2019;51:1409–31.
- Guidance, FDA. Content and format of investigational new drug applications (INDs) for phase 1 studies of drugs, including well-characterised, therapeutic, biotechnology-derived products. FDA. 1995.
- Hancock WS, Battersby JE. A new micro-test for the detection of incomplete coupling reactions in solid-phase peptide synthesis using 2, 4, 6-trinitrobenzene-sulphonic acid. *Anal Biochem*. 1976;71:260–4.
- Hlebowicz E, Andersen AJ, Andersson L, Moss BA. Identification of Fmoc-beta-Ala-OH and Fmoc-beta-Ala-amino acid-OH as new impurities in Fmoc-protected amino acid derivatives. *J Pept Res*. 2005;65:90–7.
- Isidro-Llobet A, Kenworthy MN, Mukherjee S, Kopach ME, Wegner K, Gallou F, Smith AG, Roschangar F. Sustainability challenges in peptide synthesis and purification: from R&D to production. *J Org Chem*. 2019;84:4615–28.
- Itakura K, Tadaaki H, Crea R, Riggs AD, Heyneker HL, Bolivar F, Boyer HW. Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. 1977. *Biotechnology*. 1992;24:84–91.
- Kaiser E, Colescott RL, Bossinger CD, Cook PI. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal Biochem*. 1970;34:595–8.
- Kim SJ, McAlpine SR. Solid phase versus solution phase synthesis of heterocyclic macrocycles. *Molecules*. 2013;18:1111–21.
- Krogmeier SL, Reddy DS, Vander Velde D, Lushington GH, Siahaan TJ, Middaugh CR, Borchardt RT, Topp EM. Deamidation of model beta-turn cyclic peptides in the solid state. *J Pharm Sci*. 2005;94:2616–31.
- Ladisch MR, Kohlmann KL. Recombinant human insulin. *Biotechnol Prog*. 1992;8:469–78.
- Lam XM, Patapoff TW, Nguyen TH. The effect of benzyl alcohol on recombinant human interferon-gamma. *Pharm Res*. 1997;14:725–9.
- Lau JL, Dunn MK. Therapeutic peptides: historical perspectives, current development trends, and future directions. *Bioorg Med Chem*. 2018;26:2700–7.
- Li B, Borchardt RT, Topp EM, VanderVelde D, Schowen RL. Racemization of an asparagine residue during peptide deamidation. *J Am Chem Soc*. 2003;125:11486–7.
- Li Y, Hu Y, Logsdon DL, Liu Y, Zhao Y, Cooks RG. Accelerated forced degradation of therapeutic peptides in levitated microdroplets. *Pharm Res*. 2020;37:138.
- McGregor DP. Discovering and improving novel peptide therapeutics. *Curr Opin Pharmacol*. 2008;8:616–9.
- Moffat JG, Vincent F, Lee JA, Eder J, Prunotto M. Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nat Rev Drug Discov*. 2017;16:531–43.
- Mohs RC, Greig NH. Drug discovery and development: role of basic biological research. *Alzheimers Dement (N Y)*. 2017;3:651–7.
- Muttenthaler M, King GF, Adams DJ, Alewood PF. Trends in peptide drug discovery. *Nat Rev Drug Discov*. 2021;20:309–25.
- Niraghatam VV. Regulatory issues concerning the preclinical testing of synthetic peptides. Eastern Michigan University; 2018. Ypsilanti, Michigan
- Palomo JM. Solid-phase peptide synthesis: an overview focused on the preparation of biologically relevant peptides. *RSC Adv*. 2014;4:32658–72.
- Pang J, Blanc T, Brown J, Labrenz S, Villalobos A, Depaolis A, Gunturi S, Grossman S, Lisi P, Heavner GA. Recognition and identification of UV-absorbing leachables in EPREX pre-filled syringes: an unexpected occurrence at a formulation-component interface. *PDA J Pharm Sci Technol*. 2007;61:423–32.
- Rasmussen JH. Synthetic peptide API manufacturing: a mini review of current perspectives for peptide manufacturing. *Bioorg Med Chem*. 2018;26:2914–8.
- Rastogi S, Shukla S, Kalavani M, Singh GN. Peptide-based therapeutics: quality specifications, regulatory considerations, and prospects. *Drug Discov Today*. 2019;24:148–62.

- Reynolds VL, Butler P, Abernathy MM, Aschenbrenner L, Best DD, Blank J, Crosby M, Custer L, Escobar PA, Kolaja K, Moggs J, Shuey D, Snyder C, Van Vleet T, Zhou J, Hart TK. Nonclinical safety assessment of epigenetic modulatory drugs: current status and industry perspective. *Regul Toxicol Pharmacol*. 2020;117:104746.
- Rosenberg AS, Sauna ZE. Immunogenicity assessment during the development of protein therapeutics. *J Pharm Pharmacol*. 2018;70:584–94.
- Shankar G, Arkin S, Cocea L, Devanarayan V, Kirshner S, Kromminga A, Quarmby V, Richards S, Schneider CK, Subramanyam M, Swanson S, Verthelyi D, Yim S, Scientists American Association of Pharmaceutical. Assessment and reporting of the clinical immunogenicity of therapeutic proteins and peptides-harmonized terminology and tactical recommendations. *AAPS J*. 2014;16:658–73.
- Shelukhina IV, Zhmak MN, Lobanov AV, Ivanov IA, Garifulina AI, Kravchenko IN, Rasskazova EA, Salmova MA, Tukhovskaya EA, Rykov VA, Slashcheva GA, Egorova NS, Muzyka IS, Tsetlin VI, Utkin YN. Azemiopsin, a selective peptide antagonist of muscle nicotinic acetylcholine receptor: preclinical evaluation as a local muscle relaxant. *Toxins (Basel)*. 2018;10:10(1):34
- Stahelin HF. The history of cyclosporin A (Sandimmune) revisited: another point of view. *Experientia*. 1996;52:5–13.
- Stalmans S, Gevaert B, Verbeke F, D'Hondt M, Bracke N, Wynendaele E, De Spiegeleer B. Quality control of cationic cell-penetrating peptides. *J Pharm Biomed Anal*. 2016;117:289–97.
- Swietlow A, Rode H, Szajek A, Verlander M, Eggen I, Gregg B. Control strategies for synthetic therapeutic peptide APIs-Part I: analytical consideration. *Pharm Technol*. 2014;38:48–55.
- Thybaud V, Kasper P, Sobol Z, Elhajouji A, Fellows M, Guerard M, Lynch AM, Sutter A, Tanir JY. Genotoxicity assessment of peptide/protein-related biotherapeutics: points to consider before testing. *Mutagenesis*. 2016;31:375–84.
- Uhlig T, Kyprianou T, Martinelli FG, Oppici CA, Heiligers D, Hills D, Calvo XR, Verhaert P. The emergence of peptides in the pharmaceutical business: from exploration to exploitation. *EuPA Open Proteom*. 2014;4:58–69.
- US FDA. FDA guidance for industry. Changes to an approved NDA or ANDA. 2004.
- US Food and Drug Administration. ANDAs for certain highly purified synthetic peptide drug products that refer to listed drugs of rDNA origin, guidance for industry. Silver Spring: US Food and Drug Administration; 2017.
- Villalobos AP, Gunturi SR, Heavner GA. Interaction of polysorbate 80 with erythropoietin: a case study in protein-surfactant interactions. *Pharm Res*. 2005;22:1186–94.
- Wang X, Meng N, Wang S, Lu L, Wang H, Zhan C, Burgess DJ, Lu W. Factors influencing the immunogenicity and Immunotoxicity of cyclic RGD peptide-modified Nanodrug delivery systems. *Mol Pharm*. 2020;17:3281–90.
- Wu L. Regulatory considerations for peptide therapeutics. In: *Peptide therapeutics*. London: Royal Society of Chemistry; 2019.
- Wu L, Smith H, Zheng H, Yu LX. Drug product approval in the United States and international harmonization. In: *Developing solid oral dosage forms*. London: Elsevier; 2017a.
- Wu LC, Chen F, Lee SL, Raw A, Yu LX. Building parity between brand and generic peptide products: regulatory and scientific considerations for quality of synthetic peptides. *Int J Pharm*. 2017b;518:320–34.
- Zane D, Feldman PL, Sawyer T, Sobol Z, Hawes J. Development and regulatory challenges for peptide therapeutics. *Int J Toxicol*. 2021;40:108–24.
- Zapadka KL, Becher FJ, Gomes Dos Santos AL, Jackson SE. Factors affecting the physical stability (aggregation) of peptide therapeutics. *Interface Focus*. 2017;7:20170030.
- Zaykov AN, Mayer JP, DiMarchi RD. Pursuit of a perfect insulin. *Nat Rev Drug Discov*. 2016;15:425–39.
- Zeng K, Boyne MT, Toby TK, Ruzicka C. Impurity characterization and quantification by liquid chromatography–high-resolution mass spectrometry. In: *Peptide therapeutics*. London: Royal Society of Chemistry; 2019.

Index

A

Abbreviated new drug application (ANDA), 300–302
Absorption enhancers, 192
Accelerated blood clearance (ABC), 224
ACE enzyme, 6
ACE receptor, 6
Acetylation, 68
Active targeting liposomes, 221
AddCompositionConstraintMover, 113
Adsorptive-mediated transcytosis (AMT), 248, 249
ADTC5 peptide, 258
AeroDose, 191
AERx, 191
Agonists/antagonists, 6–8
Air oxidation, 63
Alanine scanning, 8, 11, 35
Alefacept, 34
Alleviate pain, 16
Alpha helix, 12
ALRN-6924, 186
Alzheimer's disease (AD), 238, 259, 261, 262
Amide proton atom, 121
Amino acid, 3, 13, 14, 52, 53, 72
 proline, 20
 racemization, 60
 side chain, 14
Aminopeptidases, 10
Ammonium bicarbonate, 64
Angiogenesis, 23
Angiopep-2 peptide (ANG), 248
Angiotensin I, 18, 19
Angiotensin II, 18, 19
Angiotensin-converting enzyme (ACE), 18–21
Angiotensinogen, 18, 19

Angiotensin receptor, 6
ANG-TAT-PTX conjugate, 248
Antiangiogenic molecules, 23
Antibodies, 33
Antibody-drug conjugates (ADCs), 246
Antibody mapping, 24
Antigen-presenting cells (APC), 33
Anti-infective agents, 16
Antimicrobial peptides (AMPs), 16, 29, 30
Apoferitin (APO), 250
Approval process, 302
Arg-Gly-Asp (RGD motif), 23
Arg-Gly-Asp-Ser, 24
Arg-Gly-Asp-X-pro-Cys, 25
Arg48-Lys49-Glu50-Lys51, 35
Asp87-Thr88-Lys89-Gly90, 35
Assisted Model Building with Energy Refinement (AMBER) family, 91
Atoms, 93
Autoimmune diseases, 14, 15, 33
Automation of SPPS, 64

B

Backbone cyclization, 70
Backbone dihedral angles, 3, 4
Backbone head to tail cyclization, 11
Backbone/side-chain modification, 12, 13
Bacterial cell membrane, 29
Baculoviral IAP repeat (BIR), 274
Batch synthesizers, 64
Benzyl succinic acid, 20
Beta-amino acids, 13
Beta-strand structure, 70
 β -turn mimetics, 14
Beta-turns/beta-hairpin bends, 11

- Binding affinity, 9
 Binding assays, 17, 18
 Binding cavity, 15
 Binding epitope, 24–26
 Binding site and 3D structure, 6, 7
 Biochemical pathways, 5, 6, 33, 34
 Biologically active conformation, 9
 Biological product, 290
 Biologics, 166, 289–290
 BioOral system, 189
 Biotin-labeled lysine, 61
 Blood, 176
 Blood-brain barrier (BBB)
 BBBD/osmotic brain delivery method, 253
 BBBMs of intercellular junction proteins, 254–258
 diffusion pathway of peptide, 238
 in vitro and in vivo models for brain delivery, 240
 nanoparticles, 250
 nasal delivery of peptides, 259–262
 passive diffusion, transcellular pathway, 242–245
 peptide conjugates for drug delivery, 248, 249
 permeability assessment, 241, 242
 receptor-mediated transcytosis of peptides, 245–248
 structure of, 240, 241
 Blood-brain barrier disruption (BBBD) method, 253
 Blood-brain barrier modulators (BBBMs), 254, 258
 Bortezomib, 273
 Bovine brain microvessel endothelial cell (BBMEC), 255
 Brain capillary endothelial cells (BCEC), 250
 Breast cancer resistance protein (BCRP), 240
 BuriedUnsatHbonds, 118
 BYDUREON BCise, 187
- C**
- CAAX motif, 271
 Caco-2 cells, 36
 Cadherin peptides, 255
 Camptothecin-glutamic acid (CPT-Glu) conjugate, 257
 Cancer, 14, 15
 CAP-PEG-insulin-casein (CAPIC), 189
 Captopril, 21
 Carbobenzoxy, 53
 Carbon nanotubes (CNTs), 207
 Carbonyl oxygen atom, 121
 Carboxylic acid, 53
 Carboxypeptidases, 10, 11
 Cardiovascular diseases, 14
 Carrier systems, 186, 189, 196, 197
 Cationic partitioning, 280
 C-backbone cyclization, 11
 CD2-CD58 complex, 34, 35
 C2 decapeptide, 251
 Cell adhesion, 24
 Cell adhesion inhibition, 24
 Cell attachment assay, 24
 Cell-cell adhesion, 35
 Cell-penetrating peptides (CPPs), 68, 248, 271
 Cell permeability, 277, 280
 CellTiter-Glo assay, 17
 Central processing unit (CPU), 84
 Cetorelix, 192
 CFTR-associated ligand (CAL), 276
 Chemically synthesized polypeptide, 289
Chemistry at Harvard Molecular Mechanics (CHARMM) family, 91
 ChemMatrix® resin, 65
 Chitosan-based micelle, 207
 Cholesterol control, 5
 Cholesteryl hemisuccinate (CHEMS), 222
 Chromatogram, 67, 69, 73
 Cilengitide, 26–28
 Circular dichroism, 9, 11
 Cold diethyl ether, 73
 Colistin, 31, 32
 Collagen-induced arthritis (CIA), 38
 Collagen receptors, 23
 Conformational constraints, 9
 Continuous flow synthesizers, 64
 Conventional micelles, 207
 Coupling, 53
 and additives, 60, 62
 chemistry, 60
 of FAM, 68
 and reagents, 60, 62, 65
 Covid-19 mRNA vaccines, 227
 Critical Assessment of Protein Structure Prediction (CASP), 132
 C2-9r-siRNA complex, 252
 Crude cyclic peptide, 73
 Cryo-electron microscopy (cryo-EM), 99
 Cryopreserved hepatocytes, 177
 Cryoprotectants, 215
 Cryo-TEM, 212
 Crystal structure, CD2-CD58 complex, 34, 35
 C-terminal amidation, 172
 C-terminal amino acid, 53

C-terminal functional groups, 60
C-terminal transactivation domain (CTAD), 275
Cyclic peptides, 11, 35, 39, 63, 172, 173
 cyclo (SIAD-dp-PDDIK), 70–72
 N-methylation, 27
Cyclic RGD peptides, 25
Cyclization, 10–12, 63, 172, 243
Cyclosporine A (CsA), 268–270, 277, 278
Cyclotides, 16
Cys amino acids, 11
Cystic fibrosis transmembrane conductance regulator (CFTR) gene, 276
Cytokines, 33
Cytoplasmic membranes, 29
Cytoplasmic tail, 33
Cytosolic inhibitor of apoptosis (cIAP), 274

D

Dab side chain, 31
D-amino acids, 9, 10, 16, 26, 95, 168, 238, 239, 251
Dead-end elimination (DEE), 108
Dead-end elimination with perturbations (DEEPer) approach, 108
Deep learning, 132
Deep neural networks (DNN), 131–133
Degarelix, 175
Density functional theory (DFT), 93
Desacyl ghrelin (DAGs), 247
Design-centric guidance functions, 111
Desmopressin, 188
Diabetes mellitus type 2, 16
 γ -Diaminobutyric acid (Dab) residues, 30
Dicarboxylate-containing inhibitors, 21, 22
Dichloromethane (DCM), 65
Diisopropylcarbodiimide, 65
Dimethylformamide (DMF), 65
Dipalmitoyl phosphatidylcholine (DPPC) TEL, 217
Dipeptide glycine-glycine, 52
Dipeptidyl peptidase 4 (DPP-4), 171
1- α -Distearoylphosphatidylcholine (DSPC), 216
Disulfide bond, 11, 63, 64, 72
Dose-response curve, 17
Double liposomes (DL), 216
Doxorubicin (Dox), 249
Dox-pVEC-gHo conjugate, 249
D-Phe-Cys-Tyr-D-Trp-Arg-Thr-Pen-Thr-NH₂ (CTAP), [D-penicillamin^{2,5}] enkephalin (DPDPE), 246

Drug discovery, 15
Drug-like molecules, 24
Drug product, 301
Drug screening, 15
Drug substance, 300
Dry powder inhalers (DPI), 190, 191
Dynamic light scattering (DLS), 213

E

ECM proteins, 24
Egg phosphatidylcholine (EPC), 217
Elbasvir, 273
Electroporation, 195
Electrospray, 65
Elution, 66
Enfuvirtide, 10
Entrada therapeutics, 278
Enzyme, 15
Enzyme activity inhibition assay, 18
Enzyme-based drug design, 18–21
Enzyme inhibition, 175
Enzyme-linked immunosorbent assay (ELISA), 18, 24
Eptifibatid, 17
E-rosetting, 36
Ethylenediaminetetraacetic acid (EDTA), 176, 189
Ethylene glycol tetraacetic acid (EGTA), 189
Exo- and endo-peptidases, 238
Exopeptidases, 10
Extracellular matrix (ECM), 22, 23

F

FAM conjugation, 68
Farnesyltransferase (FTase), 271
FastDesign, 106, 108
FastRelax, 106
Fatty acyl amino acid, 31
FDA guidelines, 3
Fibronectin, 24
First-generation peptides, 34–37
Flanking residue, 35
Flat-bottom microplate reader, 24
Flip-flop kinetics, 175
Fluorescent-labeled peptides, 18
Fluorescently labeled peptide
 Ac-RWVOWIO(FAM)QVR-dP-G-NH₂, 68–69
Fmoc-amino acids, 70
Fmoc chemistry, 13
Fmoc-protected amino acid, 60

Fmoc protecting group, 57
 Fmoc solid-phase peptide synthesis, 58–59
 Formulation development, 184
 Fragment insertion, 94
 Fragment molecular orbital (FMO)
 approach, 93
 Freeze-drying, 215

G

GABA_A receptors, 98
 Gadopentetic acid (Gd-DTPA), 257, 258
 Ganirelix, 10
 Gastric acid, 187
 Gastric intestinal stability assays, 178
 General AMBER Force Field (GAFF), 91
 General Atomic and Molecular Electronic
 Structure System (GAMESS), 94
 Generalized kinematic closure
 (GeneralizedKIC), 102, 103
 Geranylgeranyltransferase (GGTase), 271
 Gibbs free energy, 125, 127
 GKRR-APO nanocages, 250
 Glioma-homing (gHo) peptide, 249
 Glucagon-like peptide-1 (GLP-1), 2, 16, 187
 Glutamic acid, 66
 Glycine, 101
 Gonadotropin receptor (GnR), 66
 Gonadotropin-releasing hormone (GnRH), 66
 G-protein-coupled receptors (GPCR), 5
 Gram-negative bacteria, 28
 Graphics processing unit (GPU), 84
 Grazoprevir, 273
 Greener additive, 65
 Green solvents, 65
 Growth factors, 16
 Growth hormone secretagogue receptor
 (GHSR), 247, 248
 GTPase, 268, 275
 Gut epithelium, 188

H

Hartree-Fock methods, 93
 HATU, 73
 HAV6 peptide, 258
 HCV protease inhibitor, 273
³H-daunomycin, 243, 256, 257
 Heparan sulfate proteoglycan (HSPG), 250
 Hepatic stability assays, 177
 Heptapeptide backbone, 31
 Heptapeptide ring, 31
 Heteropolymer docking, 94

HIF-1 α , 274, 275
 High-affinity inhibitors of histone deacetylases
 (HDACs), 96
 High-intensity-focused ultrasound
 (HIFU), 223
 High serum protein binding, 224
 High-throughput screening, 16, 17
 Histone deacetylases (HDACs), 123
 HIV fusion inhibitor peptide, 10
 HIV protease, 6
 HIV protease enzyme, 6
 HIV-1 protease inhibitors, 272
 HIV-1 Tat₄₉₋₅₇ peptide, 271
 Homology modeling, 4
 Hormones, 6, 16
 HPLC chromatogram
 Ac-RWVOWIO(FAM)QVR-dP-G-NH₂, 69
 cyclo (C(s-Anapa)RKSIPPR (s-Anapa)
 CFPDDDF), 74
 cyclo (SIADpPDDIK), 71
 EHWSY-dK-LRPG-NH₂, 67
 HPLC method
 and fractions, 73
 preparatory, 67, 69
 preparatory-scale, 65
 reversed-phase, 65
 H-Pro-CTC resin, 73
 Human fibroblast, 24
 Hydrofluoroalkanes (HFA), 191
 Hydrogels, 190
 Hydrogen-bonding, 20, 118
 Hydrogen fluoride (HF), 57
 Hydrophilic mucoadhesive polymers, 189
 Hydrophobic and hydrophilic amino acids, 16
 Hydrophobic motif, 31
 Hyperglycosylation, 186
 Hypoxia-inducible transcription factors
 (HIFs), 275

I

iMinDEE algorithm, 108
 Immunohistochemistry, 23
 Inflammatory diseases, 33
 Integrin-ligand interactions, 24
 Intercellular junction proteins, 254
 Intracellularly targeted peptides
 bortezomid, 273
 CFTR-CAL, 276
 clinical trials, 280
 CPP progenitor, HIV-1 TAT₄₉₋₅₇, 271
 cyclosporine A, 269
 HCV protease inhibitor, 273

- HIF1 α , 275
- HIV-1 protease inhibitors, 272
- K-Ras GTPase, 275
- MCL-1, 275
- NEMO inhibitor, 276
- p53-MDM2/4, 276
- PTP1B phosphatase, 274
- Ras farnesyl transferase inhibitor, L-744,832, 271
- Src SH2 antagonist AP22408, 270
- XIAP-BIR3, 274
- Intracellular signaling, 22
- Intranasal delivery method, 259
- Investigational new drug (IND) application, 290–293
- Iodine oxidation, 63
- Iontophoresis, 194

- J**
- Japanese encephalitis virus (JEV), 251
- Jet nebulizer, 191
- Jurkat cells, 36

- K**
- Kidney stability assays, 177
- Kinetic stability, 92
- K-Ras GTPase, 275
- KRpep-2d, 275
- KS-58, 275

- L**
- Lactam bridges, 11
- Laminin receptors, 23
- L-amino acids, 9, 95, 168
- L- and D-amino acids, 101
- Lanreotide, 175
- L-asparaginase aminohydrolase, 186
- L-captopril, 100
- Left common carotid artery (LCCA), 241
- Leukocyte-specific receptors, 23
- Ligand-binding sites, 15
- Ligands, 23
- Light-sensitive liposomes, 223
- Linaclotide, 17
- Linear peptide, 70
 - EHWSY-dK-LRPG-NH₂, 66–67
 - residue, 70, 73
- Linear tripeptide segment, 31
- Lipidation of peptides, 14
- Lipid hydration, 212
- Lipinski's Rule of Five, 242
- Lipopeptides, 28
- Lipophilic partitioning, 280
- Lipopolysaccharide (LPS), 30, 31
- Liposomes
 - active targeting, 221
 - characterization of, 212
 - formulation and manufacturing strategies
 - of, 210, 213
 - light-sensitive, 223
 - limitation
 - ABC, 224
 - aggregation of, 226
 - characterization of surface-functionalized, 225
 - high cost, 226
 - high serum protein binding, 224
 - large-scale production of surface-functionalized, 225
 - liposome clearance, 224
 - of liposome-mediated peptide drug delivery in clinical trials, 226
 - masking of surface ligands by polymers, 225
 - RES, 224
 - stringent storage conditions, 225
 - magnetic field-responsive, 223
 - for peptide drug delivery, 215, 219, 221
 - peptide liposomes composition, preparation, and characteristics, 218
 - pH-sensitive, 222
 - stability, 214
 - stabilization with lyophilization, 215
 - stimuli-responsive, 221
 - structure and composition, 209
 - thermosensitive, 222
 - ultrasound-responsive, 223
- Lisinopril, 22
- Liver microsomes, 177
- Low critical solution temperature (LCST), 222
- Low-density lipoprotein-1 (LRP1) receptors, 248
- LPS-binding affinity, 30
- LupronTM, 2
- Lupron Depot®, 187
- Lyophilization, 215

- M**
- Machine learning (ML) methods, 131–133
- Macrocyclic peptides, 268, 269, 276–278, 280
- Macrocycle, 97
- Magento-liposomes (MLs), 223

- Major histocompatibility complex (MHC), 33, 90
- MALDI mass spectrometry, 65
- MALDI-TOF, 67, 69
- Ac-RWVOWIO(FAM)QVR-dP-G-NH₂, 69
- cyclo (C(s-Anapa)RKSIPPR (s-Anapa) CFPDDF), 74
- cyclo (SIADpPDDIK), 72
- EHWSY-dK-LRPG-NH₂, 67
- Mammalian cell membrane, 29
- Mannosylation, 186
- Marine peptides, 16
- Mcl-1, 268, 271, 274, 275
- MDM2, 277
- MDM4, 277
- Medi-jector®, 195
- Membrane translocating sequences (MTSs), 272
- Mepact®, 226
- Merrifield resin, 57
- Merrifield's resin polystyrene, 57
- Metabolic diseases, 2, 14
- Metered-dose inhalers (MDI), 190, 191
- Methotrexate (MTX), 253
- 1-Methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP), 252
- 2-Methylsuccinyl-L-proline, 20
- Metropolis criterion, 87
- Metropolis-Hastings algorithm, 87
- Microneedles (MN), 193
- Mifamurtide, 226
- Minimizer, 107
- Moderna, 209
- Molecular dynamics (MD), 90–93
- Molecular Operating Environment (MOE), 96
- Molecular replacement (MR) phasing, 133
- Molecular weight, 173
- Molecular weight cut-off (MWCO), 214
- Møller-Plesset perturbation theory, 93
- Monomethoxypolyethylene glycol (mPEG), 186
- Monte Carlo methods, 87
- Monte Carlo trajectory, 117, 123
- MTT, 17
- MUC-1, 226
- Mucoadhesive polymeric systems, 190
- Mucociliary clearance, 190
- Multidrug resistance (MDR), 242
- Multidrug resistance-associated protein (MRP), 240
- Multidrug-resistant gram-negative bacterial infections, 28
- Multidrug-resistant pathogens, 28
- Multi-lamellar vesicles (MLV), 212, 216
- Multiple reaction vessels, 64
- Multiple sclerosis (MS), 238
- Multiple sequence alignment (MSA), 133
- Muramyl dipeptide (MDP), 226
- N**
- N-acetylated somatostatin analogs, 172
- NADH-quinine oxidoreductase, 30
- N- and C-termini, 10, 171
- Nanoemulsions, 189
- Nanoparticles, 250
- advantages, 206
- carbon nanotubes, 207
- dendrimers, 207
- liposomes (*see* Liposomes)
- micelles, 207
- quantum dots, 208
- Nanoscale Molecular Dynamics (NAMD), 91
- Nasal administration, 196
- Nasal cavity, 259
- Nasal delivery of peptides, 259, 261
- Native GnRH, 66
- Natural amino acids, 170
- Natural peptides, 8, 12, 16
- N-carboxyl methyl group, 21
- NCBI Reference Sequence Database, 133
- Near IR fluorescence (NIRF) imaging, 242
- Nebulizers, 190, 191
- Nectins, 252
- Nektar T-326/TOBI®Podhaler, 191
- NEMO-targeted peptide, 276
- Neurodegeneration, 14
- Neurological/cardiovascular diseases, 16
- Neurotransmitters, 6
- New Delhi metallo- β -lactamase 1 (NDM-1), 96
- New drug application (NDA), 292–294
- chemistry, 294–297
- immunogenicity and immunotoxicity, 299
- nonclinical pharmacology and toxicology, 299
- process-related impurities, 297, 298
- raw material impurities, 295, 297, 298
- stability, 298, 299
- Nicotinic acetylcholine receptor (nAChR), 250
- N-methylation, cyclic peptide, 27
- N-methylmorpholine (NMM), 68, 73
- N-methyl-2-pyrrolidone (NMP), 65
- N-octyl-N-arginine chitosan (OACS), 207
- Noncovalent π - π interactions, 175

- Noninvasive NIR fluorescence (NIRF)
 imaging, 250
- Non-labeled protein assays, 18
- Nonmalignant angiogenic disorders, 23
- NS3/4A protease inhibitors, 273
- NS5A inhibitors, 273
- N-terminal acetylation, 172
- N-terminal amino acid, 53
- Nuclear factor kappa B (NF- κ B), 276
- Nuclear magnetic resonance (NMR)
 spectroscopy, 4, 7, 9, 11, 95, 99
- Nutlins, 277
- NVP DPP728, 175
- O**
- Octreotide, 10, 169
- Oncaspar® (pegaspargase), 186
- Oncology, 2
- Opioid peptides, 246
- Oral administration, 187, 190
- Osmotic blood-brain barrier disruption
 method, 253
- OVCAR cells, 36
- OxymaPure, 65
- Oxytocin, 260
- P**
- p53, 276
- Packer palette, 106
- PackRotamersMover, 106
- Paclitaxel (PTX), 248
- PAMAM-PEG-RVG29 NPs, 251
- Paracellular pathway, 239, 252–261
- Paracellular transport, 188
- Parenteral administration, 184, 187
- Parkinson's disease (PD), 238
- Parsabiv (Etelcalcetide), 185
- Passive diffusion across BBB via transcellular
 pathway, 242, 245
- Passive transport models, 280
- PaW5-derived peptide cyclo (C(s-Anapa)
 RKSIPPR (s-Anapa)
 CFPDDF), 73, 74
- PEGylated liposomes, 224
- PEGylation, 186, 189
- Pelleted peptide, 73
- Peptide, 205, 206, 289–290
- Peptide acids, 60, 61
- Peptide amides, 60, 61
- Peptide-based ACE inhibitions, 19
- Peptide-based drug design
- agonists/antagonists, 8
 - binding assays, 17, 18
 - categories, 4
 - enzyme-based drug design, 18–21
 - known target receptors
 - agonist/antagonist, 6, 7
 - binding site and 3D structure of
 receptor cavity/surface, 6, 7
 - pharmacophore identification, 7–11
 - natural sources and modification, 16, 17
 - peptidomimetics, 12–14
 - polymyxin (*see* Polymyxin)
 and PPI (*see* Protein-protein
 interaction (PPI))
 - receptor ligand-based drug design, 22–28
 - screening assays, 17
- Peptide cleavage, 68
- Peptide conjugates, 53, 68, 248, 249
- Peptide design
- amino acids, 3
 - biochemical pathways, 5, 6
 - drug design (*see* Peptide-based
 drug design)
 - metabolic diseases, 2
 - oncology, 2
 - pharmaceutical market, 2
 - PPI, 34–40
 - and proteins, 2
 - synthetic peptides, 2
 - three-dimensional structures, 2–5
- Peptide drugs, 288, 289
- ANDA, 300–302
 - approval process, 302
 - IND application, 290–293
 - NDA, 292–294
 - abbreviated, 301
 - chemistry, 294–297
 - immunogenicity and
 immunotoxicity, 299
 - nonclinical pharmacology and
 toxicology, 299
 - process-related impurities, 297, 298
 - raw material impurities, 295, 297, 298
 - stability, 298, 299
 - post-approval, 302
- Peptide epitopes, 9
- Peptide ligand structure, 4
- Peptide macrocycle drug design
- advantages, 84
 - amino acid
 - composition, 141
 - residues, 83
 - sequence, 86

- Peptide macrocycle drug design (*cont.*)
- barriers, 82
 - binding-competent conformation, 126, 127, 129, 130
 - chemical synthesis, 124
 - clinical trials, 80
 - complexity theory, 86
 - computational design tools, 83
 - computed interaction energy, 124, 125
 - design process, 142
 - disadvantages, 84
 - DNN, 131–133
 - efficiency, 80
 - embarrassingly parallel approach, 85
 - equilibrium, 82
 - experimental validation, 130
 - FastDesign, 152
 - filters, 147
 - genic_perturb, 153
 - genic_steps mover, 147
 - Gibbs free energy, 82
 - heuristic approaches, 87–89
 - hybrid approaches, 85
 - in silico screening, 80
 - intrinsically low-throughput techniques, 124
 - layout, RosettaScripts XML script, 144
 - memory, 85
 - MinMover, 151
 - ML methods, 131–133
 - molecules, 80, 81
 - Monte Carlo refinement, 153
 - NDM-1
 - design script, 149
 - inhibitor design, 145
 - inhibitor drug design, 136
 - NP-complete, 86
 - PACKER_PALETTES, 148
 - pipelines (*see* Pipelines, peptide macrocycle drug design)
 - polynomial-time algorithm, 86
 - preclinical evaluation, 80
 - processes, 85
 - quantum computing, 134–136
 - RESIDUE_SELECTORS section, 150
 - role of computation, 81
 - RosettaScripts, 139
 - Rosetta Software, 137, 138
 - softdesign, 148
 - software approaches (*see* Software approaches, peptide drug modelling)
 - software methods, 84
 - synchronization, 85
 - synthetic peptides, 81
 - XML scripts, 138
- Peptide mapping, 8
- Peptide modifications
- analysis and purification, 65
 - automation of SPPS, 64
 - biotin-labeled lysine, 61
 - cyclization, 63
 - disulfide bond, 63, 64, 72
 - enzymology, 61
 - fluorescent and biotin labeling, 61
 - histochemistry, 61
 - immunology, 61
 - protein chemistry, 61
 - sustainability in peptide synthesis, 65
- Peptide purification, 66
- Peptide stability and prolong half-life
- backbone modification strategies, 171
 - blood, serum, and plasma stability assays, 176
 - cyclization, 173
 - enzyme inhibition, 175
 - flip-flop kinetics, 175
 - gastric intestinal stability assays, 178
 - hepatic stability assays, 177
 - increase molecular weight, 173
 - kidney stability assays, 177
 - modification strategies for, 167, 169
 - N- and C-termini, 171
 - noncovalent π – π interactions, 175
 - replacing L-amino acids with D-amino acids, 168
 - tissue stability assays, 177
- Peptide synthesis
- amino acids, 52
 - coupling chemistry, 60
 - coupling reagents and additives, 60, 62
 - C-terminal functional groups, 60
 - cyclic peptide cyclo (SIADpPDDIK), 70–72
 - dipeptide glycine-glycine, 52
 - fluorescently labeled peptide
 - Ac-RWVOWIO(FAM)QVR-dP-G-NH₂, 68–69
 - Fmoc-protected amino acid, 60
 - linear peptide EHWSY-dK-LRPG-NH₂, 66–67
 - Merrifield's resin polystyrene, 57
 - modifications (*see* Peptide modifications)
 - PawS-derived peptide cyclo (C(s-Anapa)RKSIPPR (s-Anapa)CFPDDF), 73, 74
 - peptide acids, 60, 61
 - peptide amides, 60, 61

- polymeric resins, 60
- property, 57
- proteins, 52
- resin's loading capacity, 60
- solution-phase, 53–55
- SPPS, 54, 56–59
- in sustainability, 65
- Peptidomimetics, 15, 18, 268, 269, 277, 278, 280
 - backbone/side-chain modification, 12, 13
 - classes, 12
 - definition, 12
 - secondary structure mimics, 14
 - synthesis, 13
- P-glycoprotein (Pgp), 240, 242–245, 256, 257
- Phage display method, 249
- Pharmacokinetic studies, 28
- Pharmacophore identification
 - alanine scanning, 8
 - bioanalytical assays, 7
 - conformational constraints, 9
 - cyclization, 10–12
 - D-amino acids, 9
 - molecular biology, 7
 - N- and C-termini modification, 10
 - peptide mapping, 8
 - receptor-ligand interactions, 7
 - structure of biologically active conformation, 9
 - 3D structure, 8
 - truncation of peptide, 8
- Pharmacophore portion, 9
- Phosphatase, 268, 270, 274
- Phosphoethanolamine (PEtn), 32
- Phospholipid micelles, 207
- Photodiode assay detector, 66
- pH-sensitive liposomes, 222
- Pipelines, peptide macrocycle drug design
 - amino acids, 109, 111, 112
 - approaches for flexible-backbone design, 107–109
 - design-centric guidance functions, 115
 - designing peptide sequence, 104
 - energy function, 110
 - filtering, 116–118
 - fixed-backbone design problem, 104–107
 - gradient-descent energy minimization, 115
 - high-cost simulations, 98
 - initial stub identification, 100, 101
 - isoleucine and lysine residues, 113
 - Lennard-Jones term, 114, 115
 - linear function, 112
 - Monte Carlo search, 110
 - pairwise-decomposable energy function, 110
 - peptide macrocycle therapeutics, 113
 - peptide-protein interface, 114
 - physical properties, 109
 - polar interfaces, 115
 - process of designing peptide, 109
 - PyRosetta, 109
 - reference energy, 111, 112
 - residue selector, 110
 - Rosetta scripts protocol, 118–123
 - sampling peptide conformation, 101–103
 - self-assembling peptides, 113
 - side-chain rotamers, 109
 - target modelling, 98, 99
 - user-defined magnitude, 113
- Pituitary adenylate cyclase-activating polypeptide (PCAP), 260
- Plasma or serum, 176
- Platelet aggregation assay, 24, 25
- Plenaxis (abarelix), 10, 185
- Polyacrylamide/PEG copolymer resins, 60
- Polyamidoamine dendrimers (PAMAM), 250
- Polyethylene cross-linked PEG resins, 60
- Polyethylene glycol (PEG), 57–60, 250
- Poly(lactic acid) (PLA), 192
- Poly(lactic-co-glycolic acid) (PLGA), 192
- Polymeric micelles, 207
- Polymeric resins, 60
- Polymyxin
 - AMPs, 29, 30
 - cytoplasmic membranes, 29
 - Dab residues, 30
 - gram-negative bacteria, 28
 - lipopeptides, 28
 - LPS, 30
 - mammalian cell, 30
 - mechanism of action, 29, 30
 - multidrug-resistant gram-negative bacterial infections, 28
 - multidrug-resistant pathogens, 28
 - non-ribosomal cyclic lipopeptides, 29
 - polymyxin B, 28, 29
 - polymyxin E, 28, 29
 - SAR (*see* Structure-activity relationship (SAR))
 - structure, 29
 - toxicity and modification, 32
- Polymyxin B, 28, 29, 31
- Polymyxin E, 28, 29
- Polypeptide, 3, 289–290
- Positron emission tomography (PET), 23
- PowderJect® device, 195
- Pramlintide, 17
- Preparatory-scale HPLC, 65
- Pro-Gly sequence, 37
- Proline, 20

- Pro-pro sequences, 37, 70
 Protease, 272, 273, 278, 280
 Proteasome inhibitor, 273
 Protecting group, 53–61, 63, 66, 70, 72, 73
 Protein, 2, 52, 290
 Protein chemistry, 52
 Protein Data Bank (PDB), 95
 Protein-protein interactions (PPI), 6, 7, 14, 15, 22, 24, 70, 276, 278, 280
 - APC, 33
 - autoimmune disease, 33
 - biochemical pathway, 33, 34
 - design of peptides, 34–40
 - first-generation peptides, 34–37
 - immune system, 32
 - MHC, 33
 - protein molecules, 33
 - screening assay, 36, 37
 - second-generation peptides, 37–39
 - T cells, 33
 - TCR, 33
 - TCR-MHC complex, 34
 - third-generation peptides, 39–41
- Protein receptors, 18
 Protein transduction domains (PTDs), 272
 Proteolysis, 167, 168, 172–175
 PTP1B phosphatase, 274
 Pulmonary route, 190
 pVEC peptide, 249
- Q**
- QSAR models, 277
 Quantum computing, 134–136
 Quantum dots, 208
 Quantum mechanics (QM), 93, 94
- R**
- Rabies virus glycoprotein (RVG), 250
 Ras farnesyl transferase inhibitor, L-744,832, 271
 Ras proteins, 271
 Rational design, 16
 Rational drug design, 3, 15, 99
 Rat kidney cells, 24
 Receptor-based drug design, 18
 Receptor cavity/surface, 6, 7
 Receptor ligand-based drug design
 - biochemical basis, 22–23
 - identification of binding epitope, 24–26
 - N-methylation, 27
 - preclinical and clinical evaluation, 27, 28
 - structural and development, 22
- Receptor-ligand interactions, 7
 Receptor-mediated transcytosis, 245, 248
 Renin-angiotensin system, 5, 18, 19
 Residue-level pairwise decomposability, 105
 Resin, 54
 Resin's loading capacity, 60
 Resistance, 32
 Resimat, 191
 Retro-inverso peptide, 240, 251
 Reversed-phase HPLC, 65
 RGDFV sequence, 25, 26
 RGD motif, 24, 26
 RGD peptides, 24
 RGD receptors, 23
 RGDS peptide, 24
 Rhodamine 800 (R800), 257
 Rink amide resin, 68
 Root mean squared deviation (RMSD), 127
 Rosetta FastRelax, 108
 Rosetta LoopModel, 99
 Rosetta Next-Generation-KIC, 99
 Rosetta remodel, 99
 Rosetta Scripts Protocol, 118–123
 Rosetta's GALigandDock module, 101
 Rosetta Software, 94–97, 137, 138
 Rotamer, 104
 Rotamer optimization algorithms, 117
 RVG29 peptide, 252
 Rybelsus (Semaglutide) oral tablets, 187
- S**
- Sandimmune/Neoral® (Cyclosporin A), 190
 Saquinavir, 272
 Scanning electron microscopy (SEM), 212
 Screening assays, 17, 36, 37
 sCT liposomes, 220
 Second-generation peptides, 37–39
 Self-emulsifying drug delivery systems (SEDDS), 189
 Serum, 176
 SFTI-1 framework, 72
 Sheep red blood cells (SRBC), 36
 Short-chain peptides, 9
 Side-chain cyclization, 11
 Side-chain torsion angles, 4
 Signifor Lar® (Novartis), 187
 Small-molecule, 166
 Small-molecule docking, 94
 Sodium deoxycholate (SDC), 217
 Sodium glycocholate (SGC), 217
 Sodium taurocholate (STC), 217
 Software approaches, peptide drug modelling
 - amino acids, 89
 - atomic resolution, 90
 - biased semi-random libraries, 89

- database mining, 90
 - linear programming approach, 90
 - macromolecular modelling software, 96
 - MD, 90–93
 - QM, 93, 94
 - The Rosetta software suite, 94–97
 - Solid-phase peptide synthesis (SPPS), 54, 56–59, 64
 - Solution-phase peptide synthesis, 53–55
 - SonoDerm™, 195
 - Sonophoresis, 195
 - Sonoprep®, 195
 - Spatial screening, 26
 - Src SH2 antagonist AP22408, 270
 - Stapled peptides, 186
 - Staples, 186
 - Stimuli-responsive liposomes, 221
 - Structure-activity relationship (SAR), 16, 17, 20
 - Dab side chain, 31
 - fatty acid chain and substitution, 30
 - formulations and administration, 32
 - heptapeptide backbone, 31
 - hydrophobic motif of heptapeptide ring, 31
 - linear tripeptide segment, 31
 - resistance, 32
 - toxicity and modification of polymyxin, 32
 - Supervised machine learning, 132
 - Synthetic peptide, 196
- T**
- Target-based rational design, 5
 - T-cell APC contact, 34
 - T-cell receptors (TCR), 33
 - T cells, 33
 - TCR-MHC complex, 34
 - Temperature-sensitive liposomes, 222
 - Tetraether lipids (TELs), 216
 - Theranostic approach, 10
 - Therapeutic peptide administration
 - different routes of administration, 185
 - via nasal delivery, 196
 - via oral route, 187, 190
 - via parenteral route, 184, 187
 - via pulmonary route, 190, 193
 - via rectum, 196
 - via transdermal route, 193, 196
 - via vaginal route, 197
 - Thermosensitive liposomes, 222
 - Third-generation peptides, PPI, 39–41
 - Thr38-Ser39-Asp40-Lys41, 35
 - Tissue stability assays, 177
 - Topical/transdermal administration, 193, 196
 - Transactivator of transcription (TAT) peptides, 248
 - Transcytosis, 246
 - Trans-endothelial electrical resistance (TEER) values, 241
 - Transferrin receptor (TfR), 246
 - Transferrin receptor-1 (TfR1), 250
 - Transmembrane domain, 22, 23
 - Transmission electron microscopy (TEM), 212
 - Transwell™, 241
 - Tribute Peptide Synthesizer, 66, 68, 70, 73
 - Trulance™ (Plecanatide), 187
 - Truncation of peptide, 11
 - Tumor cell expression, 23
 - Type 2 diabetes mellitus (T2DM), 167
- U**
- U-81749, 272
 - Ultrasonic nebulizer, 191
 - Ultrasound-responsive liposomes, 223
 - Undruggable sites, 15
 - Unified Decomposition-Guided Variable Neighbour Search (UGDVNS), 106
 - Unnatural amino acids, 170
 - US Food and Drug Administration (FDA), 184
- V**
- Valence electrons, 93
 - Vascular epithelial growth factor receptor 2 (VEGFR2), 215
 - Vasoactive intestinal peptide (VIP), 217, 260
 - VE-cadherins, 252
 - ViaDor™ system, 194
 - VIP-loaded liposomes (VLL), 219
 - Vitajet™, 195
- X**
- X-chromosome-linked inhibitor of apoptosis (XIAP), 274
 - Xemys, 226
 - Xenograft mice model, 28
 - X-ray crystallographic methods, 9
 - X-ray crystallography, 4, 7, 96, 99
 - X-ray crystal structures, 98, 99
- Z**
- Zeta potential, 214
 - Ziconotide, 17
 - Zomajet®, 195