Topics in Medicinal Chemistry 21

Zuben E. Sauna
Chava Kimchi-Sarfaty *Editors*

Protein Therapeutics

Springer

21
Topics in Medicinal Chemistry

Aims and Scope

Drug research requires interdisciplinary team-work at the interface between chemistry, biology and medicine. Therefore, the new topic-related series Topics in Medicinal Chemistry will cover all relevant aspects of drug research, e.g. pathobiochemistry of diseases, identification and validation of (emerging) drug targets, structural biology, drugability of targets, drug design approaches, chemogenomics, synthetic chemistry including combinatorial methods, bioorganic chemistry, natural compounds, high-throughput screening, pharmacological in vitro and in vivo investigations, drug-receptor interactions on the molecular level, structure-activity relationships, drug absorption, distribution, metabolism, elimination, toxicology and pharmacogenomics.

In general, special volumes are edited by well known guest editors.

In references Topics in Medicinal Chemistry is abbreviated Top Med Chem and is cited as a journal.

More information about this series at http://www.springer.com/series/7355

Zuben E. Sauna • Chava Kimchi-Sarfaty
Editors

Protein Therapeutics

With contributions by

S. Ahuja · J. Bender · M.R. Chéhadé · S. Ehrenforth · D. Hatton · A. Hunter · N. Kirschbaum · A.A. Komar · K. Lamberth · J. Liu · E. Marszal · H. Østergaard · K. Ram · A.S. Rathore · M. Shapiro · S.K. Singh · E.B. Struble · R. Turner · K.N. Weldingh

Springer

Editors
Zuben E. Sauna
Division of Plasma Protein Therapeutics
Office of Tissues and Advanced Therapies
FDA | CBER | OTAT
Federal Research Center at White Oak
Silver Spring, Maryland, USA

Chava Kimchi-Sarfaty
Division of Plasma Protein Therapeutics
Office of Tissues and Advanced Therapies
FDA | CBER | OTAT
Federal Research Center at White Oak
Silver Spring, Maryland, USA

ISSN 1862-2461 ISSN 1862-247X (electronic)
Topics in Medicinal Chemistry
ISBN 978-3-319-41816-2 ISBN 978-3-319-41818-6 (eBook)
DOI 10.1007/978-3-319-41818-6

Library of Congress Control Number: 2017945394

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Introduction

Chava Kimchi-Sarfaty, Aikaterini Alexaki, and Zuben E. Sauna

Protein therapeutics were first developed only a few decades ago but now dominate pharmaceutical sales. For example, protein therapeutics accounted for about 7% of revenues from the top ten best selling drugs in 2001 but generated 70% of revenues a decade later [1]. This class of drugs which represents a core component of modern pharmacotherapy thus includes some of the most expensive drugs on the market. Tables 1 and 2 provide an overview of the diversity of protein molecules used as therapeutics and the different platform technologies used.

Table 1 Number of recombinant therapeutics (not including antibodies) that are marketed or under development

	Cytokines	Hormones	Coagulation factors	Inhibitors	Enzymes	Status
Unmodified human protein	34	16	13	5	17	Marketed
	47	32	21	8	35	In pipeline
	6	7	3(1)	1(1)	2	Biosimilars
Pegylated protein	4	1	1	2	2	Marketed
				2	7	In pipeline
	1(1)					Biosimilars
Polyxen fusion protein						Marketed
	2	2				In pipeline
						Biosimilars

(continued)

C. Kimchi-Sarfaty • A. Alexaki • Z.E. Sauna
Division of Plasma Protein Therapeutics
Office of Tissues and Advanced Therapies
FDA | CBER | OTAT
Federal Research Center at White Oak
Silver Spring, Maryland, USA

Table 1 (continued)

	Cytokines	Hormones	Coagulation factors	Inhibitors	Enzymes	Status
Fc fusion protein			2	7	1	Marketed
				3		In pipeline
				(1)		Biosimilars
Albumin fusion protein	2	1	2			Marketed
				1		In pipeline
						Biosimilars
XTEN fusion protein						Marketed
		6	3		1	In pipeline
						Biosimilars

Numbers in parenthesis indicate biosimilars under development which have not yet been approved. Data were derived from Pharmaprojects (https://fda-pipeline.citeline.com)

Table 2 Number of therapeutic antibodies that are marketed or under development

	Marketed	In pipeline	Marketed biosimilars
Unconjugated antibody	52	118	1(19)
Pegylated antibodies	1	3	1
Toxin-conjugated antibodies	3	17	
Radio-immuno-conjugates	10	5	

Numbers in parenthesis indicate biosimilars under development which have not yet been approved. Data were derived from Pharmaprojects (https://fda-pipeline.citeline.com)

Proteins differ from small molecule drug both in terms of their characteristics (see below) and the manner in which they are manufactured. For example, a typical protein therapeutic is much larger, exhibits complex secondary and tertiary structures, and cannot be synthesized by chemical processes. As the products are synthesized by cells, complex extraction and purification processes are involved which could potentially introduce modifications in the protein. It is therefore not surprising that the manufacture of a typical protein therapeutic is a far more complex process than that of a small molecule drug [2]. A sense of just how much more complex manufacturing a protein therapeutic can be is illustrated by the information provided in Table 3. Compared to small molecule drugs, the number of batch records, product quality tests, critical process steps, and process data entries are all at least an order of magnitude higher for protein therapeutics.

Table 3 The complexity of manufacturing protein therapeutics compared to small molecule drugs

	Small molecule drugs	Protein therapeutics
Batch records	<10	>250
Product quality tests	<100	>2,000
Critical process steps	<100	>5,000
Process data entries	<4,000	>60,000

Several forces at play in recent years have added even more complexity to the drug-development and manufacturing processes for protein therapeutics. Experience with the first generation of native or wild-type proteins indicated that there was an unmet need for molecules with better clinical outcomes, improved patient convenience, or simplified and more reliable manufacturing processes [3]. Driven by scientific innovations and new technologies, a new generation(s) of bioengineered protein molecules, that seek to fulfill these needs, has entered the drug-development pipeline [4]. These advances have come at the same time as the manufacturing process and sourcing of materials has become more global. Together these changes have added additional challenges to the development, licensure, and manufacture of protein therapeutics. This book provides a high level view of what the specific challenges are and how they are being met.

Characteristics of Protein Therapeutics and How They Differ from Small Molecule Drugs

Characterization by analytical methods is generally considered to be a good predictor of the biological and clinical properties of small molecule drugs. This experience, for example, is the key reason why generic versions of these drugs can be rapidly and inexpensively developed and marketed. The same is not true for protein therapeutics due to the following distinctive characteristics.

Size: The most prominent difference between a small molecule drug and a protein therapeutic is the size; the latter being 100–1,000 times larger. Due to their size and complexity, currently, protein therapeutics cannot be synthesized by chemical processes and have to be manufactured in living cells (see the chapters, "Protein Production in Eukaryotic Cells" and "Production of Protein Therapeutics in the Quality by Design (QbD) Paradigm"). Thus cell characteristics such as choice of the cell line, species origin of the cell line, and culture conditions all affect the final product characteristics [5]. In addition the use of materials of biological origin increases the potential risk of the final product. Finally the large size of the product complicates drug delivery as well as storage and distribution. Proteins are also much more likely to elicit an immune response in patients which is an important safety issue (see the chapter "Immunogenicity Lessons Learned from the Clinical Development of Vatreptacog Alfa, a Recombinant Activated Factor VII Analog, in Hemophilia with Inhibitors") [6].

Structure: To possess biological activity, proteins have to adopt the correct three-dimensionally folded secondary, tertiary, and quaternary structures. Thus the development, regulation, and manufacture of protein therapeutics all require very sophisticated analytical techniques (see the chapter "Characterization of Therapeutic Proteins"). Analytical techniques are becoming increasingly sophisticated but this also means that more biophysical or biochemical changes that are not clinically relevant or are not a safety issue can be identified. Quality-by-Design (QbD) [7] offers a means of identifying and monitoring critical quality attributes (see chapter "Production of Protein Therapeutics in the Quality by Design (QbD) Paradigm").

Structure–Function relationship: In small molecules, it is often known that every atom of the molecule will play a role in defining the clinical profile of the compound; the structure–function relationship is usually unknown, or at best partially known, for proteins. Thus, the impact of differences in the molecular structure in most cases cannot be predicted [8]. This can result in safety issues late in the drug-development cycle (see the chapter "Immunogenicity Lessons Learned from the Clinical Development of Vatreptacog Alfa, a Recombinant Activated Factor VII Analog, in Hemophilia with Inhibitors" for an example).

Stability: Proteins are inherently unstable molecules, and may be altered structurally by heat, prolonged storage, denaturants, organic solvents, oxygen, pH changes, and other factors, which are all part of the manufacturing process [9]. This can be a consistent challenge and emerging strategies like QbD [2] can help to maintain consistency at a manufacturing facility and also when manufacturing facilities are moved or added (see the chapter "Production of Protein Therapeutics in the Quality by Design (QbD) Paradigm"). Proteins can be bioengineered [10] to enhance their utility as drugs. However, this can be a double-edged sword and results in unanticipated and undesirable outcomes (see the chapters "Immunogenicity Lessons Learned from the Clinical Development of Vatreptacog Alfa, a Recombinant Activated Factor VII Analog, in Hemophilia with Inhibitors" and "The Art of Gene Redesign and Recombinant Protein Production: Approaches and Perspectives").

Microheterogeneity: Proteins are modified both biologically by the producing cell (e.g., glycosylation, acylation, sulfatation, phosphorylation, and proteolysis) and by the process conditions (e.g., oxidation, deamidation, reaction with auxiliary substances, partial denaturation, and aggregation) [9]. Further heterogeneity may arise if the protein is intentionally modified, for example, by multi-site pegylation [11]. Thus, even highly purified proteins never consist of a single molecular entity but are mixtures of many closely related molecular species. This microheterogeneity can be substantial. It has been estimated that up to 10^8 different species exist in an immunoglobulin G molecule [2]. This inherent variability in the reference molecule itself makes the identification of clinically relevant variations extremely difficult. Thus the phrase, "the process is the product" [5, 8, 12] is often used in the context of the manufacture of protein therapeutics. The emphasis therefore has been on the genetic stability of the expression system and the reproducibility of the production process.

Protein Therapeutics and Immunogenicity

A significant concern unique to the development and licensure of protein therapeutics is the risk of developing anti-drug antibodies (ADAs). Such antibodies are rarely, if ever, a concern during the development of small molecule drugs. However, the development of ADAs against protein therapeutics can lead to adverse events and also make the biologic less effective for its intended use. Thus, immunogenicity assessments are now an integral part of the development, licensure, and use of this class of products [6]. Factors influencing immunogenicity against a

protein therapeutic are both patient- and product-related [13] and the latter are often associated with the manufacturing process. There are several examples where small changes in the protein or manufacturing processes have resulted in significant increase in immunogenicity [14, 15]. The chapter "Immunogenicity Lessons Learned from the Clinical Development of Vatreptacog Alfa, a Recombinant Activated Factor VII Analog, in Hemophilia with Inhibitors" provides a detailed case study where three amino acid substitutions in Factor VIIa resulted in the development of ADAs in >10% of patients in the phase 3 trial. In contrast the parent molecule which was not engineered has been used as a drug for almost two decades with no reports of ADAs, consequently the development of this bioengineered analog was discontinued [15].

Several factors influence the immunogenicity of a protein product. The presence of impurities such as host-cell impurities (host-cell proteins, DNA, lipids, viruses, or endotoxins), protein aggregates, or leachates can affect immunogenicity. Also of concern are sequence modifications made in protein products that are bioengineered to improve yields or product characteristics such as increased circulating half-life [3]. Methodology used for measuring immunogenicity includes immunoreactivity assays (radioimmunoassay, surface plasma resonance, or enzyme-based solid-phase immunoassay) or functional cell-based bioassays. It must be emphasized that despite considerable technological progress in non-clinical approaches the current state of the art does not permit immunogenicity assessments to be made in the absence of clinical trials. This again emphasizes the importance of developing manufacturing processes that are consistent and identify and monitor critical quality attributes.

Protein Therapeutics: Biosimilars

The use of proteins as therapeutics has revolutionized the treatment of many disease areas but these medications are some of the most expensive in the market place. As many biopharmaceutical products are poised to go off patent it has been recognized that replicating the highly successful generic model to contain the costs of these therapeutics is a desirable goal [7, 12]. However, the primary function of regulatory agencies is to ensure that patient safety is not compromised. Given the complexity of protein therapeutics, as well as of the manufacturing process it is unlikely that in the near term the development process for biosimilars can be abridged to quite the extent as that for classical generics [16, 17]. Significant challenges remain in developing analytical techniques to comprehensively characterize protein therapeutics. Moreover, unlike small molecule drugs protein therapeutics exhibit considerable micro-heterogeneity and thus development of more sensitive and accurate technological analytical tools alone may not be sufficient. It has been recognized that biosimilars are not generics (as the nomenclature implies) and will not be identical to the reference drug.

We do not address the thorny legal and scientific issues surrounding biosimilars. However, the difficult questions surrounding biosimilars arise due to the characteristics and complexities of protein molecules and these are adequately addressed in this volume.

Overview of This Book

In this book we have endeavored to provide a broad overview of developing and manufacturing therapeutic proteins. The individual chapters written by experts can be used as a source of information on specific topics. However, the book as a whole also provides a narrative that describes the art and science of developing a protein therapeutic in a rapidly globalizing marketplace. The book begins with the basics; Ram et al. describe the nuts and bolts of manufacturing a recombinant protein in eukaryotic cells. However, protein therapeutics are increasingly being manufactured in a global setting. This means that the same product could be manufactured at different locales; parts of the manufacturing process may be outsourced, etc. Managing the quality and consistency of a complex product in such a setting is critically important and extremely challenging. Rathore and Singh introduce the concept of QbD in the context of protein therapeutics. The importance of identifying the underlying relationship between the quality attributes of the product and clinical safety and efficacy is the ultimate goal of QbD and likely to play a critical role in maintaining product quality in an increasingly global market. Protein therapeutics require complex and sophisticated manufacturing processes; but the molecules themselves are also inherently complex. Struble et al. provide a comprehensive survey of the tools and strategies for the characterization of proteins. More importantly they discuss these in the context of the regulatory framework which is essential for translating a molecule with promise into a successful drug. The ability to engineer proteins permits the incorporation of characteristics sought after in a drug such as enhanced serum half-life, a better safety-efficacy profile, patient convenience, and delivery to target. However, these manipulations can also sometimes result in unintended consequences and termination of the drug-development process. Rather than an abstract discussion of this topic Lamberth et al. present a case study where a Factor VIIa analog with an improved safety-efficacy profile was discontinued from further development because of the identification of unwanted anti-drug antibodies in phase 3 trials. Finally, Komar provides an in-depth discussion of a single platform technology, namely codon optimization and discusses the potential consequences (both desirable and potentially hazardous) based on rapid, recent progress in basic sciences.

Disclaimer

Our contributions are an informal communication and represent our own best judgment. These comments do not bind or obligate FDA.

References

1. Waltz E (2014) It's official: biologics are pharma's darlings. Nat Biotechnol 32:1
2. Rathore AS (2009) Follow-on protein products: scientific issues, developments and challenges. Trends Biotechnol 27:698–705
3. Kimchi-Sarfaty C, Schiller T, Hamasaki-Katagiri N, Khan MA, Yanover C, Sauna ZE (2013) Building better drugs: developing and regulating engineered therapeutic proteins. Trends Pharmacol Sci 34:534–548
4. Kontermann RE (2011) Strategies for extended serum half-life of protein therapeutics. Curr Opin Biotechnol 22:868–876
5. Schellekens H (2004) When biotech proteins go off-patent. Trends Biotechnol 22:406–410
6. Shankar G, Shores E, Wagner C, Mire-Sluis A (2006) Scientific and regulatory considerations on the immunogenicity of biologics. Trends Biotechnol 24:274–280
7. Rathore AS, Winkle H (2009) Quality by design for biopharmaceuticals. Nat Biotechnol 27:26–34
8. Kresse GB (2009) Biosimilars–science, status, and strategic perspective. Eur J Pharm Biopharm 72:479–486
9. Jenkins N, Murphy L, Tyther R (2008) Post-translational modifications of recombinant proteins: significance for biopharmaceuticals. Mol Biotechnol 39:113–118
10. Lazar GA, Marshall SA, Plecs JJ, Mayo SL, Desjarlais JR (2003) Designing proteins for therapeutic applications. Curr Opin Struct Biol 13:513–518
11. Caliceti P, Veronese FM (2003) Pharmacokinetic and biodistribution properties of poly(ethylene glycol)-protein conjugates. Adv Drug Deliv rev 55: 1261–1277
12. Schneider CK, Kalinke U (2008) Toward biosimilar monoclonal antibodies. Nat Biotechnol 26:985–990
13. De Groot AS, Moise L (2007) Prediction of immunogenicity for therapeutic proteins: state of the art. Curr Opin Drug Discov Devel 10:332–340
14. Chirino AJ, Mire-Sluis A (2004) Characterizing biological products and assessing comparability following manufacturing changes. Nat Biotechnol 22:1383–1391
15. Mahlangu JN, Weldingh KN, Lentz SR, Kaicker S, Karim FA, Matsushita T, Recht M, Tomczak W, Windyga J, Ehrenforth S et al (2015) Changes in the

amino acid sequence of the recombinant human factor VIIa analog, vatreptacog alfa, are associated with clinical immunogenicity. J Thromb Haemost 13:1989–1998
16. Kozlowski S, Woodcock J, Midthun K, Sherman RB (2011) Developing the nation's biosimilars program. N Engl J med 365:385–388
17. Mellstedt H, Niederwieser D, Ludwig H (2008) The challenge of biosimilars. Ann Oncol 19:411–419

Contents

Top Med Chem (2017) 21: 1–40
DOI: 10.1007/7355_2016_3

Published online: 26 April 2016

Protein Production in Eukaryotic Cells

Kripa Ram, Diane Hatton, Sanjeev Ahuja, Jean Bender, Alan Hunter, and Richard Turner

Abstract The scientific and engineering aspects of design, development, scale-up, and manufacture of monoclonal antibodies are summarized in this chapter by outlining the key elements in the development of the expression cell line, cell culture, cell harvest, and protein purification process and exploring the effect of process technologies on production economics.

Keywords Bioreactor, Bioreactor sterility, Cell culture scale up, Cell harvest, Cell line screening, Chinese hamster ovary cell line, Continuous centrifugation, Cost modelling, Cost of goods, Depth filtration, Disc stack centrifuge, Economies of scale, Genetic stability, Host cell protein, Ion exchange, Medium optimization, Phenotypic stability, Process economics, Protein expression, Run rate, Sigma factor chromatography, Viral clearance

Contents

K. Ram (✉), S. Ahuja, J. Bender, and A. Hunter
MedImmune, Gaithersburg, MD, USA
e-mail: ramk@medimmune.com

D. Hatton and R. Turner
MedImmune, Cambridge, UK

1 Introduction

As the role of biologics in treating human diseases has evolved dramatically over the past decade, so has the technology to manufacture, test, and administer these products. This has been driven largely by improved understanding of the biological mechanism of action, the role of the structure of the molecules on the mechanism, and the role of the manufacturing process on the structure of the molecule. Among the various types of biologics, monoclonal antibodies now represent the largest segment of biopharmaceutical proteins in terms of sales. A 2015 study shows 8,182 kg of biopharmaceutical monoclonal antibody products were produced, representing nearly $60 billion in sales in 2013 [1]. Given this large demand for mAb products, along with increasing price pressures, and most of all an urgency to bring these biopharmaceuticals to patients, process scientists and engineers have developed “platform” processes with extremely high productivity; analytical scientists have developed sophisticated techniques to decipher attributes critical to quality; formulation scientists have developed formulations that allow storage at high concentrations for several years; and device engineers continue to develop modes of administration more convenient to the patient.

Depending on the commercial demand for the product and the available manufacturing facility, different manufacturing technologies are employed. The most prevalent among these is the fed-batch mode of bioreactor operation and batch mode of chromatography operation. However, the use of continuous manufacturing that relies on perfusion-based bioreactor operation and continuous flow chromatography operation is being developed to increase plant throughput [2]. This chapter will focus on batch manufacturing technology in order to anchor the reader in the most common process scheme used today. Also, this chapter will focus on manufacture of drug substance only. The reader is encouraged to refer to a review on recent progress in formulation development for protein therapeutics by Razinkov [3].

The scientific and engineering aspects of design, development, scale-up, and manufacture of monoclonal antibodies are summarized in this chapter by outlining the key elements in the development of the expression cell line, cell culture, cell harvest, and protein purification process and exploring the effect of process technologies on production economics.

2 Cell Line

2.1 Choice of a Host Cell Line

Protein therapeutics are generally produced from a recombinant production cell line, which is constructed by the introduction of the DNA encoding the therapeutic protein into a host cell line and harnessing the synthetic capacity of the cell to express the recombinant protein. The choice of host cell line determines critical aspects of therapeutic protein production, including yield, purity and quality of the protein, timelines, scale-up, and cost of goods.

Simple polypeptides, such as insulin, can be produced from bacterial systems, such as *Escherichia coli* (*E. coli*), which, owing to its short doubling time, has the advantages of rapid development of recombinant cell lines as well as low cost and short production bioreactor processes (reviewed by [4]). However, there are significant limitations to *E. coli* expression systems owing to the inability to perform disulfide bonding to assemble complex proteins and the lack of posttranslational modifications (PTMs) such as glycosylation, carboxylation, hydroxylation, sulfation, and amidation [5].

In contrast to bacterial cells, eukaryotic cells are equipped with the cellular machinery for the folding and assembly of complex proteins, as well as for performing PTMs. Yeast and insect cells can be used for production of therapeutic proteins [6] and, owing to their rapid cell division and scalability, have some of the same advantages as bacterial expression systems. However, they produce glycoproteins with carbohydrate structures that are different from human-type glycosylation and have the potential to impact both the in vivo activity and immunogenicity

of expressed proteins [5]. Systems for plant-made pharmaceuticals (PMPs) have been developed using transgenic plants [7], transgenic moss [8], or plant cell suspension cultures [9]. Again, some of the natural plant glycan structures are species specific, containing terminal beta (1,2) xylose and alpha (1,3) fucose residues that are not found in humans [10]. Efforts are ongoing to engineer PTMs in microbial [11], insect [12], and plant [13] cell systems in order to make these systems more suitable for protein therapeutic production.

Immortalized mammalian cells are most commonly used for recombinant therapeutic protein production [14]. Rodent host cell lines provide more humanlike glycosylation; they include mouse myeloma (NS0 and Sp2/0), baby hamster kidney, and Chinese hamster ovary (CHO) cells. However, these rodent cell lines also produce nonhuman alpha-gal and *N*-glycolylneuraminic acid (NGNA) glycoform structures that can affect in vivo clearance and immunogenicity, thereby impacting the pharmacokinetics of the therapeutic protein [15, 16]. Avoiding this issue, there are human host cell lines available for the production of fully human PTMs (reviewed by [17, 18]). Cell lines derived from the human embryonic kidney cell line (HEK-293) are used for the production of a number of approved therapeutic proteins, including recombinant clotting factors and fusion proteins, where additional PTMs such as gamma-carboxylation and sulfation are important for bioactivity [19]. The human HT-1080 cell line, derived from a fibrosarcoma, is used for the production of approved enzyme therapies, iduronate-2-sulfatase, agalsidase alfa, and velaglucerase alfa. Other human cell lines, PER.C6 (derived from human embryonic retinal cells; [20]) and CAP-T (derived from human amniocytes; [21]), are also being used to produce therapeutic proteins that are currently in preclinical and clinical development. Although human host cell lines offer advantages in terms of PTMs and associated bioactivity, there are theoretical concerns about the potentially increased risk of contamination of therapeutic proteins with adventitious viruses capable of infecting human patients. However, this has been addressed by incorporating the appropriate mitigation strategies during manufacturing, combining viral inactivation and clearance steps with suitable testing strategies [17, 19].

The most commonly used mammalian expression host is that of CHO cells and accounts for the production of >70% of approved therapeutic proteins [22], building a strong regulatory track record for this expression host cell line. As a rodent cell line, there is a species barrier to the production of viruses that can infect humans, and studies have confirmed that CHO cells are resistant to infection with many viruses that can infect humans [23]. Furthermore, the development of CHO-based expression platforms has enabled the removal of animal-derived components, such as serum, from the cell culture medium for cell line development and bioreactor processes, both reducing costs and concerns about adventitious virus contamination. In addition, not only are CHO cells amenable to the introduction of therapeutic protein genes, additional genetic engineering can be performed to modify the growth and metabolism of the cells and product quality attributes (reviewed by [24]). These engineering approaches have been enabled by the development of "omics" technologies and the accruing reference data for CHO

cells (reviewed by [25]) and by the availability of new genome editing tools for genetic engineering such as zinc finger nucleases, meganucleases, Talens, and the CRISPR-Cas9 system (review by [25, 26]). Tangible examples of these engineering approaches are the knockout of FUT-8 genes in CHO to enable the production of fucosyl antibodies conferring improved antibody-dependent cytotoxicity [27] and expression of single [28] or multiple heterologous genes [29] to improve glycan sialylation, which impacts glycoprotein half-life in vivo [28, 29].

As the CHO host cell is a current industry-standard platform, the remainder of this section will focus on CHO production systems.

There are in fact a plethora of different CHO host cell lines, as described by Wurm [30] and Lewis [31]. The original CHO cell line was derived from the ovary of an adult Chinese hamster by Puck [32] and later underwent cloning and other manipulations to derive a number of different CHO cell lines, including those commonly used for the production of therapeutic proteins. These include CHO-K1, CHO DG44, CHO-S, and CHO DUXB11, and these were originally cultured in medium containing serum. However, with the concerns about the cost, variation in the performance of different batches of serum, and the risks of contamination with adventitious viruses, these cell lines have been adapted to grow in animal-component-free media. The choice of individual CHO host is partly driven by the expression system used for recombinant protein production. The CHO DG44 cell line is deficient in dihydrofolate reductase (DHFR) and so is typically used with the DHFR expression system that complements this deficiency, as described below. The different CHO host cell lines can show differences in productivity phenotype, for example, recombinant cell lines from CHOK1 showed higher productivities for two difficult-to-express antibodies compared with cell lines created using DUXB11 [33]. However, the host cell lines are themselves heterogeneous, containing cells that show a diversity of characteristics, such as growth and ability to perform posttranslational modifications [34, 35]. These heterogeneous phenotypes result from genetic diversity, which can be observed as the varied karyology profiles of individual cells in a host population [30, 36] and from epigenetic variation. This genetic plasticity enables the adaptation of CHO cells to different culture media or culture conditions and contributes to the versatility of CHO cells for therapeutic protein production.

2.2 Gene Transfection

2.2.1 Stable Transfection and Selectable Markers

Typically, stable recombinant CHO cell lines are created by integrating linearized plasmid DNA encoding the therapeutic protein into the host genome so that, as the cell divides, the transgenes are transmitted to each daughter cell as the chromosomal DNA is replicated and segregated. Transfection methods to enable the transfer of the plasmid DNA to transfer across the cell membrane include

electroporation, calcium phosphate, cationic polymers, and lipid-based reagents. The process of integration of the plasmid DNA into the host genome is highly inefficient, and therefore, screening for recombinants is usually facilitated by including a selectable marker on the plasmid. This can be an antibiotic resistance gene, which can be selected by the addition of antibiotic to the cell culture medium or a metabolic gene such as DHFR or glutamine synthetase (GS), which can be selected by supplementing the cell culture medium with the specific enzyme inhibitors methotrexate and methionine sulfoximine, respectively (reviewed by [37]). Typically, transfectants that have incorporated the selectable marker and are resistant to the selection reagent also usually express the protein therapeutic as the genes are physically linked. Strong viral or housekeeping gene promoters are used to drive high levels of transcription of the therapeutic protein genes to maximize expression [38].

A number of strategies to improve the efficiency of the generation of high-producing transfectants have been developed and have been reviewed elsewhere [39]. These include the use of higher-stringency selection by using CHO host cells that have a knockout in the GS or DHFR endogenous metabolic marker that is used for selection [40] and/or attenuation of the selectable marker carried by the expression plasmid [41]. Other approaches incorporate chromosomal architectural sequences, such as ubiquitous chromatin opening elements (UCOE) or matrix attachment regions (MARS) into the expression plasmid to help promote expression of the therapeutic transgenes by enhancing or maintaining transcription once the plasmid is integrated into the host genome [38, 42]. Transposon-based expression systems can increase the frequency of high-yielding cell lines by increasing gene copy number and insertion at transcriptionally favorable sites [43]. Alternatively, expression plasmid DNA can be targeted by site-specific recombination to particular regions in the CHO genome that are known to be favorable for expression [44]. Similarly, expression plasmids can be targeted by site-specific recombination to regions of artificial chromosomes that have been developed to be maintained in CHO cells [45].

2.2.2 Transient Expression

Stable cell line development is time-consuming and resource intensive, and, at early stages of development, more-rapid methods with higher throughput, based on transient expression, are generally used to make therapeutic protein candidates for early characterization studies. Historically, transient expression systems based on the human HEK-293 cell line have been used for transient expression owing to its propensity for high transfection efficiencies and correspondingly high yields of recombinant protein (reviewed by [46]). Although transient expression in HEK-293 cells can be indicative of the expression levels of recombinant proteins seen in stable CHO cells [47], there are differences in product quality such as the glycosylation profiles [48, 49]. The desire to produce early-stage material that is more representative of the final production cell line has driven the development of

transient CHO systems that are capable of high yields. A number of CHO-based transient systems have now been developed involving engineering of the host cell line [50–53] and/or optimizing the transfection and production processes (reviewed by [54]). In these CHO transient systems, plasmid DNA is introduced into cells using either electroporation [55] or a range of reagents including polyethylenimine (PEI), calcium phosphate, or lipid-based systems [54, 56, 57]. The DNA that reaches the nucleus is transcribed and the mRNA processed and then transported to the cytoplasm where it is translated. No selection is applied, and the cells are allowed to grow and express over a period of a few days to a few weeks, with levels of the plasmid per cell reducing over time as the cells divide. Although CHO transient systems were initially developed for rapid expression of multiple candidates, the technology is being advanced for production at scale. With the achievement of titers exceeding 2 g/L at the 6-L scale [51], it is now becoming feasible to rapidly produce material for pharmacology, formulation, and toxicology studies without having to establish a stable cell line.

2.3 Clone Selection and Single-Cell Progeny

Stable transfection generates heterogeneous cell lines in terms of their productivity, growth, bioreactor performance, and product quality. This heterogeneity between individual transfectants arises from two sources. First is from the integration of the expression vector into different sites in the host genome and the variation in the number of copies of the expression vector, which impact the transcription of the genes encoding the therapeutic protein. Second, as described above, individual cells in the host population have heterogeneous phenotypes that determine characteristics such as doubling time, maximal cell density, metabolism, and product quality [30, 35]. This variation means that it is important to screen recombinant cell lines to ensure that candidate production cell lines express therapeutic protein with suitable product characteristics and that the growth and productivity parameters are suitable for scale-up in a production bioreactor. Finally, to ensure consistency of product quality and process performance, it is important to ensure that any cell line is clonally derived from a single originator transfectant, so that the cells in the cell line population are genotypically and phenotypically homogeneous. The processes for single-cell cloning are described below.

2.3.1 Cloning Processes

The regulatory guidelines direct that recombinant production cell lines should be "cloned from a single progenitor cell" (ICH Q5D).There are a number of different strategies and technologies to derive clonal cell lines (reviewed by [58]). Limiting dilution methods, where low concentrations of cells in suspension are distributed into multi-well plates, can be applied such that a cell line is recovered from a single

cell in an individual well. Traditionally, a statistical analysis of data on the recovery of colonies in wells has been used to support clonality of the derived cell lines [59]. However, with the advent of high-content imaging systems, limiting dilution statistics can be supplemented with detailed images that resolve and identify the originating single cell in a well at the time of plating. The capillary-aided cell-cloning method uses capillaries to transfer individual cells in a small droplet of culture medium from a dilute cell suspension into the well of a plate that is then visually assessed to confirm the presence of a single cell before addition of further medium [60]. Another approach uses the ClonePix robot system and involves dispensing cell suspensions into semisolid medium, allowing single cells to grow into colonies and then using the automated imaging and picking capabilities of the robot to transfer single, well-separated colonies into the individual wells of a multi-well plate [61]. The use of suitable fluorescent detection reagents in the semisolid medium enables enrichment of the colonies picked for those secreting recombinant product. The fluorescence-activated cell sorter (FACS) can be used for sorting of a cell suspension and single-cell deposition into individual wells of a multi-well plate [62]. The sorting capability of the FACS instrument can be leveraged by using fluorescently labeled reagents to detect either the product or a surrogate that is present on the cell surface and then sorting on the basis of the fluorescence signal [63]. Microfluidics is a developing technology that can also be applied to single-cell cloning [64]. Cells can be captured in picodroplets of culture medium in an oil emulsion, the resulting droplets can then be imaged on microfluidic chips, and those containing a single cell can be sorted and subsequently dispensed into plates. Additionally, as the recombinant protein is secreted by the cell into the culture medium in the droplet, this offers the potential to couple isolation of single cells with performing assays on the picodroplet for secreted product to assess yield or product quality [65].

2.3.2 Screening Strategies

Central to the cell line development strategy is that cell lines are screened in an established "platform" bioreactor process using production medium and feed, so that cell lines are selected as "fit to process" and minimize the requirement to perform further upstream process development before scale-up and manufacture of product for the clinic. Following cloning, the individual cell lines are assessed to find those that have suitable growth and productivity characteristics. In order to manage the assessment of the large numbers of cell lines, this process involves a screening cascade with a series of cell line evaluation steps where the numbers of cell lines reduce at each stage (Fig. 1). The first step generally simply identifies those cell lines expressing the therapeutic product, usually by detecting or measuring product secreted into the culture medium. Those cell lines expressing product are advanced to the next evaluation step that involves evaluating cell lines in fed-batch culture to assess both growth and productivity. Traditionally, this was performed using shake flask cultures, but manual handling restricts the number of

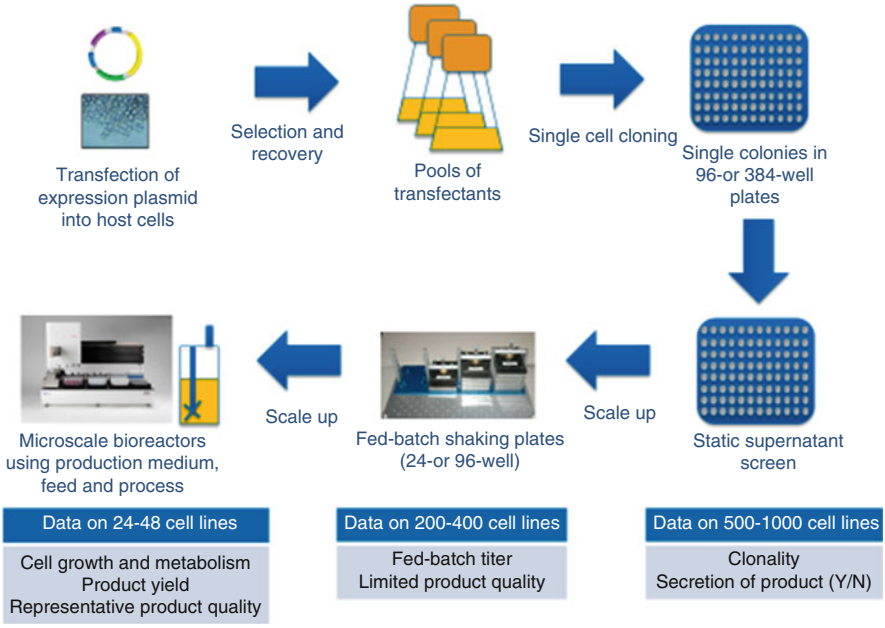


Fig. 1 A generic cell line development screening cascade. As the number of cell lines decreases at each screening step, more detailed data can be collected on each cell line in culture processes that are increasingly more predictive of the final production bioreactor production process. Example numbers of cell lines screened and the collected data at each stage are summarized in the *text boxes*

cell lines that can be evaluated in parallel to a few tens. However, the development of high-throughput, small-scale, fed-batch culture processes using multi-well plates allows hundreds of cell lines to be assessed in parallel [66]. Subsequently, further performance data can then be obtained on multiple (24–48) cell lines in parallel by using scaled-down bioreactor systems, which control pH and dissolved oxygen, generating data that are predictive of larger-scale bioreactors in terms of cell growth, productivity, and metabolism [67]. An additional output from the micro-scale reactor systems is the cell culture samples containing representative product for analysis of product quality attributes. The resulting cell line bioreactor performance and product quality data are assessed to select candidate production cell lines for further in-depth characterization before selecting the final clone for the creation of a master cell bank (MCB).

2.3.3 Product Characterization During Development of Stable Cell Lines

The cell line and the upstream process both contribute to determining the product quality attributes of protein therapeutics. Therefore, it is important to integrate product quality analysis into the cell line screening process. The generation of

analytical data is facilitated by high-throughput analysis of product within the cell culture medium [68] or by integration with high-throughput purification and analytical assays. The exact assays are determined by the properties of the product itself, but typically include evaluation of glycosylation, aggregation, fragmentation, and amino acid sequence integrity. Amino acid sequence integrity is assessed by mass spectrometry methods and peptide mapping [69] and confirms that the expected sequence is being produced. Product sequence variants containing one or more amino acid substitutions can occur as a result of mutations in the encoding DNA or misincorporation of amino acids during translation in the production cell line [69–72]. As these sequence variants are cell line specific, it is feasible to screen them out during clone selection if the causal genetic modification has occurred prior to cell cloning.

2.4 Phenotypic and Genotypic Stability

A critical requirement for the upstream production process is that it is robust and generates product of consistent quality across different batches and scales in order to satisfy both regulatory and commercial requirements. To achieve this, the production cell line must demonstrate suitable growth, productivity, and product quality over the generations required to run the manufacturing process (Fig. 2). Therefore, candidate production cell lines are typically evaluated for phenotypic and genotypic stability at different time points over approximately 70 generations before final clone selection to mitigate the business risk of selecting a cell line that is not sufficiently stable. Regulatory requirements for demonstrating cell line stability are described in the ICH Q5D guidelines.

A number of underlying genetic mechanisms of instability of product expression have been characterized, including gene loss, gene rearrangements, and gene silencing [73]. In addition, as described above, product-related variants owing to genome mutations can occur during cell line development and subsequent subculture. Initially, cells with these genetic changes can be at a very low level within a cell population, so that they are difficult to detect by standard product or genetic characterization methods. However, if these genetic changes confer a benefit to the cell in terms of reduced biosynthetic and metabolic burden, this can lead to faster cell growth, and in turn these aberrant cells can begin to increase in proportion in relation to other cells in the population and consequently impact product expression levels and/or product quality.

Regulatory approval requires that the production cell line is genetically characterized in accordance with the regulatory guidelines (ICH Q5B). This includes analysis of the integrated expression construct for copy number, typically by QPCR, and assessment of insertions or deletions of the transgenes and the number of plasmid integration sites, generally using Southern blotting. In addition to protein analytical characterization, confirmation of the protein sequence for the therapeutic protein is performed by sequencing genomic copies of the encoding

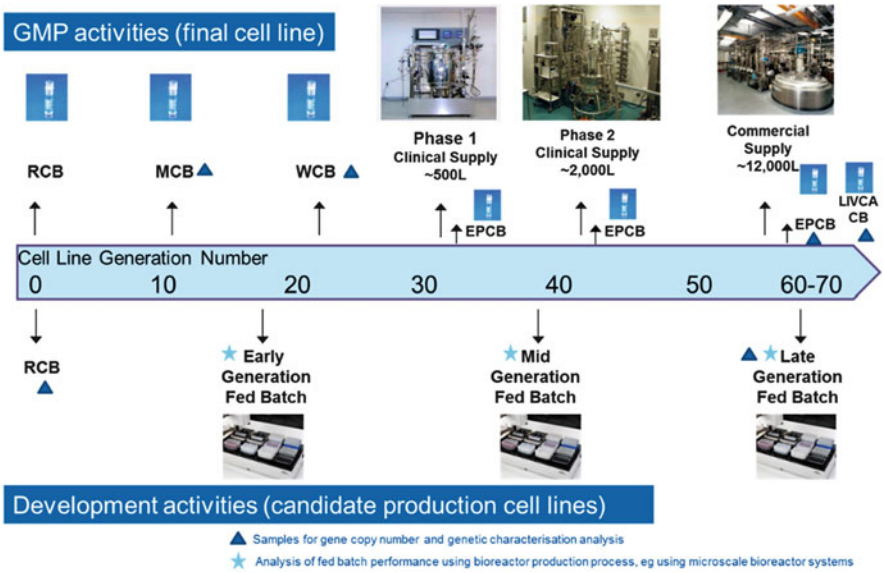


Fig. 2 Phenotypic and genotypic stability assessment in relation to manufacturing time scales. *Top panel*: seventy generations exceed the number of cell doublings required to expand a research cell bank (RCB; generation 0) to create master and working cell banks (MCB and WCB) and then to run the production bioreactors up to commercial supply scale (12,000 L). Cell banks are generated from cells from the production bioreactors – the end of production cell bank (EPCB) and also the limit of in vitro cell age cell bank (LIVCACB) that defines the upper limit of the number of cell generations used for the manufacturing bioreactor process. *Bottom panel*: candidate production cell lines are cryopreserved as RCBs and then *characterized* over a further 70 generations to demonstrate a suitably stable phenotype and genotype in cell culture processes that simulate the production process

genes or by sequencing cDNA produced from RNA transcripts. Next-generation sequencing is also beginning to be applied as a characterization tool for assessing copy number, structure, and sequences of therapeutic protein genes as well as transcript sequence [74].

2.5 Cell Banking

Production cell lines are cryopreserved to enable long-term storage and to sustain consistent manufacture of a therapeutic protein through the clinical development and commercial phases of the drug life cycle. Methods and strategies for the cryopreservation of production cell lines have been reviewed elsewhere [75]. Typically, a tiered system of frozen cell banks is used as outlined in the ICH Q5D guidelines. At the end of cell line development, candidate production cell lines are cryopreserved as a research cell bank (RCB) or pre-master cell bank

(pre-MCB).A single vial of the RCB of the final clone is then used to create an MCB containing several hundred identical vials. Cells from a high-viability culture are centrifuged, and the cell pellet is resuspended in culture medium containing a cryopreservant, typically 5–15% dimethyl sulfoxide, to create a homogeneous suspension that is then dispensed into individual vials before being frozen and stored in the vapor phase of liquid nitrogen at temperatures below –135°C. The process of preparing the vials is performed aseptically under controlled conditions to minimize the introduction of any environmental contaminants. The MCB is the starting point for manufacture of a therapeutic protein product, and so it is vital to safeguard the MCB to maintain drug supply. Hence, MCB vials are stored in at least two independent liquid nitrogen Dewars to mitigate risk of a catastrophic Dewar failure and to maintain supply of the therapeutic product. Also, in order to conserve the use of MCB vials, working cell banks (WCBs) are cryopreserved in a similar way to the MCB, each WCB consisting of a several hundred vials, being derived from a single vial of the MCB. A single vial of the WCB is then revived and expanded to provide culture to inoculate a production reactor. At early stages of clinical development or if only a few vials of the MCB are required each year for product supply, MCB vials can be used for manufacture, avoiding the need to generate a WCB. Historically, cryopreserved MCBs are very stable if stored appropriately, with cells reviving in a consistent way even after many years. The recovery data of the cell banks is tracked over time so that, if necessary, a new MCB can be generated to maintain drug supply.

Typically, cell banks are cryopreserved in small (1–2 mL) cryovials, and cells from each vial are revived into 50–100 mL of cell culture medium and then expanded into larger culture volumes over sequential passages to generate sufficient cells to inoculate the production reactor. At the commercial supply stage, there is a desire to shorten the timelines, and the duration of this expansion phase can be shortened by reviving a larger number of cryopreserved cells at the start of the process. This can be achieved by cryopreserving cells in larger cryopreservation bags (150 mL). In turn, this process requires the availability of high-density and high-viability cell cultures, which can be produced from a bioreactor process in a fed-batch or perfusion mode [75].

Cells used in the manufacturing process are also cryopreserved for the purposes of safety testing, which is described in the next section. Cell culture is removed from the production bioreactor to create the cell banks. If the cells are at high viability, the cells can be cryopreserved directly, whereas lower viability cells are serially subcultured until they reach high viability before being cryopreserved. Freezing high-viability cells ensures that the cells can later be successfully revived. These banks are termed end of production cell banks (EPCB) and limit of in vitro cell age cell banks (LIVCACB) when the cells have reached the maximal number of cell generations that will be used for the manufacturing bioreactor process.

2.6 Safety Testing

Safety testing of the cell banks and the culture harvest at the end of bioreactor production process is performed to ensure patient safety, as outlined in ICH guidelines Q5A and Q5D. Identity testing of the cell banks is performed to confirm the species of origin of the cell line and is important as the cell lines can become contaminated with other cell lines if not appropriately segregated [76]. Tests for microbial and viral contamination are performed on the cell banks and the bioreactor harvest or unprocessed bulk (UPB) according to ICH guidelines Q5A and using methods summarized by Mclean and Harbour [77]. The testing of the UPB, end of production, and limit of cell age banks (EPCB and LIVCACB; Fig. 2) confirms that microbes or adventitious viruses were not introduced during the production process, for example, through the use of contaminated raw materials in the culture process [78] or from the environment. In addition, the risk of introducing contamination with adventitious viruses is mitigated by the use of animal-component-free culture medium, screening raw materials, and heat treatment of culture medium to inactivate viruses. The development of next-generation nucleic acid sequencing technologies provides an additional method for testing and investigating potential incidences of contamination [79]. As next-generation sequencing technology gains regulatory acceptance, it has the potential to reduce the need for the in vivo testing that forms part of the traditional program of virus testing. It is well documented that CHO cell lines possess type A and type C endogenous retrovirus that express retroviral particles [80]. However, these particles have been found to be defective and noninfectious for cells from other animals, including human, and are not considered to present a safety risk [81].

3 Cell Culture Process

3.1 Inoculum Train

Manufacturing of a protein therapeutic starts with the thaw of a vial of a frozen cell bank. Cells are diluted in nutrient medium and placed in shake flasks. Cultures are progressively scaled up in larger shake flasks, by passaging to the next stage when they have grown to a predefined cell density range. Cells are then transferred to disposable bags (20 or 50 L) and subsequently to larger inoculum bioreactors, often known as seed bioreactors. For the earlier stages in shake flasks and disposable bags where pH is not actively controlled, CO_2-enriched air is used to buffer the growth media. For seed bioreactors, pH is controlled within a range by way of sparging CO_2 gas (acid) or dilute sodium hydroxide (base) as needed. Temperature is controlled throughout the inoculum train, and for seed bioreactors, dissolved oxygen (pO_2 measured in mmHg) is also controlled at a predefined set point. The

primary objective of inoculum train is to generate sufficient quantity of cells at high viability to inoculate the production bioreactor where protein production occurs.

3.2 Production Bioreactor

Nutrient and environmental conditions within the production bioreactor are optimized to ensure growth and maintenance of cells at high viability to enable maximum expression of protein of acceptable quality. Environmental conditions that need to be controlled include temperature, pH, oxygen (pO_2), carbon dioxide (pCO_2), and hydrodynamic shear. Agitation and aeration in the bioreactor are controlled such that the oxygen demand of cells is met and dissolved CO_2 is maintained within a predefined value. Agitation also ensures uniformity in concentration of cells and nutrients inside the vessel. Since many product attributes including glycosylation are primarily determined at the cell culture stage, control of bioreactor process parameters is very important. Also, culture viability before harvest is maintained sufficiently high so that the cell harvest process runs smoothly and product variants that may be formed due to released enzymes (e.g., sialidases) [82] are minimized. Low viability also results in higher levels of host cell DNA and proteins, which makes their clearance through the purification steps more challenging.

While the scale of production depends on the product demand, production stage bioreactors as large as 25,000 L are being used in the industry for suspension cell culture [83]. Production bioreactors are often stainless steel reactors although disposable reactors up to 2,500 L are also available [84] and are being increasingly used owing to their ease of implementation and routine operation, reduced change over time between batches, and flexibility in plant capacity as demand changes.

3.3 Modes of Operation

The most typical mode of operation, termed fed-batch mode, involves adding concentrated nutrient-rich feeds periodically during the course of the bioreactor run to supply nutrients such as glucose and amino acid consumed during growth and product expression. This mode is preferred for its balance of simplicity and ability to satisfy nutrient demands to very high cell densities and hence is the mode most used in the industry. It should be noted, however, that this configuration does not work for all molecular formats. If the protein molecule is unstable at the temperature and pH conditions of the bioreactor or because of enzymatic action, it becomes necessary to minimize the residence time of the product in the bioreactor, making a perfusion mode of operation preferable. For example, commercial production of recombinant human Factor VIII is performed using a perfusion system [85]. Also, if the production cell line is anchorage dependent, cells are often grown

on microcarriers. A seasonal influenza vaccine is produced in Vero cells that grow on microcarriers [86]. In addition to providing anchors for cell growth, microcarriers also protect cells from excessive shear.

An emerging trend in the industry is the interest in continuous bioprocessing [2]. The continuous system makes an attractive option to increase the throughput of the plant allowing manufacture of large amounts of product in a facility with smaller-scale bioreactors. Many cell retention devices can be used for perfusion process [87], but alternating tangential filtration (ATF) devices are particularly popular [88]. Also, a perfusion process may be used to increase the cell density in the last inoculum reactor, which in turn results in a higher seeding cell density in the production bioreactor. It should be noted that perfusion, which involves continuous addition of medium and removal of product with return of cells back to the bioreactor, does increase the volume of medium consumed as well as the complexity of operation.

3.4 Process and Media Optimization

During the design of the process, the nutrient concentrations in the basal medium as well as in the nutrient feeds are carefully optimized considering the shifting metabolic needs of the cells during the growth stage and during protein expression. Due to risks associated with transmissible spongiform encephalopathy (TSE) and other transmissible diseases of animal origin, industry has shifted away from the use of animal sera or media containing serum components. However, some serum or serum components are occasionally used during cell line development. Although some companies use only chemically defined media, many rely on animal protein-free formulations that do not have any components of animal origin but may contain complex raw materials such as hydrolysates. Industrial cell culture media contain a number of components and commonly include glucose, amino acids, vitamins, salts, trace metals, buffering agents, and antioxidants. Components such as growth hormones may also be included. Additionally, growth media contain shear protectants in the form of surface active polymers such as pluronic polyols [89]. Media optimization is usually achieved via rational understanding of cellular needs as well as through empirical experimental studies. An important criterion for medium design is to stay within a narrow range of osmolality (approximately 300–500 mOsm) where mammalian cells thrive. Equally important is the design of nutrient feeds, which often have higher nutrient concentrations compared to those in growth media. Since the growth media and nutrient feeds often contain multiple components that interact with each other, scientists rely on statistical experiments to improve process productivity and optimize product quality. Such experiments are carried out in benchtop bioreactors (2–5-L working volume) or in miniaturized bioreactors with working volumes as low as tens of milliliters [90]. Industry has been quite successful in continually increasing the titers for mammalian cell culture processes. A recent example of a significant titer increase,

accomplished by optimizing nutrient concentrations and process conditions without the need for a cell line change, was reported by Ahuja [91].

In addition to approaches mentioned above, advances on multiple fronts continue to enhance the understanding of mammalian cell metabolism. These advances relate to different omics technologies [92, 93] and flux analysis techniques [94]. These techniques help scientists generate hypotheses for further developing growth media and nutrient feeds.

3.5 Equipment Design and Maintaining Sterility

The primary goals of the mechanical design of the bioreactor are to prevent ingress of microbial and adventitious viral agents, to supply oxygen needed for cell growth and product expression, and to provide adequate mixing to ensure homogeneity. Sterility is achieved by ensuring that the bioreactors provide a closed environment with the sterile boundary and that the media, feed, and base used in the manufacturing process are sterilized through membrane filters as they enter the bioreactor. Bioreactors are commonly run at positive head space pressure (generally $\leq$5 psig) to minimize the risk of bacterial contamination. Requirement of the sterile boundary necessitates the design of a robust sterile-in-place (SIP) method and other operational procedures for bioreactor operation and maintenance, the details of which can be found elsewhere [95]. Sterilization cycles for bioreactors and autoclaves are designed to provide sterility assurance level (SAL) in excess of 10^{-6}, where SAL of 10^{-6} indicates the probability of one out of 10^{6} sterilization cycles to be unsuccessful [96].

Adventitious agents of concern include bacteria, yeast, mycoplasmas, and viruses. Use of 0.1-μm filtration is generally effective in removing bacteria, yeast, and mycoplasma. For higher assurance of inactivation of mycoplasmas and viruses, heat treatment (high-temperature short time (HTST)) could be employed. HTST treatment of media and feeds involves exposure to temperature about 102°C for approximately 10 s [97]. This treatment has been successfully shown to inactivate multiple viruses including minute virus of mice (MVM), which has been implicated in facility contaminations as recent as 2011 [98]. In addition to heat treatment, another possible technology for inactivating viruses is the use of UV-C (ultraviolet light in the C spectrum) irradiation [99], which so far has found limited use in the industry. Yet another approach is gamma irradiation, which is frequently used for inactivating viruses in sera used for biologics manufacturing [98]. Gamma irradiation is also used to sterilize disposable bags and bioreactors [100].

The bioreactor usually has an aspect ratio (diameter to height ratio) of approximately 1:1.5, which is higher compared to that of a microbial fermenter (approximately 1:3) [101]. A shorter reactor would be better for maintaining homogeneity even though it might be less efficient in oxygen transfer for a given amount of power input [102]. Bioreactors are continuously mixed to ensure homogeneity of cells, nutrients, and environmental parameters such as pH, temperature, dissolved

oxygen, and pCO_2, as well as to improve gas–liquid mass transfer by reducing the size of the gas bubble as well as that of the liquid boundary layer. Many impeller types including Rushton, hydrofoil, elephant ear, marine, and pitched blade can be seen in the industrial bioreactors [103, 104]. Compared to Rushton impellers, more efficient impellers such as axial hydrofoil impellers make an attractive choice, as the latter can be operated with lower power input to provide the desired level of mixing and mass transfer, and they also provide better axial (vertical) mixing [105]. In large-scale bioreactors, it is common to have two impellers with the bottom impeller located above the sparger, which supplies gases (air, oxygen, and CO_2) to the bioreactor. Bioreactors may also contain baffles to aid adequate mixing. Also, if possible, nutrient feeds and base used for pH control are added through subsurface addition; this has been shown to significantly reduce pH fluctuations during base addition in the bioreactors [106].

Foaming, the extent of which depends upon the medium type and sparge conditions, is usually controlled using products containing simethicone. Foaming is not desirable as it can act as a trap for cells, which in turn can be damaged by bubbles bursting on the liquid surface [107]. Moreover, excessive foaming can compromise bioreactor sterility if the foam wets the exhaust filter.

3.6 Process Scale-Up

One must be aware of different physical and chemical environments that may exist as the process is scaled up from a small lab scale to a larger production-scale bioreactor, as maintaining similar environmental conditions is the key to successful scale-up. Parameters that are routinely controlled in a bioreactor are temperature, pH, agitation, and dissolved oxygen (pO_2). Temperature and pH set points remain the same as the process is scaled up. The pO_2 set point may require changing to account for the increased pressure in a larger bioreactor in order to maintain the oxygen concentration across scales.

Agitation rate changes with scale and is commonly estimated using the scale-up criterion based on same power input per unit bioreactor volume (P/V). P/V is a common scale-up criterion since it is a broad indicator of mixing and mass transfer characteristics. Mixing time is inversely related to agitation rate up to a point, and it generally increases with scale and is usually in the order of minutes in large-scale (10,000–15,000 L) reactors [108]. As long as they are significantly lower than characteristic times of gas–liquid oxygen transfer and cellular oxygen uptake [104], differences in mixing times across scales do not pose a problem as they have been found to be satisfactory for dispersing nutrient feeds and base, which are intermittently added in the process. Shear stress experienced by cells in a bioreactor can be estimated by the knowledge of agitation speed [109] or more precisely, by using computational fluid dynamics [110]. Excessive shear, which may generate turbulent eddies (known as Kolmogoroff eddies) with sizes similar to that of cells, can damage cells directly. However, it has been shown that the lethal level of shear

needed to directly damage cells is significantly higher than what is normally experienced in the bioreactors [111]. This offers flexibility in scale-up as agitation set point calculated based on P/V can be adjusted as necessary to optimize other parameters.

Compared to agitation, a more critical factor is the effect of gas sparging on bioreactor performance. Sparge rate and/or composition (relative proportion of air and oxygen) is continuously varied to compensate for the change in the rate of oxygen consumption during the course of the culture in order to maintain the pO_2 within a narrow range. It has been well established that major damage to cells in a bioreactor occurs when bubbles burst at the top of the surface [107]. High rate of sparging, in addition to damaging cells, can also cause excessive foaming. The other important effect of gas sparge rate and composition is the level of CO_2 in culture. Cellular metabolism results in evolution of CO_2 and has to be removed at the same rate to prevent accumulation, as high levels of pCO_2 have been shown to negatively affect the product yield and product quality [112]. High levels of pCO_2 in the bioreactor reduce the pH, which in turn results in higher base addition. This increases the osmolality, which may negatively affect the culture performance. Higher levels of pCO_2 can also affect the intracellular pH, which can result in altered process performance. Appropriate control of pCO_2 in the bioreactor is therefore critical to successful scale-up.

The gases (air and oxygen) sparged to meet cellular oxygen demand in a bioreactor also serve to remove CO_2. In large-scale bioreactors, removal of CO_2 is more challenging than in smaller vessels. As indicated earlier, it is typical for process engineers to maintain a similar P/V as the criterion for scale-up of agitation. This approach ensures adequate mixing, comparable shear imparted to the cells by the impeller, and comparable mass transfer efficiencies to satisfy cellular oxygen demand. However, because the larger-scale vessel is considerably taller, the residence time of gas bubbles is significantly longer, and they reach near-saturation levels of CO_2 concentration [113], thereby losing their ability to strip out CO_2. This in turn leads to an accumulation of CO_2 in the bioreactor. To ensure similar levels of CO_2 between scales, careful consideration of equipment design and sparging strategy is important. Equipment design could include impeller design, sparger design, number of spargers, and their placement. Design of sparging strategy, that is, the profile of sparge rate and composition throughout the culture, should consider cellular CO_2 metabolism/stripping in addition to oxygen metabolism/supply.

While designing the agitation rate and sparging strategy, the process engineer must take into consideration the constraints of mixing, shear stress, damage due to sparging, oxygen supply, CO_2 accumulation, and foaming. For example, although a high gas flow rate would be preferred for keeping pCO_2 levels low, it might result in excessive foaming. Figure 3 depicts the interplay between different constraints and how they all must be looked into for optimizing operating conditions in a bioreactor.

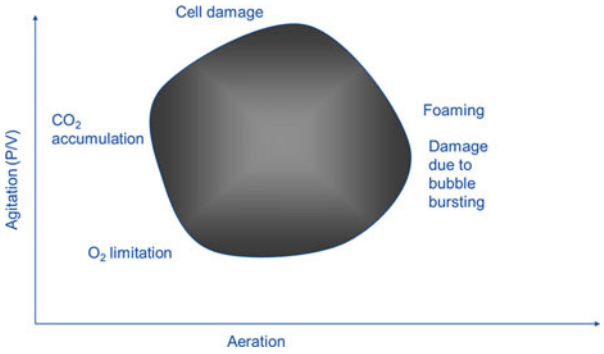


Fig. 3 Interplay of different constraints as function of agitation and aeration rates

The concerns listed above apply to fed-batch as well perfusion cultures with the latter having increased demands for oxygen transfer, CO_2 removal, and foaming due to much higher cell densities.

4 Harvest

Cell harvesting is the separation or retention of cells and cellular debris from the supernatant containing the target molecule. Selection of a harvest method depends greatly on the type of cells, mode of bioreactor operation, process scale, and characteristics of the product and cell culture fluid. Whether the bioreactor is operated in continuous (perfusion), batch, or fed-batch mode, the optimal harvest method preserves cell viability in order to avoid release of intracellular enzymes and impurities that could negatively impact product quality and complicate the purification process. Harvest techniques used for perfusion cell culture must also be designed to maintain sterility of the bioreactor throughout the duration of the perfusion operation, which may be up to several weeks. Further details on cell retention techniques for perfusion can be found in the review by Voisard et al. [114]. Most traditional harvest methods use some form of filtration, centrifugation, or a combination of both for cell separation and/or retention. Filtration methods include normal flow depth filtration and tangential flow microfiltration and achieve separation based on particle size differences. Centrifugation methods achieve primary separation based on density differences and are typically followed by depth and sterile filtration for removal of small debris from the centrate. The ability to predictably scale the selected harvest method helps ensure successful production and is critical for conducting small-scale characterization studies for confirming parameter targets and ranges.

4.1 Methods for Cell Harvest

Batch harvest methods include tangential flow filtration (TFF), depth filtration, and centrifugation. Tangential flow filtration has traditionally been used for lower solid containing <2% and higher viability cell cultures.TFF retains cells using a membrane while the clarified filtrate flows through. The flow along the membrane surface helps keep the membrane from fouling. This method is linearly scalable and, depending on the membrane pore size, can generate clarified cell culture fluid with minimal downstream filtration. Since cells are concentrated during harvest, this method is limited by the concentration factor and product yield that can be achieved. While single-use options for TFF are available, reuse is generally economically more favorable.

Depth filters separate cells based on size exclusion by forcing the liquid through a porous medium in normal flow mode. The medium traps cells and cellular debris while the liquid flows through to a collection vessel. Depth filters of different nominal pore sizes are often used in series to allow the larger cells to be trapped in the first stage and cellular debris to be trapped in the second stage. As harvest volumes increase, the amount of depth filtration area also increases, leading to a larger footprint and less economic advantages when compared to TFF and centrifugation.

Centrifugation is the most widely implemented harvest method for large-scale manufacturing. Centrifuges separate cells from the product-containing liquid by exploiting density differences between the solid cells and the liquid. In a continuous disk stack centrifuge, the feed is accelerated via a rotating bowl, and the solids are thrown toward the outer radius where they contact a series of closely placed, angled disks. The solids travel down the sides of the disks, collect at the bowl periphery, and are discharged intermittently by opening and closing the bowl. The clarified liquid travels upward and out of the centrifuge. The clarified liquid typically is filtered downstream through a combination of depth and membrane filters to ensure complete solids removal. Centrifugation provides significant advantages over filtration including high process yields, effective and consistent clarification performance, and lower resources for process development. However, among the three most commonly used harvest methods, centrifugation is operationally more complex and more challenging to scale up.

4.2 Scale-Up of Centrifugation

Centrifugation takes advantage of density differences between cellular solid particles and cell culture fluid in order to achieve separation. Particles that differ in density will settle at different rates in response to an applied gravitational or

centrifugal force. Assuming laminar flow and approximating cellular particles as spheres enable application of Stoke's law, which defines particle settling velocity as:

$$V_{\mathrm{t}} = \frac{\omega^2 r(\rho_p - \rho l)d^2}{18\mu}, \tag{1}$$

where V_{t} is the settling velocity, ω^2 is the angular velocity, r is the distance of the axis of rotation, ρ_{p} is the density of the particle, ρl is the density of the liquid, and μ is the viscosity of the liquid.

Ambler related the particle settling velocity to the flow rate and settling area of a centrifuge using:

$$V_{\mathrm{t}} = \frac{Q}{\Sigma}, \tag{2}$$

where Q is fluid flow rate and Σ is a relationship Ambler derived for various types of centrifuges [115]. For a disk stack centrifuge, Ambler defined Σ as:

$$\Sigma = \frac{2\pi N\omega^2}{3g}\left(r_2^3 - r_1^3\right)\cot\alpha, \tag{3}$$

where N is the number of disks, ω^2 is the angular velocity, r^2 is the maximum disk radius, r^1 is the minimum disk radius, and α is the half cone angle of the disk.

With these relationships, the development scientist can experimentally determine the settling velocity of the cellular solid particles using bench-scale centrifuges. Since sigma factors are known for each centrifuge, the scientist can solve for feed flow rate in Ambler's equation. In theory, operation of the centrifuge at the determined flow rate should yield a centrate with similar clarity as that found in the bench-scale centrifuge experiment. However, in reality, the centrifuge system is more complicated and additional experimentation is needed to fully predict separation performance. Prediction of performance using the sigma factor relationship is more reliable when comparing Q/Σ from similarly-designed centrifuges. For example, a laboratory-scale disk stack centrifuge may be used to predict the flow rate range needed to achieve the same separation as a production-scale disk stack centrifuge via:

$$\left(\frac{Q}{\Sigma}\right)_{\mathrm{lab}} = \left(\frac{Q}{\Sigma}\right)_{\mathrm{production}}. \tag{4}$$

Even when using a scaled-down version of the production centrifuge, scientists often need to apply a safety factor to account for the effects of differing particle size distributions, shear-induced cell lysis, higher discharge pressures, and variable solid volumes from batch to batch.

Several researchers have successfully developed small-scale models to mimic the shear created within the feed zone of a large-scale centrifuge [116–118]. When used with a laboratory-scale centrifuge, the shear models have had success predicting centrate clarity from a pilot-scale centrifuge [117].

4.3 Filter Sizing

Filter sizing for harvest operations is best determined empirically with scaled-down devices and representative feedstocks. In theory, Darcy's relationship between flow rate and the pressure drop resulting from the flow through a porous material can be used to describe filtration performance:

$$\frac{dV}{dt} = \frac{k\Delta PA}{\mu L} \tag{5}$$

where dV/dt is flow rate, k is bed permeability, μ is liquid viscosity, ΔP is pressure drop, L is bed thickness, and A is filtration area.

However, applying Darcy's law to biological filtrations is challenging because of the difficulty in calculating the bed permeability and the resistance (k/L) for these fluids. Though permeability and resistance could be estimated from the Kozeny–Carman expression [119], experimental measurements are usually a better option. Experiments can be performed either at constant pressure ($V_{\max}$) or at constant flux ($P_{\max}$) [120]. The $V_{\max}$ method applies the linear form of the pore-plugging model to predict the maximum volume that can be filtered:

$$\frac{t}{V} = \frac{1}{Q_{\mathrm{i}}} + \frac{t}{V_{\max}}, \tag{6}$$

where t is time, V is volume, Q_{i} is initial filtrate flow rate, and $V_{\max}$ is the maximum volume that can be filtered at the test pressure before the membrane fouls.

The experiment is performed by filtering at a constant pressure and measuring the filtered volume as a function of time. If the plot of t/v versus t is linear, then the gradual pore-plugging model applies and $V_{\max}$ is calculated from the inverse of the slope. If the plot of t/v versus t is nonlinear, then the constant flux method should be used. Generally, constant pressure experiments are best for screening depth filtration options and constant flux experiments are best for estimating filtration performance. Unless numerous studies are performed with multiple feedstocks, the experimentally determined filtration area is increased by 50–100% to provide a margin of safety for large-scale depth filtration harvest operations [121]. Depth filtration sizing experiments are relatively easy to perform, and generally similar harvest areas and filter types can be used across pilot and commercial scales.

Harvest operations for cell culture are typically filtration or centrifugation or a combination of both methods. Harvest performance is still heavily determined by

experimental data, although improved scale-down models and high-throughput systems continue to be developed.

5 Purification

5.1 Purification Process Targets

Purification of biopharmaceuticals must strike a balance between several competing goals to achieve an optimal result. While each drug is unique, these competing goals often encompass product quality, manufacturing facility fit, adherence to platform process principles, robustness, scalability, productivity, yield, cost of goods for manufacturing (COGM), employee safety, environmental impact, freedom to operate (intellectual property), and validation. While it is beyond the scope of this work to discuss all of these considerations in detail, a few definitions and general guidelines are provided below.

Product quality is defined as the suitability of either a drug substance or drug product for its intended use. This term includes such attributes as identity, strength, and purity (ICH Q8(R2)).Typical limits for process-related impurities and bacterial endotoxins in biopharmaceuticals are provided in Table 1. There are additional considerations for product-related substances and impurities that are specific to each biopharmaceutical. These species may need to be controlled and, therefore, must also be taken into account during design of a purification process.

Good manufacturing facility fit can be thought of as the ability to run a process optimally with few if any changes to the existing facility or equipment. Common purification process bottlenecks, such as buffer requirements that exceed storage capacity, are avoided. For monoclonal antibodies in particular, adherence to platform process principles, where most aspects of process design are fixed, can usually avoid many facility fit issues.

Process robustness is defined as the ability of a process to tolerate variability of materials, process parameters, and equipment performance without negative impact

Table 1 Typical limits for process-related impurities in biopharmaceuticals

Impurity	Limit	References
Host cell protein (HCP)	Varies based on risk assessment; often ng HCP/mg drug	[122], author experience
Small and synthetic macromolecule impurities	Varies based on risk assessment; often μg impurity/mg drug	[123], author experience
DNA	≤10 ng/dose	[124]
Endotoxin	≤5 EU/kg body weight per h[a]	USP <85>[b]

[a]For parenteral administration other than intrathecal

[b]The United States Pharmacopeia Convention <85> Bacterial Endotoxins Test. www.usp.org/usp-nf/harmonization/stage-6/bacterial-endotoxins-test

on quality (ICH Q8(R2)).As process deviations and upsets are possible occurrences during routine manufacturing, good process robustness is a critical goal of purification process development.

Process productivity (e.g., kg/year) and yield (e.g., kg/batch) are related but not identical. These parameters determine how much drug a given facility can produce per unit time and at what cost. Cost of goods for manufacturing (COGM) is impacted by productivity and yield, as well as other parameters such as raw material costs and labor costs. Process economics is discussed in greater detail later in this chapter.

5.2 *Purification of Monoclonal Antibodies*

Given the importance of the biopharmaceutical market and scale of production as indicated earlier in this chapter, tremendous effort has been devoted to mAb purification process development. Numerous comprehensive reference texts on the subject are available (e.g., [125]). Here, we will briefly describe the culmination of decades of biopharmaceutical development, resulting in the modern mAb platform purification process.

Figure 4 shows a process flow diagram for a common variant of the mAb platform purification process. This process begins with removal of cells and clarification by continuous centrifugation and depth filtration. The clarified culture broth containing the mAb is captured by Protein A affinity chromatography. This is followed by a dedicated viral clearance step – low-pH inactivation. The anion exchange flow through chromatography step is principally a viral clearance step, although it can also clear process-related impurities such as DNA. The cation exchange step serves to remove process and product-related impurities. Much of

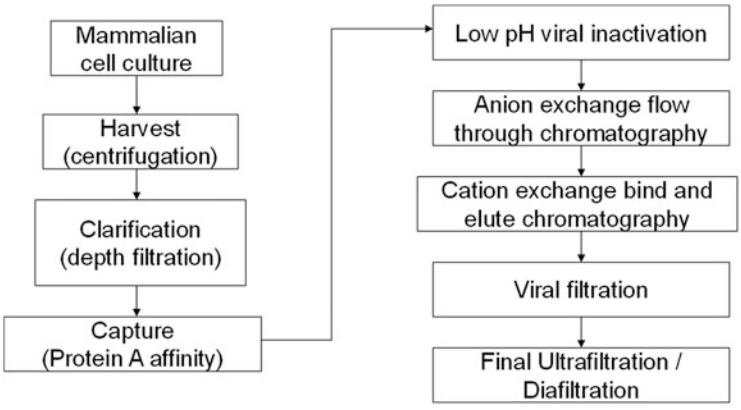


Fig. 4 Process flow diagram for a common variant of the modern mAb platform purification process

the process development effort is dedicated to optimization of this step for removal of product aggregates. Another dedicated viral clearance step, virus filtration, is performed after cation exchange. Lastly, the mAb is concentrated and buffer exchanged into formulation buffer by ultrafiltration/diafiltration.

In most instances, the Protein A column is placed first in the purification train to capture product from clarified cell culture broth [125–127]. This configuration provides an optimal balance of process similarity for different molecules, favorable COGM for routine commercial manufacturing, and process robustness. However, process models developed to predict cost of goods and facility capacity tend to be sensitive to Protein A capture column dynamic binding capacity (DBC) [128, 129]. As a result, the last decade has seen introduction of multiple generations of Protein A stationary phases designed to achieve higher DBCs.

5.3 Affinity Chromatography

Owing to the demand for mAbs, the industrial-scale use of Protein A chromatography with column diameters in excess of one meter is now routine. Staphylococcal Protein A (SpA) is a 42 kDa single-chain protein localized to the outer surface of *Staphylococcus aureus* [130–133]. Native SpA is composed of five Fc-binding domain-designated E, D, A, B, and C [134–136].

The potential of SpA to be used as an affinity ligand for protein purification was recognized early on. Initial SpA affinity resins consisted of native Protein A coupled to a base matrix usually through covalent bonding to amines. Since then, extraordinary improvements have been made in Protein A chromatography resins. Among the most innovative is the Z domain, which represents an analogue of the native B domain developed for purification of Fc-fusion proteins [137, 138]. A derivative of the Z domain engineered for greater alkaline stability, incorporated in mAb Select Sure from GE Healthcare, has gained widespread use for capture of mAbs and Fc-fusion proteins.

Protein A chromatography is most often used for capture of biopharmaceuticals. It serves to remove process-related impurities, including cell culture media components, host cell protein, and DNA. Fc-containing proteins are typically bound to Protein A at neutral or near neutral pH and eluted under acidic conditions below pH 4.0.

Other affinity matrices are used in mammalian cell culture processes. Notably, recombinant Factor VIII preparations have historically utilized custom-designed immunoaffinity resins [139]. However, as non-mAb recombinant proteins manufactured in 2013 represent less than 2% of the total mammalian cell culture production, these applications comprise a tiny fraction of the bioprocessing capacity represented by Protein A chromatography [1].

5.4 Ion Exchange Chromatography

Due to its gentle processing conditions and relatively low cost, ion exchange is among the most widely used separation modalities for purification of biopharmaceuticals. Most modern ion-exchange media are composed of a porous agarose, methacrylate, or polystyrene base matrix containing either strong or weak ionogenic groups. Ion-exchange membrane absorbers are also used for purification of cell culture-derived products, but to a lesser extent.

In most mAb platform processes, ion-exchange steps serve primarily a viral clearance function, in the case of anion exchange, and a polishing function for aggregate removal, in the case of cation exchange. For non-mAb processes, ion exchange serves a much wider variety of functions ranging from capture to polishing. Proteins are usually bound under low ionic strength conditions and eluted at higher ionic strength, although some proteins (mAbs in particular) have demonstrated higher binding capacity at intermediate ionic strength owing to electrostatic exclusion effects [140]. Binding and elution by varying pH can also be employed although this approach is not as widely used and generally requires more time and effort for process development.

Contributions from both thermodynamic and rate factors govern the performance of ion-exchange chromatography. The thermodynamic component is described by the equilibrium adsorption isotherm, while the rate is usually governed by mass transfer resistance. Early generations of ion-exchange media had relatively poor equilibrium capacity and mass transfer characteristics. By comparison, more recent iterations strike a good balance of equilibrium binding capacity and mass transfer rates for many protein biopharmaceuticals.

Mechanistic models of ion-exchange adsorption usually require a description of the equilibrium isotherm and in many instances require a numerical solution of the general diffusional transport model [141]. However, constant pattern solutions for column breakthrough are available for the external film model and the linear driving force model. For instances of strong binding, which can be approximated by a rectangular isotherm and which is frequently encountered with ion-exchange media at low ionic strength, exact analytical solutions are available [141].

5.5 Hydrophobic Interaction and Multimodal Chromatography

Hydrophobic interaction chromatography (HIC) and multimodal chromatography are widely used for purification of cell culture products. Because of the need for salts composed of kosmotropic anions and chaotropic cations (e.g., ammonium sulfate), HIC does tend to have greater drawbacks compared to other modalities. The high concentrations of salt required may necessitate special waste treatment prior to discharge from the manufacturing facility. Moreover, HIC can be quite

sensitive to a number of operating parameters and, therefore, requires more time and effort to develop a robust and scalable unit operation. Nonetheless, HIC unique selectivity often proves indispensable in biopharmaceutical purification, requiring the investment be made to overcome these challenges.

Proteins are commonly bound to HIC ligands at high ionic strength using ammonium sulfate, sodium sulfate, or sodium citrate and eluted by lowing the ionic strength. HIC is used extensively for demanding separations such as removal of product aggregates and fragments. It is most often used in a polishing configuration later in the process. While HIC can be used for product capture, this is relatively rare due to complications that can arise when adding large amounts of salt to cell culture harvest broths and the tendency of HIC resins to become fouled due to the higher proportion of impurities in these feed streams.

Multimodal chromatography utilizes ligands typically composed of both ion-exchange and HIC modalities. In general, these matrices mitigate some of the manufacturing drawbacks of HIC. For example, the use of concentrated salt solutions is often not required for multimodal chromatography. From a process development perspective, the time and expense to develop a multimodal purification step are intermediate between that of HIC and ion exchange. Therefore, it is usually evaluated before HIC for challenging separations that cannot be achieved by less selective methods.

Binding and elution of proteins on multimodal resins can occur under conditions similar to either HIC or ion exchange, depending on the specific nature of the protein. Some proteins show a strong tendency toward one mechanism, while others may be able to utilize either depending on other parameters such as solution pH. Multimodal chromatography is utilized in a wide variety of roles for removal of both process- and product-related impurities.

5.6 Scale-Up of Chromatography Steps

A number of factors must be considered prior to scale-up of process chromatography unit operations. These usually encompass equipment and hardware, column packing and pressure-flow relationships, dynamic binding capacity, yield, selectivity, and resolution [142]. Due to loss of wall support, higher backpressure is encountered in larger diameter columns as scale is increased for the same bed height. This effect is now relatively well understood and predictable using mathematical relationships for compressible media [143]. The quality of column packing is usually evaluated using relatively simple inert tracer pulse response experiments to measure height equivalent to a theoretical plate (HETP) and peak asymmetry. A practical guide to qualification of chromatographic columns and setting limits for these parameters is provided by Rathore et al. [144].

Capacity, selectivity, and resolution are usually optimized using laboratory-scale models. Dynamic binding capacity (DBC) is best measured using representative feedstocks with >1 cm diameter columns at the target bed height and mobile

phase velocity to ensure that both selectivity and resolution are maintained at scale; the process developer needs to have a good understanding of potential sources of variability and the normal operating ranges for all relevant process parameters. For example, aggregate removal by stepwise elution from ion-exchange resins tends to be sensitive to both pH and ionic strength [145]. If the process requires tighter operating ranges than can be achieved in manufacturing, process performance or product quality may suffer. Scale-up of gradient elution and prediction of optimal stepwise elution conditions using relatively simple mathematical relationships has been established by Yamamoto and coworkers [146].

5.7 Viral Safety Considerations

Viral safety of biopharmaceutical preparations is of paramount consideration. For products made using well-characterized cell lines, at least two unit operations with orthogonal clearance mechanisms have to be tested for clearance capability using model viruses [147]. Cell lines used for biopharmaceutical production often contain endogenous retroviruses. Moreover, adventitious viral contamination may occur through ingress from raw materials, cell banks, or process operations. Viral safety must be assured for both sources of viral contamination. In the authors' experience, a large proportion of questions from regulatory authorities concern viral safety matters. Therefore, attention to detail in this area during process development and execution of viral clearance studies is particularly important.

For monoclonal antibody processes, the majority of clearance is usually derived from low-pH virus inactivation (i.e., exposure to low pH, typically 3.6 or lower, sufficient to inactivate enveloped viruses), anion exchange flow through chromatography, and virus filtration (a dedicated membrane filter with a pore size, usually 20–40 nm, smaller than the size of the virus desired to be removed).The cation exchange step can also serve to clear viruses if needed [148]. Low pH, anion exchange, and virus filtration can often provide a combined log reduction value (LRV) of greater than 15 if implemented correctly for xenotropic murine leukemia virus, which is a commonly used model virus for clearance studies.

5.8 Ultrafiltration/Diafiltration (UF/DF)

In many cell culture processes, UF/DF is used as a process step to perform a simple buffer exchange and concentration function. They may also serve to remove small impurities with a high sieving coefficient. In practice, removal of many small-molecule impurities by UF/DF may not follow ideal behavior and can be described using mathematical models developed by Zydney and coworkers [149].

For monoclonal antibodies, the final UF/DF step serves to achieve the final drug substance protein concentration and exchange the product into formulation buffer.

As a result of the industry trend toward high-concentration formulations in order to reduce the volume of product dosed, this step can pose unique challenges. In particular, concentration polarization, high shear rates, and high viscosity can lead to scale-up challenges for this unit operation [150]. Furthermore, the requirement for a large amount of product can make this step challenging to develop even in a laboratory-scale setting. To avoid commonly encountered problems associated with slip-induced shear from peristaltic pumps, it is recommended that high-quality diaphragm pumps be used for both development and manufacturing.

6 Process Economics

Biotechnology is increasingly providing new and potentially effective protein-based treatments against many of humankind's most serious diseases. However, treatment in therapy areas such as oncology, neuroscience, and metabolic disease, all with significant unmet medical need, is not viable unless these novel products can be made in sufficient quantity and sold at a price that payers can afford. In addition, the market is becoming increasingly competitive with efficacy of the new drug being of primary importance but time to market and cost becoming key differentiators. Therefore, as this trend in rapid expansion of next-generation biologic portfolios continues to grow, getting to grips with process economics is very important.

There are a number of fundamental components that affect process economics including titer, scale, purification yield, time in plant, facility costs (depreciation, labor, and utilities), raw materials, and process configuration.

6.1 Titer and Economies of Scale

Recombinant protein titers vary widely depending on expression system and protein class. A well-established expression platform such as Chinese hamster ovary (CHO) expression of therapeutic antibodies can reach 10–15 g/L [126]. Other protein classes (e.g., blood factors) may be expressed at significantly lower levels. The impact on cost of goods manufactured per gram (COGM/g) by titer is very significant. As titer continues to increase, the impact of manufacturing costs become more apparent (Fig. 5). Beyond a certain titer, say above 4 g/L, the effect of COGM/g is less dramatic; however, the plant output continues to increase as titer increases, in the absence of purification bottlenecks. With a finite amount of time that can be devoted to process optimization before the program needs to move into clinical evaluation, the allocation of development effort to increasing titer versus other process improvements, downstream, needs to be carefully considered.

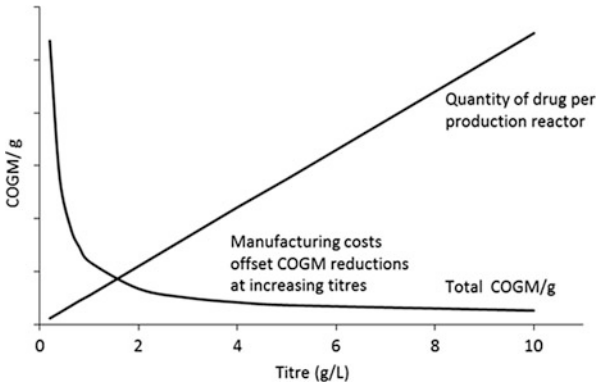


Fig. 5 Economies of scale. Typical cost trend as titer increases

6.2 Purification Yield

The overall process yield is a function of the performance of each individual purification step. Improvements in step yield through process optimization and in some cases removal of non-value added intermediate steps can increase the quantity of product per batch and thus reduce COGM. Typically the more complex the process and the more steps required to produce appropriate quality product, the lower the yield and the more expensive the final drug. For an established product type such as antibodies, process yields have risen in recent years from <50% to typically over 70% [151].

6.3 Batch Duration: Run Rate

A biologic manufacturing facility costs money whether it is used or not. Fixed costs, such as depreciation, labor, and power, are constantly accumulating and will be attributed to the COGM in that facility. Therefore, it is essential to maximize the potential number of batches that the plant can handle. Fermentation duration, number of bioreactors, number of purification suites, and the purification time are some of the key components that dictate run rate. Clearly the number of bioreactors and purification suites may not be easily changed; however, the way that the fermentation and purification process are optimized and run to maximize run rate is of paramount importance.

6.4 Material Costs and Process Configuration

Production of biologics requires the use of raw materials (chromatography resins, filters, buffer constituents, cell culture media, and water). The cost impact of raw

material used is dependent on scale and in certain cases (resins and filters) whether or not components are reused and, if reused, how many times. The decision to incorporate disposable or reuse is complex. While the disposable path may lead to higher raw material costs, this is potentially offset to some extent by reduction of operating costs and other indirect costs such as cleaning validation [152, 153]. The process configuration is ultimately determined through careful consideration and modeling of potential manufacturing scale, campaign size, and other manufacturing considerations, such as whether or not the process will move into a dedicated or a multiuse facility.

6.5 Cost Modeling

The choice of operating conditions, process performance, and scale of manufacture will all impact COGM. Cost models are typically used to help rationally determine the optimum solution for process configuration and potential process development strategies [154]. Developing a protein therapeutic has high uncertainty and is costly and subject to strict regulatory requirements. The old paradigm of "fail fast-fail cheap" in which early proof of principle was demonstrated before significant investments were made is no longer always the case. With some disease areas becoming increasingly well understood, the requirement to rapidly accelerate through the clinical phases leaves little time to make significant process modifications. Development of processes that are economically viable from the beginning of the process life cycle is becoming increasingly necessary. One way it has been possible to satisfy this contradictory expectation of high productivity, and accelerated timelines is by development of "platform processes" that work especially well for a class of molecules (e.g., mAbs). The use of cost models and other in silico process tools to evaluate and prioritize impact of cost-related factors (e.g., titer, scale, and process choice) with clinical data (potential dose, patient population) and manufacturing facilities is becoming fundamental to strategic and rational selection of process requirements and optimization goals [155].

7 Summary

The progress in bioprocess development and scale-up continues to keep up with the changing nature of treatment paradigms and economic pressures. On the one hand, a fairly solid platform process (described in this chapter) consisting of fed-batch bioreactors, three-column chromatography, and two viral clearance steps has been developed and routinely practiced at large scale for the manufacture of monoclonal antibody products to supply markets with high product demand. At the same time, the industry is also retooling itself to respond to personalized medicine approaches calling for greater number of products with lesser demand by developing disposable

bioreactors and columns. Leveraging the experience gained in developing processes for mAbs, progress is also being made in other recombinant proteins, such as Fc-fusion proteins, multi-specific proteins, and even other product modalities such as gene therapy and cell therapies.

Because of the need to evolve with the changing patient needs, bioprocess development and manufacture will never become a mature field; however, thanks to the rapid progress made over the past decade, it is well past its infancy.

References

1. Ecker DM, Jones SD, Levine HL (2015) The therapeutic monoclonal antibody market. MAbs 7(1):9–14
2. Konstantinov KB, Cooney CL (2015) White paper on continuous bioprocessing. May 20–21, 2014 continuous manufacturing symposium. J Pharm Sci 104(3):813–820
3. Razinkov VI, Treuheit MJ, Becker GW (2015) Accelerated formulation development of monoclonal antibodies (mAbs) and mAb-based modalities: review of methods and tools. J Biomol Screen 20(4):468–483
4. Huang CJ, Lin H, Yang X (2012) Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. J Ind Microbiol Biotechnol 39(3):383–399
5. Walsh G, Jefferis R (2006) Post-translational modifications in the context of therapeutic proteins. Nat Biotechnol 24(10):1241–1252
6. Felberbaum RS (2015) The baculovirus expression vector system: a commercial manufacturing platform for viral vaccines and gene therapy vectors. Biotechnol J 10(5):702–714
7. Moustafa K, Makhzoum A, Tremouillaux-Guiller J (2015) Molecular farming on rescue of pharma industry for next generations. Crit Rev Biotechnol 8:1–11
8. Decker EL, Reski R (2012) Glycoprotein production in moss bioreactors. Plant Cell Rep 31 (3):453–460
9. Huang TK, McDonald KA (2012) Bioreactor systems for in vitro production of foreign proteins using plant cell cultures. Biotechnol Adv 30(2):398–409
10. Webster DE, Thomas MC (2012) Post-translational modification of plant-made foreign proteins; glycosylation and beyond. Biotechnol Adv 30(2):410–418
11. Anyaogu DC, Mortensen UH (2015) Manipulating the glycosylation pathway in bacterial and lower eukaryotes for production of therapeutic proteins. Curr Opin Biotechnol 36:122–128
12. Contreras-Gomez A, Sanchez-Miron A, Garcia-Camacho F, Molina-Grima E, Chisti Y (2014) Protein production using the baculovirus-insect cell expression system. Biotechnol Prog 30(1):1–18
13. Gomord V, Fitchette AC, Menu-Bouaouiche L, Saint-Jore-Dupas C, Plasson C, Michaud D, Faye L (2010) Plant-specific glycosylation patterns in the context of therapeutic protein production. Plant Biotechnol J 8(5):564–587
14. Butler M, Spearman M (2014) The choice of mammalian cell host and possibilities for glycosylation engineering. Curr Opin Biotechnol 30:107–112
15. Ghaderi D, Taylor RE, Padler-Karavani V, Diaz S, Varki A (2010) Implications of the presence of N-glycolylneuraminic acid in recombinant therapeutic glycoproteins. Nat Biotechnol 28(8):863–867
16. Ghaderi D, Zhang M, Hurtado-Ziola N, Varki A (2012) Production platforms for biotherapeutic glycoproteins. Occurrence, impact, and challenges of non-human sialylation. Biotechnol Genet Eng Rev 28(1):147–175
17. Dumont J, Euwart D, Mei B, Estes S, Kshirsagar R (2015) Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. Crit Rev Biotechnol 18:1–13

18. Swiech K, Picanco-Castro V, Covas DT (2012) Human cells: new platform for recombinant therapeutic protein production. Protein Expr Purif 84(1):147–153
19. McCue J, Kshirsagar R, Selvitelli K, Lu Q, Zhang M, Mei B, Peters R, Pierce GF, Dumont J, Raso S, Reichert H (2015) Manufacturing process used to produce long-acting recombinant factor VIII Fc fusion protein. Biologicals 43(4):213–219
20. Havenga MJ, Holterman L, Melis I, Smits S, Kaspers J, Heemskerk E, van der Vlugt R, Koldijk M, Schouten GJ, Hateboer G, Brouwer K, Vogels R, Goudsmit J (2008) Serum-free transient protein production system based on adenoviral vector and PER.C6 technology: high yield and preserved bioactivity. Biotechnol Bioeng 100(2):273–283
21. Schiedner G, Hertel S, Bialek C, Kewes H, Waschutza G, Volpers C (2008) Efficient and reproducible generation of high-expressing, stable human cell lines without need for antibiotic selection. BMC Biotechnol 8:13
22. Walsh G (2014) Biopharmaceutical benchmarks 2014. Nat Biotechnol 32(10):992–1000
23. Berting A, Farcet MR, Kreil TR (2010) Virus susceptibility of Chinese hamster ovary (CHO) cells and detection of viral contaminations by adventitious agent testing. Biotechnol Bioeng 106(4):598–607
24. Fischer S, Handrick R, Otte K (2015) The art of CHO cell engineering: a comprehensive retrospect and future perspectives. Biotechnol Adv 33(8):1878–1896
25. Kildegaard HF, Baycin-Hizal D, Lewis NE, Betenbaugh MJ (2013) The emerging CHO systems biology era: harnessing the omics revolution for biotechnology. Curr Opin Biotechnol 24(6):1102–1107
26. Lee JS, Grav LM, Lewis NE, Faustrup Kildegaard H (2015) CRISPR/Cas9-mediated genome engineering of CHO cell factories: application and perspectives. Biotechnol J 10(7):979–994
27. Malphettes L, Freyvert Y, Chang J, Liu PQ, Chan E, Miller JC, Zhou Z, Nguyen T, Tsai C, Snowden AW, Collingwood TN, Gregory PD, Cost GJ (2010) Highly efficient deletion of FUT8 in CHO cell lines using zinc-finger nucleases yields cells that produce completely nonfucosylated antibodies. Biotechnol Bioeng 106(5):774–783
28. Lin N, Mascarenhas J, Sealover NR, George HJ, Brooks J, Kayser KJ, Gau B, Yasa I, Azadi P, Archer-Hartmann S (2015) Chinese hamster ovary (CHO) host cell engineering to increase sialylation of recombinant therapeutic proteins by modulating sialyltransferase expression. Biotechnol Prog 31(2):334–346
29. Yin B, Gao Y, Chung CY, Yang S, Blake E, Stuczynski MC, Tang J, Kildegaard HF, Andersen MR, Zhang H, Betenbaugh MJ (2015) Glycoengineering of Chinese hamster ovary cells for enhanced erythropoietin N-glycan branching and sialylation. Biotechnol Bioeng 112(11):2343–2351
30. Wurm FM (2013) CHO quasispecies—implications for manufacturing processes. Processes 1 (3):296–311
31. Lewis NE, Liu X, Li Y, Nagarajan H, Yerganian G, O'Brien E, Bordbar A, Roth AM, Rosenbloom J, Bian C, Xie M, Chen W, Li N, Baycin-Hizal D, Latif H, Forster J, Betenbaugh MJ, Famili I, Xu X, Wang J, Palsson BO (2013) Genomic landscapes of Chinese hamster ovary cell lines as revealed by the *Cricetulus griseus* draft genome. Nat Biotechnol 31 (8):759–765
32. Puck TT (1957) The genetics of somatic mammalian cells. Adv Biol Med Phys 5:75–101
33. Hu Z, Guo D, Yip SS, Zhan D, Misaghi S, Joly JC, Snedecor BR, Shen AY (2013) Chinese hamster ovary K1 host cell enables stable cell line development for antibody molecules which are difficult to express in DUXB11-derived dihydrofolate reductase deficient host cell. Biotechnol Prog 29(4):980–985
34. Davies SL, Lovelady CS, Grainger RK, Racher AJ, Young RJ, James DC (2013) Functional heterogeneity and heritability in CHO cell populations. Biotechnol Bioeng 110(1):260–274
35. O'Callaghan PM, Berthelot ME, Young RJ, Graham JW, Racher AJ, Aldana D (2015) Diversity in host clone performance within a Chinese hamster ovary cell line. Biotechnol Prog 31(5):1187–1200

36. Derouazi M, Martinet D, Besuchet Schmutz N, Flaction R, Wicht M, Bertschinger M, Hacker DL, Beckmann JS, Wurm FM (2006) Genetic characterization of CHO production host DG44 and derivative recombinant cell lines. Biochem Biophys Res Commun 340(4):1069–1077
37. Jostock T, Knopf HP (2012) Mammalian stable expression of biotherapeutics. Methods Mol Biol 899:227–238
38. Ho SC, Mariati, Yeo JH, Fang SG Yang YS (2015) Impact of using different promoters and matrix attachment regions on recombinant protein expression level and stability in stably transfected CHO cells. Mol Biotechnol 57(2):138–144
39. Lai T, Yang Y Ng SK (2013) Advances in Mammalian cell line development technologies for recombinant protein production. Pharmaceuticals (Basel) 6(5):579–603
40. Fan L, Kadura I, Krebs LE, Hatfield CC, Shaw MM, Frye CC (2012) Improving the efficiency of CHO cell line generation using glutamine synthetase gene knockout cells. Biotechnol Bioeng 109(4):1007–1015
41. Chin CL, Chin HK, Chin CS, Lai ET, Ng SK (2015) Engineering selection stringency on expression vector for the production of recombinant human alpha1-antitrypsin using Chinese Hamster ovary cells. BMC Biotechnol 15:44
42. Saunders F, Sweeney B, Antoniou MN, Stephens P, Cain K (2015) Chromatin function modifying elements in an industrial antibody production platform–comparison of UCOE, MAR, STAR and cHS4 elements. PLoS One 10(4), e0120096
43. Matasci M, Baldi L, Hacker DL, Wurm FM (2011) The PiggyBac transposon enhances the frequency of CHO stable cell line generation and yields recombinant lines with superior productivity and stability. Biotechnol Bioeng 108(9):2141–2150
44. Mayrhofer P, Kratzer B, Sommeregger W, Steinfellner W, Reinhart D, Mader A, Turan S, Qiao J, Bode J, Kunert R (2014) Accurate comparison of antibody expression levels by reproducible transgene targeting in engineered recombination-competent CHO cells. Appl Microbiol Biotechnol 98(23):9723–9733
45. Kennard ML, Goosney DL, Monteith D, Zhang L, Moffat M, Fischer D, Mott J (2009) The generation of stable, high MAb expressing CHO cell lines based on the artificial chromosome expression (ACE) technology. Biotechnol Bioeng 104(3):540–553
46. Geisse S, Fux C (2009) Recombinant protein production by transient gene transfer into Mammalian cells. Methods Enzymol 463:223–238
47. Diepenbruck C, Klinger M, Urbig T, Baeuerle P, Neef R (2013) Productivity and quality of recombinant proteins produced by stable CHO cell clones can be predicted by transient expression in HEK cells. Mol Biotechnol 54(2):497–503
48. Bohm E, Seyfried BK, Dockal M, Graninger M, Hasslacher M, Neurath M, Konetschny C, Matthiessen P, Mitterer A, Scheiflinger F (2015) Differences in N-glycosylation of recombinant human coagulation factor VII derived from BHK, CHO, and HEK293 cells. BMC Biotechnol 15:87
49. Croset A, Delafosse L, Gaudry JP, Arod C, Glez L, Losberger C, Begue D, Krstanovic A, Robert F, Vilbois F, Chevalet L, Antonsson B (2012) Differences in the glycosylation of recombinant proteins expressed in HEK and CHO cells. J Biotechnol 161(3):336–348
50. Cain K, Peters S, Hailu H, Sweeney B, Stephens P, Heads J, Sarkar K, Ventom A, Page C, Dickson A (2013) A CHO cell line engineered to express XBP1 and ERO1-Lalpha has increased levels of transient protein expression. Biotechnol Prog 29(3):697–706
51. Daramola O, Stevenson J, Dean G, Hatton D, Pettman G, Holmes W, Field R (2014) A high-yielding CHO transient system: coexpression of genes encoding EBNA-1 and GS enhances transient protein expression. Biotechnol Prog 30(1):132–141
52. Kunaparaju R, Liao M, Sunstrom NA (2005) Epi-CHO, an episomal expression system for recombinant protein production in CHO cells. Biotechnol Bioeng 91(6):670–677
53. Rajendra Y, Hougland MD, Alam R, Morehead TA, Barnard GC (2015) A high cell density transient transfection system for therapeutic protein expression based on a CHO GS-knockout cell line: process development and product quality assessment. Biotechnol Bioeng 112 (5):977–986

54. Jager V, Bussow K, Schirrmann T (2015) Transient recombinant protein expression in mammalian cells. In: Al-Rubeai M (ed) Animal cell culture. Springer, Dordrecht
55. Steger K, Brady J, Wang W, Duskin M, Donato K, Peshwa M (2015) CHO-S antibody titers >1 gram/liter using flow electroporation-mediated transient gene expression followed by rapid migration to high-yield stable cell lines. J Biomol Screen 20(4):545–551
56. Geisse S, Voedisch B (2012) Transient expression technologies: past, present, and future. Methods Mol Biol 899:203–219
57. Sou SN, Polizzi KM, Kontoravdi C (2013) Evaluation of transfection methods for transient gene expression in Chinese hamster ovary cells. Adv Biosci Biotechnol 04(12):1013–1019
58. Browne SM, Al-Rubeai M (2007) Selection methods for high-producing mammalian cell lines. Trends Biotechnol 25(9):425–432
59. Coller HA, Coller BS (1986) Poisson statistical analysis of repetitive subcloning by the limiting dilution technique as a way of assessing hybridoma monoclonality. Methods Enzymol 121:412–417
60. Onadipe AO, Metcalfe HK Freeman PR, James C (2001) Capillary-aided cell cloning: a technique for one step cloning with high probability of monoclonality. In: Lindner-Olsson EC, Lüllau N (eds) Animal cell technology: from target to market. Springer, Netherlands
61. Nakamura T, Omasa T (2015) Optimization of cell line development in the GS-CHO expression system using a high-throughput, single cell-based clone selection system. J Biosci Bioeng 120(3):323–329
62. Evans K, Albanetti T, Venkat R, Schoner R, Savery J, Miro-Quesada G, Rajan B, Groves C (2015) Assurance of monoclonality in one round of cloning through cell sorting for single cell deposition coupled with high resolution cell imaging. Biotechnol Prog 31(5):1172–1178
63. DeMaria CT, Cairns V, Schwarz C, Zhang J, Guerin M, Zuena E, Estes S, Karey KP (2007) Accelerated clone selection for recombinant CHO CELLS using a FACS-based high-throughput screen. Biotechnol Prog 23(2):465–472
64. Mazutis L, Gilbert J, Ung WL, Weitz DA, Griffiths AD, Heyman JA (2013) Single-cell analysis and sorting using droplet-based microfluidics. Nat Protoc 8(5):870–891
65. Joensson HN, Zhang C, Uhlen M, Andersson-Svahn H (2012) A homogeneous assay for protein analysis in droplets by fluorescence polarization. Electrophoresis 33(3):436–439
66. Silk NJ, Denby S, Lewis G, Kuiper M, Hatton D, Field R, Baganz F, Lye GJ (2010) Fed-batch operation of an industrial cell culture process in shaken microwells. Biotechnol Lett 32 (1):73–78
67. Rameez S, Mostafa SS, Miller C, Shukla AA (2014) High-throughput miniaturized bioreactors for cell culture process development: reproducibility, scalability, and control. Biotechnol Prog 30(3):718–727
68. Paul AJ, Schwab K, Hesse F (2014) Direct analysis of mAb aggregates in mammalian cell culture supernatant. BMC Biotechnol 14:99
69. Yang Y, Strahan A, Li C, Shen A, Liu H, Ouyang J, Katta V, Francissen K, Zhang B (2010) Detecting low level sequence variants in recombinant monoclonal antibodies. MAbs 2 (3):285–298
70. Ambrogelly A, Liu YH, Li H, Mengisen S, Yao B, Xu W, Cannon-Carlson S (2012) Characterization of antibody variants during process development: the tale of incomplete processing of N-terminal secretion peptide. MAbs 4(6):701–709
71. Harris RP, Kilby PM (2014) Amino acid misincorporation in recombinant biopharmaceutical products. Curr Opin Biotechnol 30:45–50
72. Khetan A, Huang YM, Dolnikova J, Pederson NE, Wen D, Yusuf-Makagiansar H, Chen P, Ryll T (2010) Control of misincorporation of serine for asparagine during antibody production using CHO cells. Biotechnol Bioeng 107(1):116–123
73. Kim M, O'Callaghan PM, Droms KA, James DC (2011) A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies. Biotechnol Bioeng 108(10):2434–2446

74. Zhang S, Bartkowiak L, Nabiswa B, Mishra P, Fann J, Ouellette D, Correia I, Regier D, Liu J (2015) Identifying low-level sequence variants via next generation sequencing to aid stable CHO cell line screening. Biotechnol Prog 31(4):1077–1085
75. Seth G (2012) Freezing mammalian cells for production of biopharmaceuticals. Methods 56 (3):424–431
76. Capes-Davis A, Theodosopoulos G, Atkin I, Drexler HG, Kohara A, MacLeod RA, Masters JR, Nakamura Y, Reid YA, Reddel RR, Freshney RI (2010) Check your cultures! A list of cross-contaminated or misidentified cell lines. Int J Cancer 127(1):1–8
77. Mclean C, Harbour C (2013) Contamination detection in animal cell culture. In: Flickinger MC (ed) Upstream industrial biotechnology. Wiley, New Jersey
78. Moody M, Alves W, Varghese J, Khan F (2011) Mouse Minute Virus (MMV) contamination–a case study: detection, root cause determination, and corrective actions. PDA J Pharm Sci Technol 65(6):580–588
79. Cabannes E, Hebert C, Eloit M (2014) Whole genome: next-generation sequencing as a virus safety test for biotechnological products. PDA J Pharm Sci Technol 68(6):631–638
80. Shepherd AJ, Wilson NJ, Smith KT (2003) Characterisation of endogenous retrovirus in rodent cell lines used for production of biologicals. Biologicals 31(4):251–260
81. Dinowitz M, Lie YS, Low MA, Lazar R, Fautz C, Potts B, Sernatinger J, Anderson K (1992) Recent studies on retrovirus-like particles in Chinese hamster ovary cells. Dev Biol Stand 76:201–207
82. Gramer MJ, Goochee CF (1993) Glycosidase activities in Chinese hamster ovary cell lysate and cell culture supernatant. Biotechnol Prog 9(4):366–373
83. Ozturk SS (2014) Equipment for large-scale mammalian cell culture. Adv Biochem Eng Biotechnol 139:69–92
84. Arnaud CH (2015) Disposable plastic bioreactors lead to savings—and challenges—for biopharma firms. Chem Eng News 93(46):10–13
85. Wurm FM (2004) Production of recombinant protein therapeutics in cultivated mammalian cells. Nat Biotechnol 22(11):1393–1398
86. Milian E, Kamen AA (2015) Current and emerging cell culture manufacturing technologies for influenza vaccines. Biomed Res Int 2015:504831
87. Woodside SM, Bowen BD, Piret JM (1998) Mammalian cell retention devices for stirred perfusion bioreactors. Cytotechnology 28(1–3):163–175
88. Robin J (2013) Case study: challenges and learning in implementing ATF perfusion process. Integrated continuous biomanufacturing ECI conference, Castelldefels, 20–24 October 2013. http://www.engconf.org/staging/wp-content/uploads/2013/12/jarno_ICB-13AQ-Monday.pdf
89. Wu J (1995) Mechanisms of animal cell damage associated with gas bubbles and cell protection by medium additives. J Biotechnol 43(2):81–94
90. Hsu WT, Aulakh RP, Traul DL, Yuk IH (2012) Advanced microscale bioreactor system: a representative scale-down model for bench-top bioreactors. Cytotechnology 64(6):667–678
91. Ahuja S, Bui T, Chen J, Chen J, Dorotheo R, Jain S, Lee A, Russell B, Singh S, Qu L (2011) Development and scale-up of a high titer cell culture process. Abstracts of papers of the American Chemical Society, American Chemical Society, Washington, DC
92. Datta P, Linhardt RJ, Sharfstein ST (2013) An omics approach towards CHO cell engineering. Biotechnol Bioeng 110(5):1255–1271
93. Gupta P, Lee KH (2007) Genomics and proteomics in process development: opportunities and challenges. Trends Biotechnol 25(7):324–330
94. Antoniewicz MR (2015) Methods and advances in metabolic flux analysis: a mini-review. J Ind Microbiol Biotechnol 42(3):317–325
95. Hines M, Holmes C, Schad R (2010) Simple strategies to improve bioprocess pure culture processing. Pharm Eng 30(3):1–11
96. Tsui V, Wiederhold W (2007) A practical approach to steam autoclave cycle development. J Validation Technol 13(2):124

97. Schleh M, Lawrence B, Park T, Rosenthal S, Hart R, Dehghani H (2010) Effectiveness of upstream barrier technologies for inactivation of adventitious contaminants of cell culture. Am Pharm Rev 13(7):72
98. Goetschalckx S, Fabre V, Wynants M, Bertaux L, Plavsic M, Boussif O, Laenen L (2014) A holistic biosafety risk mitigation approach. Am Pharm Rev 17(4):48–56
99. Yen S, Sokolenko S, Manocha B, Blondeel EJ, Aucoin MG, Patras A, Daynouri-Pancino F, Sasges M (2014) Treating cell culture media with UV irradiation against adventitious agents: minimal impact on CHO performance. Biotechnol Prog 30(5):1190–1195
100. Weber A, Husemann U, Chaussin S, Adams T, De Wilde D, Gerighausen S, Greller G, Fenge C (2014) Development and qualification of a scalable, disposable bioreactor for GMP-compliant cell culture. Bioprocess Int 12(S5):47
101. Keijzer T, Kakes E, Van Halsema E (2011) Advances in the design of bioreactor systems. Innov Pharm Technol 60–64. http://www.iptonline.com/articles/public/advancesinthedesignofbioreactorsystems.pdf
102. Benz GT (2011) Bioreactor design for chemical engineers. Chem Eng Prog 107:21–26
103. Mirro R, Voll K (2009) Which impeller is right for your cell line. BioProcess Int 7(1):52–58
104. Nienow AW (2006) Reactor engineering in large scale animal cell culture. Cytotechnology 50(1-3):9–33
105. Nienow AW (1996) Gas-liquid mixing studies: a comparison of Rushton turbines with some modern impellers. Chem Eng Res Design 74(A4):417–423
106. Langheinrich C, Nienow AW (1999) Control of pH in large-scale, free suspension animal cell bioreactors: alkali addition and pH excursions. Biotechnol Bioeng 66(3):171–179
107. Hu W, Berdugo C, Chalmers JJ (2011) The potential of hydrodynamic damage to animal cells of industrial relevance: current understanding. Cytotechnology 63(5):445–460
108. Hu W, Wiltberger K (2014) Industrial cell culture process scale-up strategies and considerations. In: Hauser H, Wagner R (eds) Animal cell biotechnology: in biologics production. Walter de Gruyter GmbH & Co. KG, Berlin, pp 455–488
109. Perez JAS, Porcel EMR, Lopez JLC, Sevilla JMF, Chisti Y (2006) Shear rate in stirred tank and bubble column bioreactors. Chem Eng J 124(1–3):1–5
110. Villiger TK (2015) Bioprocess engineering framework to control protein N-linked glycosylation, Diss., Eidgenössische Technische Hochschule ETH Zürich, Nr. 22727
111. Mollet M, Ma N, Zhao Y, Brodkey R, Taticek R, Chalmers JJ (2004) Bioprocess equipment: characterization of energy dissipation rate and its potential to damage cells. Biotechnol Prog 20(5):1437–1448
112. Meghrous J, Khramtsov N, Buckland BC, Cox MM, Palomares LA, Srivastava IK (2015) Dissolved carbon dioxide determines the productivity of a recombinant hemagglutinin component of an influenza vaccine produced by insect cells. Biotechnol Bioeng 112 (11):2267–2275
113. Sieblist C, Hageholz O, Aehle M, Jenzsch M, Pohlscheidt M, Lubbert A (2011) Insights into large-scale cell-culture reactors: II. Gas-phase mixing and CO_2 stripping. Biotechnol J 6 (12):1547–1556
114. Voisard D, Meuwly F, Ruffieux PA, Baer G, Kadouri A (2003) Potential of cell retention techniques for large-scale high-density perfusion culture of suspended mammalian cells. Biotechnol Bioeng 82(7):751–765
115. Ambler CM (1959) The theory of scaling up laboratory data for the sedimentation type centrifuge. J Biochem Microbiol Technol Eng 1(2):185–205
116. Boychyn M, Yim SSS, Shamlou PA, Bulmer M, More J, Hoare A (2001) Characterization of flow intensity in continuous centrifuges for the development of laboratory mimics. Chem Eng Sci 56(16):4759–4770
117. Hutchinson N, Bingham N, Murrell N, Farid S, Hoare M (2006) Shear stress analysis of mammalian cell suspensions for prediction of industrial centrifugation and its verification. Biotechnol Bioeng 95(3):483–491

118. Tait AS, Aucamp JP, Bugeon A, Hoare M (2009) Ultra scale-down prediction using microwell technology of the industrial scale clarification characteristics by centrifugation of mammalian cell broths. Biotechnol Bioeng 104(2):321–331
119. Carman P (1937) Cyclic permeability of granular material. Transl Inst Chem Eng 15:150–167
120. Badmington F, Wilkins R, Payne M, Honig ES (1995) Vmax testing for practical microfiltration train scale-up in biopharmaceutical processing. Pharm Technol 19(9):64–76
121. Lutz H (2009) Rationally defined safety factors for filter sizing. J Membr Sci 341 (1–2):268–278
122. de Zafra CLZ, Quarmby V, Francissen K, Vanderlaan M, Zhu-Shimoni J (2015) Host cell proteins in biotechnology-derived products: a risk assessment framework. Biotechnol Bioeng 112(11):2284–2291
123. CMC-Biotech-Working-Group (2009) A-Mab: a case study in bioprocess development. CASSS, Emeryville
124. Yang H (2013) Establishing acceptable limits of residual DNA. PDA J Pharm Sci Technol 67 (2):155–163
125. Gottschalk U (2008) Bioseparation in antibody manufacturing: the good, the bad and the ugly. Biotechnol Prog 24(3):496–503
126. Kelley B (2009) Industrialization of mAb production technology: the bioprocessing industry at a crossroads. MAbs 1(5):443–452
127. Liu HF, Ma J, Winter C, Bayer R (2010) Recovery and purification process development for monoclonal antibody production. MAbs 2(5):480–499
128. Kelley B (2007) Very large scale monoclonal antibody purification: the case for conventional unit operations. Biotechnol Prog 23(5):995–1008
129. Trexler-Schmidt M, Sze-Khoo S, Cothran AR, Thai BQ, Sargis S, Lebreton B, Kelley B, Blank GS (2009) Purification strategies to process 5 g/L titers of monoclonal antibodies. Biopharm Int 22:8–15
130. Gouda H, Shiraishi M, Takahashi H, Kato K, Torigoe H, Arata Y, Shimada I (1998) NMR study of the interaction between the B domain of staphylococcal protein A and the Fc portion of immunoglobulin G. Biochemistry 37(1):129–136
131. Kim HK, Thammavongsa V, Schneewind O, Missiakas D (2012) Recurrent infections and immune evasion strategies of *Staphylococcus aureus*. Curr Opin Microbiol 15(1):92–99
132. Silverman GJ (1998) B cell superantigens: possible roles in immunodeficiency and autoimmunity. Semin Immunol 10(1):43–55
133. Surolia A, Pain D, Khan MI (1982) Protein A: nature's universal anti-antibody. Trends Biochem Sci 7(2):74–76
134. Graille M, Stura EA, Corper AL, Sutton BJ, Taussig MJ, Charbonnier JB, Silverman GJ (2000) Crystal structure of a Staphylococcus aureus protein A domain complexed with the Fab fragment of a human IgM antibody: structural basis for recognition of B-cell receptors and superantigen activity. Proc Natl Acad Sci U S A 97(10):5399–5404
135. Hober S, Nord K, Linhult M (2007) Protein A chromatography for antibody purification. J Chromatogr B Anal Technol Biomed Life Sci 848(1):40–47
136. Starovasnik MA, O'Connell MP, Fairbrother WJ, Kelley RF (1999) Antibody variable region binding by Staphylococcal protein A: thermodynamic analysis and location of the Fv binding site on E-domain. Protein Sci 8(7):1423–1431
137. Nilsson B, Moks T, Jansson B, Abrahmsen L, Elmblad A, Holmgren E, Henrichson C, Jones TA, Uhlen M (1987) A synthetic Igg-binding domain based on staphylococcal protein-A. Protein Eng 1(2):107–113
138. Tashiro M, Tejero R, Zimmerman DE, Celda B, Nilsson B, Montelione GT (1997) High-resolution solution NMR structure of the Z domain of staphylococcal protein A. J Mol Biol 272(4):573–590
139. Boedeker B (2001) Production processes of licensed recombinant factor VIII preparations. Semin Thromb Hemost 27(4):385–394

140. Zydney AL, Harinarayan C, van Reis R (2009) Modeling electrostatic exclusion effects during ion exchange chromatography of monoclonal antibodies. Biotechnol Bioeng 102 (4):1131–1140
141. Carta G, Ubiera AR, Pabst TM (2005) Protein mass transfer kinetics in ion exchange media: measurements and interpretations. Chem Eng Technol 28(11):1252–1264
142. Hagel L, Jagschies G, Sofer GK (2007) Handbook of process chromatography: development, manufacturing, validation and economics, 2nd edn. Academic/Elsevier, Amsterdam
143. Stickel JJ, Fotopoulos A (2001) Pressure-flow relationships for packed beds of compressible chromatography media at laboratory and production scale. Biotechnol Prog 17(4):744–751
144. Rathore AS, Kennedy RM, O'Donnell JK, Bemberis I, Kaltenbrunner O (2003) Qualification of a chromatographic column – why and how to do it. Biopharm Int Appl Technol Biopharm Dev 16(3):30
145. Suda EJ, Thomas KE, Pabst TM, Mensah P, Ramasubramanyan N, Gustafson ME, Hunter AK (2009) Comparison of agarose and dextran-grafted agarose strong ion exchangers for the separation of protein aggregates. J Chromatogr A 1216(27):5256–5264
146. Ishihara T, Yamamoto S (2005) Optimization of monoclonal antibody purification by ion-exchange chromatography – application of simple methods with linear gradient elution experimental data. J Chromatogr A 1069(1):99–106
147. Remington KM (2015) Fundamental strategies for viral clearance part 2: technical approaches. Bioprocess Int 13(5)
148. Connell-Crowley L, Nguyen T, Bach J, Chinniah S, Bashiri H, Gillespie R, Moscariello J, Hinckley P, Dehghani H, Vunnum S, Vedantham G (2012) Cation exchange chromatography provides effective retrovirus clearance for antibody purification processes. Biotechnol Bioeng 109(1):157–165
149. Shao J, Zydney AL (2004) Optimization of ultrafiltration/diafiltration processes for partially bound impurities. Biotechnol Bioeng 87(3):286–292
150. Shire SJ, Shahrokh Z, Liu J (2004) Challenges in the development of high protein concentration formulations. J Pharm Sci 93(6):1390–1402
151. Werner RG (2004) Economic aspects of commercial manufacture of biopharmaceuticals. J Biotechnol 113(1-3):171–182
152. Sinclair A, Monge M (2002) Quantitative economic evaluation of single use disposables in bioprocessing. Pharm Eng 22(3):20–34
153. Sinclair A, Monge M (2005) Concept facility based on single-use systems, part 2. BioProcess Int 3(9)
154. Hill C, Sinclair A (2007) Maximizing the use of process data from development to manufacturing. Biopharm Int 20(7):38–42
155. Farid SS, Novais JL, Karri S, Washbrook J, Titchener-Hooker NJ (2000) A tool for modeling strategic decisions in cell culture manufacturing. Biotechnol Prog 16(5):829–836

Top Med Chem (2017) 21: 41–68
DOI: 10.1007/7355_2015_5004

Published online: 15 March 2016

Production of Protein Therapeutics in the Quality by Design (QbD) Paradigm

Anurag S. Rathore and Sumit K. Singh

Abstract Biotech products and processes are complex. Our understanding of how the process affects product quality is incomplete and that of how the various quality attributes of the product affect the clinical safety and efficacy is even more limited. Quality by Design (QbD)-based process and product development aims at improving this understanding. In this chapter, we briefly introduce the concept of QbD in the context of biotherapeutics. Next we discuss the various unit operations that together make a typical process. Recent advancements in the manufacturing of biotech therapeutics will also be presented. The importance of identifying the underlying relationship between the quality attributes of the product and clinical safety and efficacy for ultimate realization of QbD goals is discussed in the last section of the chapter. Future perspective of the increasingly important role that QbD is likely to play for the manufacturing of drugs for an increasingly global market is presented as the concluding note of this chapter.

Keywords Biosimilars, Downstream processing, Process analytical technology, Quality by design, Safety and efficacy, Upstream processing

Contents

A.S. Rathore (✉) and S.K. Singh
Department of Chemical Engineering, Indian Institute of Technology, Hauz Khas, New Delhi 110016, India
e-mail: asrathore@biotechcmz.com

1 Introduction

Recombinant therapeutics, monoclonal antibodies, and industrial enzymes together create a multibillion dollar market, global sales of which are expected to rise up to US$169 billion in 2015 [1, 2]. In this scenario, accelerating process development and achieving consistency in process performance and product quality while reducing treatment costs is a goal that is widely shared among manufacturers of the major biotech therapeutics. Quality by Design (QbD) has been demonstrated to be a possible solution to this conundrum as an approach toward process and product development [3, 4]. Recent and ongoing patent expirations and the focus on cost reduction have also fuelled interest in development of biosimilars (follow-on versions after patent expiry of the originator) [5].

QbD aims to develop products using a well-defined and controlled process based on the deep understanding of the product itself [3, 4]. The production of biotech therapeutics under this paradigm involves designing the production process such that the quality target product profile (qTPP) and the critical quality attributes (CQA) of the biotherapeutic are well defined and the desired profile is consistently achieved. Considering the fact that even innovators are known to observe differences in product quality of batches manufactured in the same facility [6], a key challenge is deciding on which differences are acceptable and are likely not to have significant clinical ramifications [7, 8].

A typical biologic is a complex molecule with numerous quality attributes, many of which are CQA and can have a significant effect on the clinical safety and/or efficacy of the product [8]. For instance, a monoclonal antibody (mAb) molecule expressed in mammalian cells can have as many as 40–50 attributes that define its structure and function. This compares to a much smaller number for a typical small-molecule (pharmaceutical) drug entity which may have a handful of CQA. This complexity brings along with it a significant analytical burden and generally necessitates the use of high-end analytical tools that are, in most cases, expensive and require a skilled workforce [9]. Catering to these needs, evolution and

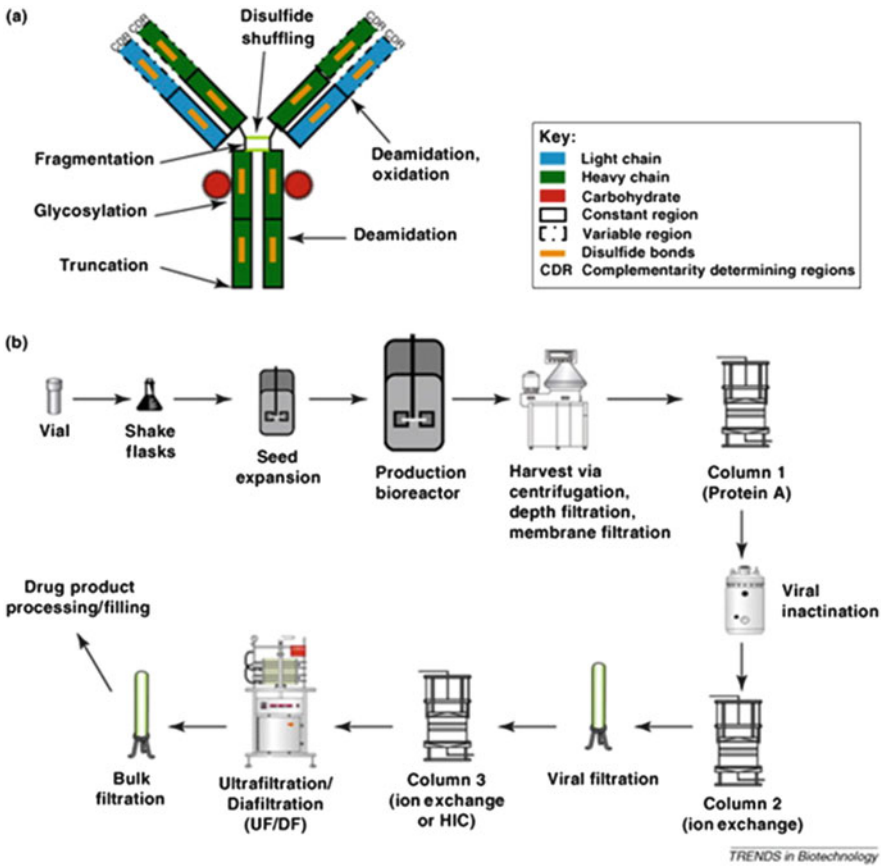


Fig. 1 Illustration of the various steps in biologic manufacturing and the inherent sources of variability in each of these steps. The large number of process steps and the variability at each of these steps make the entire bioprocessing a complex activity and thereby affecting CQAs of the final product in a rather unanticipated manner. (**a**) A typical mAb molecule with multiple sites of modifications. (**b**) Number of process steps involved in manufacturing of a typical biotech product. Figure adapted with copyright permission from [8]

advancements in analytical methodologies that are capable of deciphering differences even at single amino acid resolution have occurred over the past decade [10–14]. Despite all these advancements, we are far off from achieving complete understanding of how each of the different CQA impacts the clinical performance of a drug [15]. Figure 1 illustrates the complexity involved in a biologic manufacturing.

As a consequence of the aforementioned, implementation of QbD principles can certainly overcome and simplify the overwhelming challenges confronted during the manufacturing of biotherapeutics [8]. Increasing adoption of process analytical technology (PAT) in manufacturing settings is a positive step in this direction

[16, 17]. Numerous publications have reported advancements in bioprocessing as well as safety and efficacy testing for biologics [18–23].

2 What Is QbD?

Quality by Design is a *systematic approach* that begins with *predefined objectives* and emphasizes *product and process understanding and process control*. It is an approach based on principles *sound science and quality risk management* [24]. Various steps that are involved in the implementation of QbD process have been illustrated in Fig. 2 [3, 4]. Product knowledge is derived during the drug discovery process and the preclinical and the clinical trials and is continually enhanced as the prospective drug candidate evolves through its development life cycle [4]. This

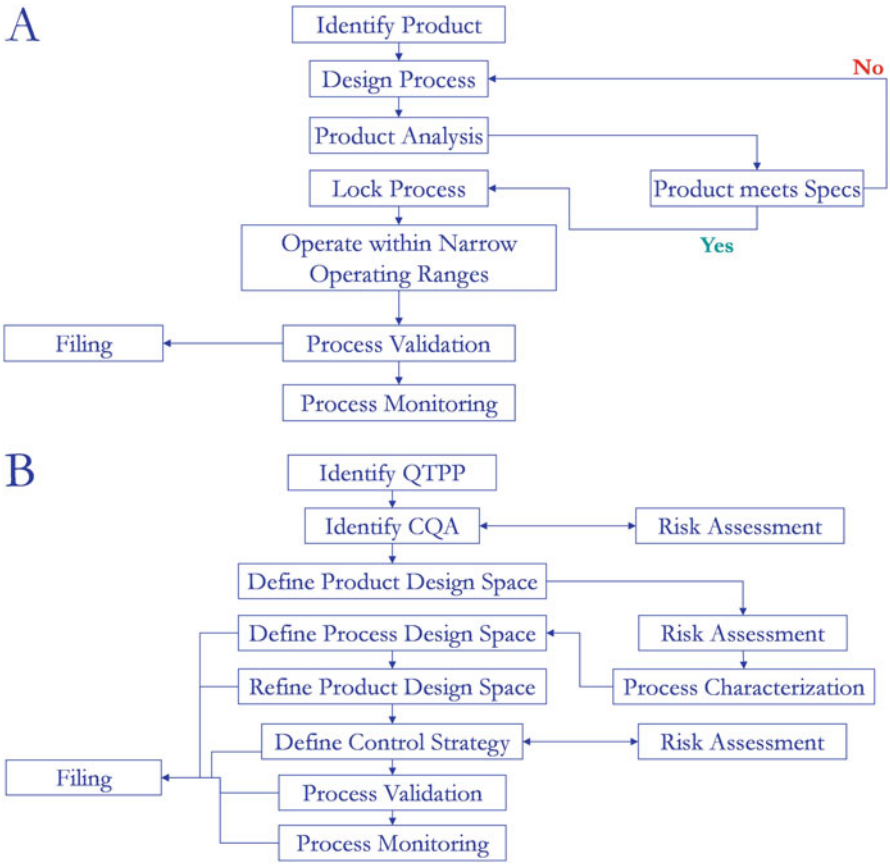


Fig. 2 Illustration of various steps in process and product development by (**a**) traditional approach and (**b**) QbD paradigm. Republished with copyright permission from [25]

Table 1 Key terms used in QbD implementation for the production of biotherapeutics

Term	Definition
Quality target product profile (*qTPP*)	A prospective and dynamic summary of the quality characteristics of a drug product that ideally is sought be achieved to ensure that the desired quality, and thus the safety and efficacy, of a drug product is realized
Critical quality attribute (*CQA*)	A physical, chemical, biological, or microbiological property or characteristic that should be within an appropriate limit, range, or distribution to ensure the desired product quality
Critical process parameters (*CPP*)	A process parameter whose variability has an impact on a critical quality attribute and, therefore, should be monitored or controlled to ensure the process produces the desired quality
Design space	The multidimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality
Quality risk management (*QRM*)	A systematic process for the assessment, control, communication, and review of risks to the quality of the drug (medicinal) product across the product life cycle
Real-time release testing (*RTRT*)	The ability to evaluate and ensure the quality of in-process and/or final product based on process data, which typically include a valid combination of measured material attributes and process controls
Continuous process verification	An alternative approach to process validation in which manufacturing process performance is continuously monitored and evaluated
Knowledge management	Systematic approach to acquiring, analyzing, storing, and disseminating information related to products, manufacturing processes, and components

Definitions taken as per (ICH Guideline Q11 [26] and Q10 [27])

product knowledge is in turn capitalized for identifying the quality target product profile (qTPP) and the critical quality attributes (CQA). Process understanding is then created and leveraged to identify the critical process parameters (CPP). Table 1 lists out the definitions of some of the important terminologies used in the context of QbD. Product understanding and process capability are then utilized to design specifications for in-process, drug substance, and drug product attributes (defining process and product design space). Sources of information that come in handy in establishing such specifications include clinical and nonclinical studies, prior published literature, and process capability with respect to variability observed in the manufactured lots [4]. In the QbD paradigm, it is also important that the process be controlled appropriately so as to reduce variability in the product quality not just in the final drug product but rather at every stage of production. This is the basis for PAT and requires the use of appropriate process monitoring tools and robust process controls so as to steer the process on the predefined trajectory in real time [28–30]. Other elements that play a key role in the implementation of QbD for product of biotherapeutics include characterization and control of raw materials, knowledge management, and continuous risk assessment and management.

In the sections to follow, QbD concepts that are being practiced by the biotechnology industry in upstream and downstream segments of their manufacturing are discussed. Analytical approaches that are capable of elucidating the crucial link between quality attributes of product and the safety and efficacy of the product are also discussed. Finally, the future roadmap for the implementation of QbD for biotech products is presented.

3 QbD in Upstream Processing

Upstream processing in a biopharma facility involves series of steps from cell-line development, media and feed development, shake flask operations, and production bioreactor [31]. The aim at each of these steps is to maximize the productivity of cells (and the product) without negatively impacting product quality.

3.1 Raw Material Characterization

A very large number of raw materials are used in upstream processing and at times some of them (such as yeastolates) may not be chemically defined. These components are known to significantly contribute to the variability in process performance and that in product quality [32]. Over the last several decades, the industry has spent considerable efforts toward understanding how different process parameters affect bioreactor performance. As a result, it is not uncommon to find raw material variability as a principal contributor to the variability in process performance and that in product quality. The complexity arises from the fact that it is not practically feasible to examine the effect of each component (anywhere from 20 to 50) and each process parameter (anywhere from 10 to 20) and their interactions. Common practice in industry is to resort to some form of risk analysis to identify the process parameters and raw materials that can be examined via lab-scale experimentation (typically less than 10).

In the QbD paradigm, a risk-based approach is undertaken for raw material management depending on the development stage of the product [33]. Under this strategy, raw materials are classified into critical, key, and non-key [3]. Those components falling under the critical and key categories are known to have significant impact on product quality or process consistency, respectively. Thus, they warrant heightened focus in terms of characterization and monitoring. In contrast, components in the non-key category can be managed through the supplier assessment procedures as defined by the company's internal quality system. Figure 3 presents the various assessment approaches that are followed for determining the criticality of raw materials.

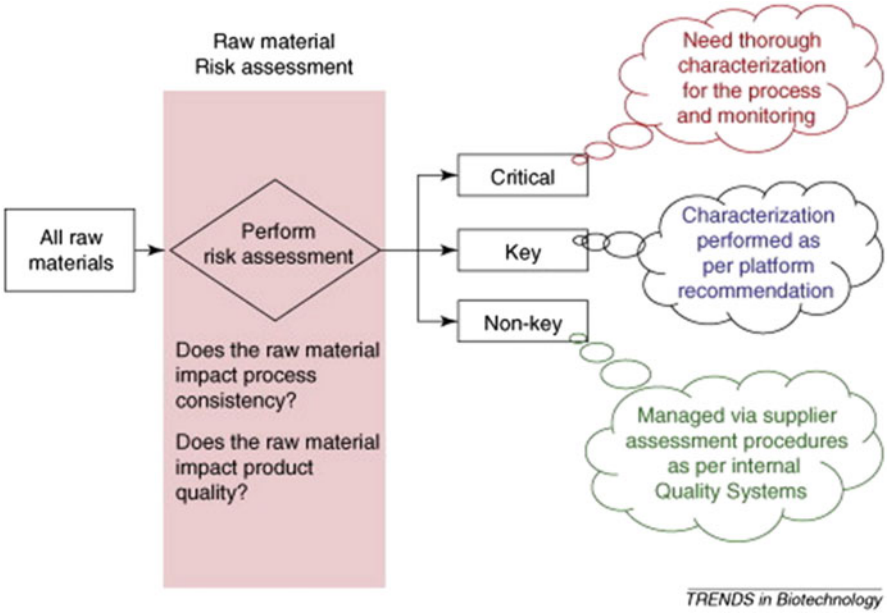


Fig. 3 Different risk assessment approaches to determine the criticality of raw materials. Republished with copyright permission from [3]

3.2 Product and Process Understanding

Upstream processing is known to significantly and in a lot of cases solely impact several of the quality attributes of biotherapeutic products. These include glycosylation profile, charge variant (acidic/basic) distribution, and other product-related impurities [34]. The cell line used (microbial, yeast, insect, plant, and mammalian) has been known to have a considerable effect on many of the CQAs of the product [8, 35].

The implementation of QbD in bioprocessing involves understanding how the various upstream input process parameters and raw material attributes affect the process outcome. This information is typically expressed in the form of correlations between upstream process outputs such as product yield, viable cell density, product quality attributes such as glycosylation profile, and the various process parameters and raw material attributes [15]. Since QbD recommends a bottom-up approach (product defines process) where the process is designed keeping in sight the product profile as per the end-user (patient) requirement, upstream processing being far away from the end product may have an indirect impact on many of the CQA [32]. Still, as mentioned earlier, there are several CQA (most notably the glycosylation profile) that are primarily defined by the upstream process. In order to achieve consistency in these CQA, critical process parameters (CPP) such as pH, temperature, dissolved oxygen level, substrate, and other nutrient ratios need to be

Table 2 CQA and CPP reported for some of the commonly used expression systems for manufacturing of biotherapeutic products. The list is not meant to be exhaustive but rather illustrative

Expression system	Product characteristics	Typical CQA	Typical CPP
E. coli	Low molecular weight with limited disulfide bonding and no glycosylation	Misfolds, HCP, endotoxin, PEG variants, misincorporations, truncations	pH, temperature, dissolved oxygen, substrate, and other nutrient ratio
Pichia/ yeast	Complex proteins with minimal glycosylation	Mannose residues, HCP, misincorporations, truncations	Viable cell density, dissolved oxygen, pH, temperature, harvest time
CHO	Complex glycosylated proteins	Sialic acid, galactose, xylose, fucose residues, glycation, HCP	Viable cell density, specific/ volume productivity, apoptotic %, agitation, cell-line stability, CO_2%, harvest time

Information summarized from [32, 36, 38]

controlled appropriately [36, 37]. Table 2 presents a review of the reported CPP and CQA associated for some of the commonly used expression systems.

3.3 Monitoring and Control of Upstream Processes

A robust monitoring and an advanced, adaptive control system (rather than fixed controls) forms the lynchpin for the successful implementation of QbD paradigm in upstream processing [34]. A list of different process variables and output parameters that are typically monitored and controlled during a typical bioprocess operation is presented in Fig. 4.

A few recent examples of QbD implementation in upstream processing include:

1. Use of near-infrared spectroscopy (NIRS) for quantitative prediction of single analyte in less than 2 min [39, 40].
2. Use of mid-infrared spectroscopy (MIRS) for monitoring multiple analytes in complex fermentation broth [41].
3. Use of online biomass probe for the estimation of cell density and cell viability [42, 43].
4. Use of off-gas analyzers for real-time measurement of carbon dioxide and oxygen concentration of bioreactor exhaust gas [36, 43].
5. Use of fluorescence probes for accurate assessment of actual metabolic state of the cells in the culture [44].
6. Use of dielectric capacitance spectroscopy (DSC) for the estimation of viable cell concentration based on the measurement of capacitance [34, 45].
7. Use of automated flow cytometry for gaining multidimensional information on cell population [34, 44–46].

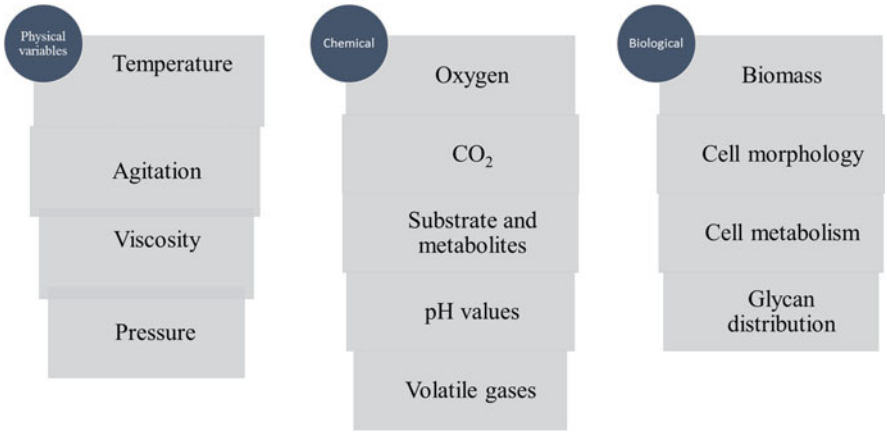


Fig. 4 List of process parameters and product quality attributes that are typically monitored during bioprocess

8. Use of nuclear magnetic resonance (NMR) spectroscopy for identifying and quantification of levels of various nutrient components in mammalian cell culture for mAb production (using parallel bioreactors) [47]. The authors referred to this approach as "fermentanomics," and it included monitoring of all nutrients (except cysteine) in the representative culture via HPLC and NMR [48]. However, as we have pointed out later in the text, like IR spectroscopy, the utility of NMR as robust and reliable bioprocess monitoring tool hinges on the data analysis using multivariate modeling.

As is evident from the list of tools presented above, infrared-based measurements have gained prominence because of their nondestructive nature and the feasibility of fabricating in situ autoclavable probes (FTIR is an exception) that can be directly inserted into the bioreactor [43]. However, it should be noted that mere IR (near, mid-range, or far) spectra measurements do not yield any meaningful information unless mathematical deconvolution is performed to extract the necessary sample details [49, 50]. Specifically, two kinds of information are sought from IR measurements. One is the determination of the concentration of a particular component in the culture broth and the other is the characterization of overall behavior of the process. For process characterization, multivariate data analysis (MVDA) based on reduction in data dimensionality by the use of projection using latent variables is increasingly being employed [51–55]. The increasing popularity of MVDA stems from its ability to effectively deal with the complexities arising from the multicollinearity, missing data, and the variability emanating from experimental error and noise, which are otherwise difficult to handle [53–56]. A survey of the literature reveals a slew of MVDA applications involving the use of advanced mathematical and statistical algorithms for multivariate statistical process monitoring. These include processing of pharmaceutical granules [57], prediction of product attributes [53–55], quantitative assessment of process comparability [51, 53–55,

58], root cause analysis [58], raw material characterization [59], and prediction of column integrity and impurity clearance during reuse of chromatography resin [60].

3.4 Recent Advancements

Another interesting concept that has come to fore in upstream processing is that of mini QbD assessment [61]. Popularly known as Molecular Design Intent (MDI), the program involves the use of an amalgam of what can be referred to as smart technologies encompassing in silico methods/bioinformatics, micro-bioreactors and/or disposable bioreactors, and protein engineering to obtain a complete fingerprint of the promising leads before proceeding to the bioprocess scale-up. The outcome from such an extensive assessment is typically in the form of ranking of the tested candidates on a scale representing the relative risk of manufacturing as low, moderate, and high. Candidates with low risk of manufacturing are then taken up further for scale-up. Again instead of using traditional approach of using shake flask and bench top bioreactors for optimization of process parameters during clone selection and media formulation, bioprocess scientists have started using disposable versions of these platforms to speed up the entire *process* via automation and parallelization [62].

In addition to this, there is a recent upsurge in the development and use of fully automated high-throughput bioreactors that mimic production scale bioreactors [63–66]. The development of PAT tools for these high-throughput bioreactors is of enormous interest presently [29, 30]. Multi-PAT tools are being eyed upon both by industry as well as academic researchers [67]. In the field of process control, adaptive controls (e.g., decoupled controllers) are gaining popularity due to increased use of robust PAT tools to determine process components (CPPs and state variables) [34].

Despite the aforesaid advancements, the use of these monitoring and control techniques as PAT tools in upstream processing has been rather restricted. One of the biggest factors is rigorous calibration requirements due to the transient nature of cell culture and continuously changing data sets generated upon implementation of these sophisticated instrumentation [36]. This is a nontrivial task as accurate calibration requires a large data set generated by a large number of experiments. In such a scenario, the decision-making time is often significantly less than the actual analysis time, and this results in batch failure much before corrective controls are applied/implemented [28]. Another constraint is with regard to the development of analytical tools and standardization of the method to determine CQA [8]. The correlations between the CQA and CPP are also not entirely known, and hence further advancements in the development of data visualization and real-time decision-making methods in upstream bioprocessing are the need of the hour.

4 QbD in Downstream Processing

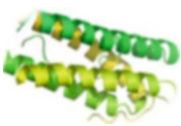
After the desired product has been successfully produced in the production bioreactor, the process stream undergoes clarification (harvest) followed by a series of purification steps that are chosen to purify the protein from a variety of process impurities, host-cell impurities, and product-related variants and impurities [31]. The following discussion aims to summarize QbD applications in key downstream processing unit operations, namely, protein refolding, process chromatography, and viral filtration.

4.1 Protein Refolding

Most biotech products are either of microbial or mammalian origin. Those that are expressed in microbial systems are comparatively simpler in their structure and are generally produced as inclusion bodies (inactive, insoluble, and amorphous aggregate form). On the other hand, the mammalian cell culture-derived products are typically complex structurally as well as functionally and are generally secreted out of the production cells in active form (Table 3).

Microbial products, if expressed in hosts such as *E. coli*, require a refolding step before further processing in order to attain the native functional state [68]. The biggest concern while performing protein refolding is that protein is susceptible to develop both subtle and catastrophic changes if the optimal refolding time is exceeded (oxidation, aggregation, reduction, etc.). Further, in vitro refolding of proteins is also influenced by physical parameters like pH, ionic strength, redox potential, and temperature [69, 70]. With so many process parameters, one has to evaluate not only their main effects on protein refolding but also their interactions.

Table 3 Comparison of key characteristics of microbial and mammalian biotherapeutic products

Feature	Microbial products (e.g., GCSF, insulin) 	Mammalian products (e.g., monoclonal antibodies) 
Molecular weight	Small (5–20 kDa)	High (~150 kDa)
Expression yield	High (1–5%)	Very low (<1%)
Protein folding	Required if expressed in *E. coli*	Not required typically
Glycosylation	None in *E. coli*, limited in *Pichia*	Yes
CQA	Few (10–20)	Many (~20–50)

Though not all product variants adversely impact safety and efficacy of the product, any such heterogeneities should either be removed in the subsequent downstream processing steps or if retained need to be justified by appropriate scientific data explaining their impact in the clinic [71–73]. This adds to both the cost and product development timelines and as a result would deprive patients or society at large from the likely benefits (affordable, safe, and efficacious drug) that could have been accrued from the otherwise successful development and commercialization of biotherapeutics [74].

QbD and PAT concepts have been used to overcome these challenges [69]. The approach involves performing DOE studies to evaluate the main effects and interactions between the process parameters with regard to refold recovery and product quality [75, 76]. Risk-based analysis [69], DOE [76], and high-throughput refolding [62] are some of the approaches that have facilitated QbD implementation in protein refolding.

4.2 Process Chromatography

Process chromatography has long been considered the workhorse of therapeutic protein purification [77] as it allows separation of the component of interest from a mixture by exploiting difference in properties such as charge [78], size [79], affinity [80], hydrophobicity [81], and/or a mixture of both. However, performing chromatographic purification process development is nontrivial due to the several associated complexities arising from complex and poorly characterized raw materials and feed materials, low feed concentration, product instability, and incomplete mechanistic understanding of chromatographic steps [82]. On the other hand, the implementation of QbD requires an improved fundamental understanding of the process especially of the protein and resin matrix interactions for chromatographic operations [83].

Traditional way of process development for chromatography step involves series of steps performed in hierarchical manner: resin screening, process parameter optimization, and process characterization studies [62]. Information from these studies is then used to define the design space for the process. As in case of protein refolding, several operational parameters may interact among themselves and affect the overall performance. Therefore, the traditional method of process development where few parameters are chosen for investigation based on the user's prior experience is likely to result in picking the local optima and not the global one.

QbD-based implementation would involve a DOE-based approach that examines the main effects and interactions between the various process parameters, feed material attributes, and raw material attributes. Examples of QbD and PAT implementation in process chromatography are listed here:

- Estimation of appropriate sample loading based on the % breakthrough set point via the use of at-line stopped-flow analyzers [84]

- Use of multiwavelength UV spectroscopy for real-time monitoring of protein mixtures during purification [85]
- Monitoring of the chromatographic eluent for the presence of any product-related impurities (deamidated isoforms) which if detected could be eliminated by applying appropriate corrective measures in real time [86]
- Performing pooling of process chromatography columns based on feed-forward control for a high-volume product [87]
- Deployment of online HPLC [88], online UPLC [89], or at-line tryptophan fluorescence [22] for creating a PAT-based pooling of process chromatography fractions

A major challenge in process chromatography is the relatively small window of decision making [28]. A typical elution in process chromatography takes 30–60 min, and within this the window of decision making for where to start or end pooling could be less than a minute. This makes it a challenge to design an appropriate PAT control scheme as high-resolution product analysis tools such as HPLC take significantly longer to analyze. However, latest advancements seem to suggest a way out of this conundrum [89]. The fact that some of the major biotech companies are actually practicing these solutions in their manufacturing plants indicates the significance and maturity of these advancements [87, 88].

4.3 Membrane Applications

One of the prime objectives of downstream processing is to achieve certain level of product quality and desired formulation conditions apart from the removal of process and product impurities [29, 30]. Membrane processes such as ultrafiltration (UF), diafiltration (DF), microfiltration (MF), and nanofiltration (NF) are reported to be extensively used for this purpose. The ubiquity of these steps in biotech processes is due to their ease of operation, robustness toward normal variations in feed materials, and operating parameters and lower capital cost requirements versus other processing options [90].

MF is employed for the separation of fine particles in the size range of 0.1–10 μm, while UF/DF membranes typically retain proteins and other macromolecules (pore size 1–100 nm). While the former is widely used for clarification during harvest, the latter is commonly used for concentrating protein solutions and buffer exchange. Diafiltration in UF/DF is commonly performed for a fixed number of diavolumes. The number of diavolumes is based on process development studies and a safety factor. However, the method suffers from a major drawback of:

- Over-usage of expensive buffers
- Stability issues of the product due to extended filtration time

Recent publication describing PAT implementation for this unit operation suggests a possible solution for the problem [91]. The paper describes a process control

strategy based on monitoring concentration of species in the retentate (via pH or conductivity probes) as a function of number of diavolumes. The reported strategy allowed making real-time decision as to at what point in time the UF-DF step was to be stopped. According to the authors, the decision to end the process is made when the salt and buffer concentration becomes equal in both retentate and buffer feed. However, reliability of the probes used for making measurements in such applications is a crucial factor in achieving the desired goal.

4.4 Viral Filtration

Viral filtration is an important step of biotherapeutic development. This is especially applicable to products derived using mammalian expression system (such as mAbs) as they are more prone to contamination with adventitious viruses and endogenous retroviral particles [92]. Thus, for successful commercialization of these biotherapeutic products, a comprehensive viral clearance assessment is a regulatory need. Traditionally, this is performed using viral filters (very costly) in combination with other orthogonal methods such as pH inactivation and chromatography with the pre-validated data from the manufacturer of these filters obtained during process validation with some model viruses [83]. Filtration characteristics such as the extent of fouling/flux decay are considered as critical process parameters and are used as a measure of the extent of viral clearance [53–55].

A major issue with this step is the cost of the filters as they are single use. This is worsened by the practice of oversizing the filters in order to account for the variability in the feed stream characteristics. It is not a surprise that, after chromatography resin, virus filters are the second major contributor to the cost of downstream processing. In this respect, the use of mechanistic models for optimal sizing of virus filters has been suggested [53–55]. It is hoped that as our overall process understanding evolves and deepens, further improvements can be made for optimal design of this step.

4.5 Recent Advancements

4.5.1 Continuous Processing

A major advancement that the biotech industry is likely to see in near future is the adoption of continuous processing (vs. batch processing) for production of protein therapeutics [16, 17, 93, 94]. Despite the complexity of biotech processes and products and the resulting technical challenges that one needs to overcome for successful implementation of continuous processing for production of biotherapeutics, regulatory authorities (including USFDA) have been very supportive of its adoption. However, widespread implementation will take time as

development of robust control systems that can tackle and set in corrective actions in real time catches up with the need.

Integration of multiple steps has been offered as a step forward in this direction. A recent publication has demonstrated robust removal of mAb aggregates by combining ion exchange chromatography and hydrophobic interaction chromatography in series [19]. This two-stage purification scheme resulted in the baseline separation for monomers as well as aggregates, thereby improving the overall yield apart from better clearance of host-cell impurities.

4.5.2 Membrane Chromatography

An alternative to the traditional packed-bed-based process chromatography that has recently emerged is that of membrane chromatography [95]. The rising interest in this alternative comes from the superior mass transfer characteristics that membranes offer due to minimal diffusion-related limitations versus the traditional resin-based separation modules [90, 96, 97]. In addition to this, reduced buffer requirements, lowered pressure drops, ease of manufacturing, and reusability are some of the characteristics of membrane chromatography that has significantly contributed toward increased bioprocessing efficiencies [98]. The most successful use of membrane chromatography is its use in flow-through mode for the removal of process-related contaminants like host-cell proteins, endotoxins, and viruses from the large volume of final processing streams (polishing). The use of membrane chromatography as a capture step is also being explored though the major challenge is the low-binding capacity in comparison to the conventional packed-bed chromatography [99].

Recent major advancements in membrane chromatography involve newer membrane prototypes which have been designed for bind and elute processes that exhibit higher dynamic capacity for large molecules [100], higher selectivity against process-related contaminants [101–103], and higher recovery yields [104]. Hydrophobic interaction chromatography (HIC) membranes for purification of mAbs have been developed with binding capacities comparable to the conventional HIC resins [105].

It can be concluded that membrane chromatography will play a larger role in protein purification in the times to come.

4.5.3 High-Throughput Process Development (HTPD)

Another development that has increasingly important implication for the implementation of QbD principles in bioprocessing is that of high-throughput process development. The primary motivation of using HTPD platforms is to gain understanding of the impact that the combination of process variables exerts on the process outcome in shorter time with relatively much lesser resources [62]. This objective is further facilitated by the amenability of HTPD tools to be fabricated as

miniaturized (handling samples on microliter/nanoliter scale) and automated (robotic handling of process fluids) platforms.

HTPD platforms have been predominantly developed for resin screening of various chromatographic modes [62, 106–109] for determining binding/eluting conditions and adsorption isotherms; selection of optimum membrane type, transmembrane pressure, and cross flow rates for membrane filtration unit operations [62]; studying viral clearance by anion exchange chromatography [110]; evaluation of protein solubility [111]; membrane filtration of plasmid DNA [112]; determination of appropriate conditions for protein precipitation [113]; and determination of optimum formulation conditions [114].

While the development of HTPD platforms for bioprocessing applications continues to evolve at a rapid pace, more needs to happen to facilitate the required analysis of the samples so that the full potential of HTPD can be realized. Also, a clearer understanding of how the HTPD platform will be used for developing QbD understanding of the industrial bioprocess is missing presently, and a clearer framework needs to be developed.

Figure 5 illustrates some of the other major recent developments in downstream bioprocessing [115]:

- Use of multimodal materials for antibody purification and removal of process-related impurities [116]
- Employment of combinatorial ligands for improved affinity separations and replacing protein A capture step [117]
- Use of disposable formats both in chromatographic as well as non-chromatographic applications for maintaining process consistency [118]

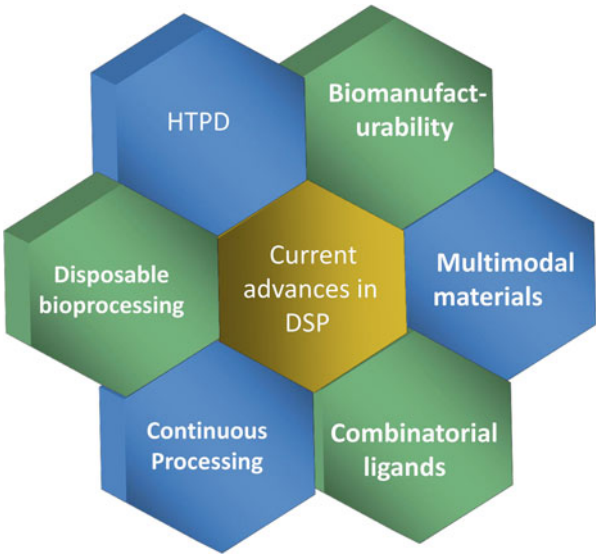


Fig. 5 Recent advances in downstream processing (DSP) of protein therapeutics

- Designing protein molecules with tailored stability and aggregation characteristics using computational tools [119–121]

5 QbD-Based Analytical Characterization and Evaluation of Safety and Efficacy of Biotherapeutic Products

Process and product characterizations are essential components for successful development and commercialization of biotherapeutics. Analytical support in form of various high-end, sophisticated, high-resolution instruments that can derive a detailed fingerprint of the target product allows for performing such characterization and monitoring the product quality throughout the product life cycle [122]. This is particularly relevant to biosimilars as demonstrating analytical comparability with the innovator drug is the foundational first step for receiving regulatory leniency with respect to the extensiveness of the clinical trials. Table 4 summarizes some of the quality attributes that have been reported in the literature for affecting safety and efficacy of a biotherapeutic and the analytics that are commonly employed for characterization.

Table 4 List of quality attributes of biotherapeutics and analytical tool used for their characterization

Characteristic	Quality attribute	Analytical tool
Primary structure	Amino acid sequence	RP-HPLC, LC-ESI-MS, peptide mapping
Higher-order structure	Disulfide bond modifications/free thiols, conformation	LC-ESI-MS, FTIR, CD, X-ray crystallography
Purity	Aggregation	SEC-HPLC, SEC-MALS, AUC, CE-SDS, DLS
Charge heterogeneity/ amino acid modification	Charged isoforms	IEF, IEC-HPLC
	Deamidation/oxidation/C- and N-terminal modifications	LC-MS peptide mapping
Glycosylation	*N*-glycans, *O*-Glycans, fucosylation, galactosylation, and oligomannose forms, sialylation	LC-MS, CE-SDS, HPLC, ESI-Q-TOF-MS
Host-cell impurities	Residual host-cell protein, residual DNA	ELISA, 2D-DIGE, LC-MS, CE-MS

Information for the table adapted from [8]
HPLC high performance liquid chromatography, *ESI-MS* electrospray ionization mass spectrometry, *RP* reversed phase, *CD* circular dichroism, *SEC* size exclusion chromatography, *IEF* isoelectric focusing, *IEC* ion exchange chromatography, *AUC* analytical ultracentrifugation, *CE-SDS* capillary electrophoresis-sodium dodecyl sulfate, *DLS* dynamic light scattering, *ELISA* enzyme-linked immunosorbent assay, *CE-MS* capillary electrophoresis mass spectrometry

Table 5 Summary of impact of CQA of biotherapeutics on product safety and efficacy from scientific literature

Sl. no.	Attributes	Biological activity	CDC	ADCC	Half-life	Safety	Immunogenicity
1	Aggregation	+/−				−	−/0
2	Fragmentation	−	−	−	−		0
3	C-/N-terminal truncation	0	0	0	0	0	−/0
4	Oxidation	−/0			−	−	−
5	Deamidation	−/0			0		−/0/+
6	Glycosylation	0/+	+	+	0/+		0
7	Glycation	−/0/+			−	−	0
8	Conformation						−
9	DNA					−	
10	HCP						−
11	Raw material impurities					−	
12	Adventitious agents, endotoxins					−	

Adapted with copyright permission from [123]
+ Positive impact, − negative impact, 0 no impact, blank unreported/unknown
CDC complement-dependent cytotoxicity, *ADCC* antibody-dependent cellular cytotoxicity, *HCP* host-cell proteins

Recently an excellent review was published on the topic of how the various attributes relate to the product's clinical safety and efficacy [123]. This information is summarized in Table 5. The authors reviewed the existing literature and concluded that our understanding of how product quality attributes impact safety and efficacy of the product in clinic is not only incomplete but also inaccurate in many cases. For example, protein aggregation is generally considered as a CQA due to its possible role in eliciting adverse immune response in the patients. But as is evident from Table 5, aggregation can also enhance biological activity without any immunogenicity. Of late, it has been proved using a combination of analytical characterization and nonclinical and clinical studies that protein aggregates per se do not pose immunogenicity threats, and the chances of such adverse response is a function of type of aggregates, inherent immunogenicity of the monomer, and the number and size of aggregate particles [124].

There are only a few quality attributes for which there is universal recognition with regard to their impact on safety and efficacy. Thus, successful QbD implementation requires a careful, tedious approach of isolating the concerned product variant or impurity and then testing them on a case-by-case basis. However, ethical and other logistical considerations prevent us from performing in vitro and in vivo experiments to derive such relationships [123]. In this scenario, we see great potential of developing newer platforms that have less reliance on clinical/animal trials. Genomics and proteomic platforms together with bioinformatics would be highly instrumental in achieving this goal.

6 Chemometrics/Multivariate Data Analysis

With the complexities that are inherent to bioprocessing, multivariate data analysis (MVDA) of the bioprocess data has gradually been accepted across industry as an approach for enhancing our process and product understanding. Today, every major biotech company is engaged in this practice.

Multivariate data analysis uses projection techniques such as principal component analysis (PCA) and partial least squares (PLS) to extract useful information from data sets. Both PCA and PLS act by reducing the dimensionality of the original data set into lower number of uncorrelated variables (principle components/latent variables) that explains most of the variance contained in the data set [53–56].

Among many others, some of the common applications of MVDA in bioprocessing include:

- Modeling of protein A column decay utilizing multiple transitions in chromatogram by PCA [125]
- Development of scale-down bioreactor model for characterization and validation of cell culture process [126]
- Analysis of data originating from spectroscopic measurements [56]
- Analysis of data profiles from unit operations such as cell culture [52, 127] and chromatography [60]
- Quantitative assessment of process comparability [53–55]
- Root cause analysis [53–55, 58]
- Raw material characterization [59]

Figure 6 illustrates the facilitating role of MVDA in QbD and PAT implementation in bioprocess development [52].

7 QbD for Biosimilars

The rising demand for biosimilars are in part due to the attempts to contain the increasing cost of healthcare and offer solutions for complex diseases for which existing treatments are inadequate or not available. However, the path for the development of a biosimilar product is not as simple and straightforward as it is for the production of a small-molecule generic drug. Factors that contribute to this complexity include increased complexity of production processes and the product itself, structural heterogeneity inherent in biotech products, lack of complete analytical characterization, sensitivity of product quality toward seemingly small changes in the manufacturing process, significant impact that variability in raw material quality can have on product quality, and last, but perhaps the most important, our limited understanding of how the various critical quality attributes of a biotech therapeutic affect its clinical safety and efficacy [5, 8, 74, 128–131].

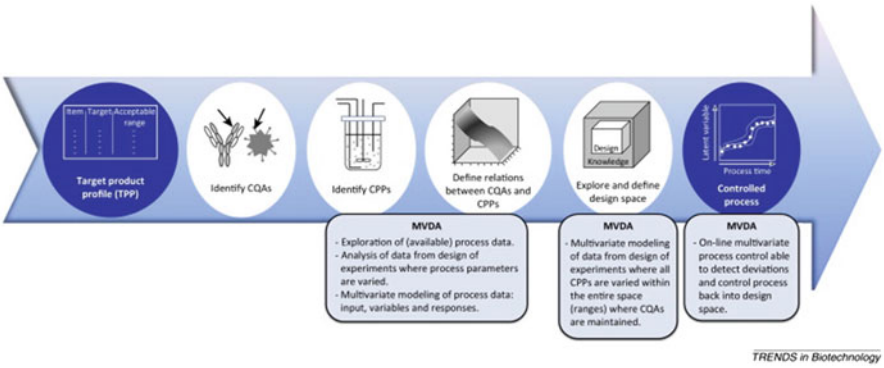


Fig. 6 Typical approach for QbD and PAT implementation in bioprocess development and role of MVDA in this process. Republished with copyright permission from [52]

With all of the above-stated issues that come with manufacturing and commercialization of biosimilars, the implementation of QbD in biosimilar manufacturing seems to be the long-term solution for the abovementioned challenges. It is expected that the implementation of QbD will enhance our understanding of how the starting materials impact the performance of the process, identify the critical process parameters (CPP) of the process that impact the quality attributes of the product, and identify the critical quality attributes of the product that impact safety and efficacy in the clinic. This in return would lead to significant reduction in clinical data required for achieving regulatory approval for biosimilars.

Despite this, the topic of approval of biosimilars will stay controversial till a clear and uniform regulatory framework is put in place across the globe with parallel developments in the analytical front for the characterization of these complex entities.

8 Future Perspective

The aim of QbD is really to be able to understand the product and the process so well that we can predict the impact of defined as well as unknown sources of variability on product quality and thereby help in significantly improving process consistency using an enhanced process knowledge-driven approach. As the biopharma industry continues to embrace the idea of in-line, online, and/or at-line sensors and real-time characterization of physiological state of the cells for process monitoring and control, the existing gaps with regard to our ability of monitoring multiple parameters/variables associated with the upstream process will be alleviated over time. These efforts can be further accelerated with simultaneous advances in the data processing and visualization, development of mechanistic models for bioprocess unit operations, and incorporation of statistical approaches such as DOE

[75] and multivariate data analysis in scale-down experiments [132]. Also, with the experience that has been gained with about a decade of QbD implementation for biologic manufacturing, it is evident that said limitations, though formidable, are surmountable.

The major benefit of following the QbD approach involving enhanced process and product understanding, sophisticated monitoring and control strategies, risk-based assessment for identifying CPP and CQA, and finally integration of mechanistic and multivariate models is that a design space can be defined within which changes in operating conditions have been demonstrated to cause minimal change in product quality [15]. This in turn translates into increased manufacturing flexibility as minor variations in operating conditions, feed material attributes, and raw material attributes (within the design space) would not require as much regulatory scrutiny as those changes that are outside the design space.

Table 6 lists out key differences between traditional development of bioprocesses and QbD-based process development.

We hope to convince those working on development of biotech processes and products that the implementation of QbD will result in benefits for the manufacturers (more flexibility, less maintenance during commercial manufacturing, fewer lot failures), regulators (improved product and process understanding, better consistency in product quality, better risk management), and most of all the patients (more consistent product quality, better product understanding).

Table 6 Key differences between traditional process development and QbD-based process development

Feature	Traditional approach	QbD approach
Approach	Process defines product	Product defines process
Product specification	Set based on process performance	Set based on process and product knowledge
Control strategy	Process is rigid, control meant to run process in narrow operating range	Process is flexible, controls can adapt to incoming variability in raw materials or process or facility
Product quality assessment	Performed retrospectively	Performed in real time
Process validation (PV)	All parameters are included in PV protocol	Only critical parameters identified based on risk assessment are included in PV protocol
Decision making	Function based, more delays and recycles	Team based, few delays and recycles
Filings	Describes process with less focus on product and process knowledge	Focus on product and process knowledge with and establish relationship between quality attribute with safety and efficacy
Post-approvals	High regulatory burden with little scope of process improvements post-approval	Low regulatory burden with quicker noncompliance resolution due to already existing process knowledge

References

1. Weinacker D et al (2013) Applications of recombinant *Pichia pastoris* in the healthcare industry. Braz J Microbiol 44(4):1043–1048
2. Walsh G (2014) Biopharmaceutical benchmarks 2014. Nat Biotechnol 32(10):992–1000
3. Rathore AS (2009) Roadmap for implementation of quality by design (QbD) for biotechnology products. Trends Biotechnol 27(9):546–553
4. Rathore AS, Winkle H (2009) Quality by design for biopharmaceuticals. Nat Biotechnol 27(1):26–34
5. Tsuruta LR, Lopes dos Santos M, Moro AM (2015) Biosimilars advancements: moving on to the future. Biotechnol Prog 31(5):1139–1149
6. Schellekens H (2004) How similar do 'biosimilars' need to be? Nat Biotechnol 22(11): 1357–1359
7. CHMP (2005) Guideline on similar biological medicinal products, European Medicines Agency, 437, London
8. Rathore AS (2009) Follow-on protein products: scientific issues, developments and challenges. Trends Biotechnol 27(12):698–705
9. Cai X, Wake A, Gouty D (2013) Analytical and bioanalytical assay challenges to support comparability studies for biosimilar drug development. Bioanalysis 5(5):517–520
10. Beck A et al (2013) Analytical characterization of biosimilar antibodies and Fc-fusion proteins. TrAC Trends Anal Chem 48:81–95
11. Chelius D et al (2010) Structural and functional characterization of the trifunctional antibody catumaxomab. MAbs 2(3):309–319
12. Rosati S et al (2013) Tackling the increasing complexity of therapeutic monoclonal antibodies with mass spectrometry. TrAC Trends Anal Chem 48:72–80
13. Singleton CA (2014) MS in the analysis of biosimilars. Bioanalysis 6(12):1627–1637
14. Xie H et al (2010) Rapid comparison of a candidate biosimilar to an innovator monoclonal antibody with advanced liquid chromatography and mass spectrometry technologies. MAbs 2(4):379–394
15. Rathore AS, Mhatre R (2008) In: Rathore AS, Mhatre R (eds) Quality by design for biopharmaceuticals: principles and case studies. Wiley, Hoboken
16. Rathore AS et al (2015) Continuous processing for production of biopharmaceuticals. Prep Biochem Biotechnol 45(8):836–849
17. Rathore AS et al (2015) Process analytical technologies in biopharmaceutical process development. J Chem Technol Biotechnol 90(2):213–214
18. Hemmerich J et al (2014) Comprehensive clone screening and evaluation of fed-batch strategies in a microbioreactor and lab scale stirred tank bioreactor system: application on *Pichia pastoris* producing *Rhizopus oryzae* lipase. Microb Cell Fact 13(1):36
19. Kumar V, Rathore AS (2014) Two-stage chromatographic separation of aggregates for monoclonal antibody therapeutics. J Chromatogr A 1368:155–162
20. Legmann R et al (2009) A predictive high-throughput scale-down model of monoclonal antibody production in CHO cells. Biotechnol Bioeng 104(6):1107–1120
21. Persad A et al (2013) Comparative performance of decoupled input–output linearizing controller and linear interpolation PID controller: enhancing biomass and ethanol production in *Saccharomyces cerevisiae*. Appl Biochem Biotechnol 169(4):1219–1240
22. Rathore AS et al (2009) Case study and application of process analytical technology (PAT) towards bioprocessing: use of tryptophan fluorescence as at-line tool for making pooling decisions for process chromatography. Biotechnol Prog 25(5):1433–1439
23. Roychoudhury P et al (2007) Multiplexing fibre optic near infrared (NIR) spectroscopy as an emerging technology to monitor industrial bioprocesses. Anal Chim Acta 590(1):110–117
24. ICH Guidelines Q 11, 20AD. Pharmaceutical development Q8. ICH Harmonised Tripartite Guideline, 8(August), pp 1–28

25. Rathore AS (2014) QbD/PAT for bioprocessing: moving from theory to implementation. Curr Opin Chem Eng 6:1–8
26. ICH Guideline (2012) Development and manufacture of drug substances (chemical entities and biotechnological/biological entities) Q11, European medicines agency, May 2011
27. ICH Guideline (2008) Pharmaceutical quality system Q10, current step, 4
28. Rathore AS, Bhambure R, Ghare V (2010) Process analytical technology (PAT) for biopharmaceutical products. Anal Bioanal Chem 398(1):137–154
29. Read EK, Park JT et al (2010) Process analytical technology (PAT) for biopharmaceutical products: part I. Concepts and applications. Biotechnol Bioeng 105(2):276–284
30. Read EK, Shah RB et al (2010) Process analytical technology (PAT) for biopharmaceutical products: part II. Concepts and applications. Biotechnol Bioeng 105(2):285–295
31. Gronemeyer P, Ditz R, Strube J (2014) Trends in upstream and downstream process development for antibody manufacturing. Bioengineering 1(4):188–212
32. Haigney S (2013) QbD and PAT in upstream and downstream processing. BioPharm Int 26(7):28–37
33. ICH Expert Working Group (2005) Quality risk management Q9. ICH Harmonised Tripartite Guideline (November), pp 1–23
34. Gomes J, Chopda VR, Rathore AS (2015) Integrating systems analysis and control for implementing process analytical technology in bioprocess development. J Chem Technol Biotechnol 90(4):583–589
35. Raju TS (2003) Glycosylation variations with expression systems and their impact on biological activity of therapeutic immunoglobulins. Bioprocess Int 1:44–53
36. Biechele P et al (2015) Sensor systems for bioprocess monitoring. Eng Life Sci 15(5): 469–488
37. Jain E, Kumar A (2008) Upstream processes in antibody production: evaluation of critical parameters. Biotechnol Adv 26(1):46–72
38. Rouiller Y et al (2012) Application of quality by design to the characterization of the cell culture process of an Fc-fusion protein. Eur J Pharm Biopharm 81(2):426–437
39. Goldfeld M et al (2014) Advanced near-infrared monitor for stable real-time measurement and control of *Pichia pastoris* bioprocesses. Biotechnol Prog 30(3):749–759
40. Sandor M et al (2013) NIR-spectroscopy for bioprocess monitoring & control. BMC Proc 7(Suppl 6):P29
41. Roychoudhury P, Harvey LM, McNeil B (2006) The potential of mid infrared spectroscopy (MIRS) for real time bioprocess monitoring. Anal Chim Acta 571(2):159–166
42. Schmidt-Hager J et al (2014) Noninvasive online biomass detector system for cultivation in shake flasks. Eng Life Sci 14(5):467–476
43. Zhao L et al (2015) Advances in process monitoring tools for cell culture bioprocesses. Eng Life Sci 15(5):459–468
44. Craven S, Whelan J (2015) Process analytical technology and quality-by-design for animal cell culture. In: Al-Rubeai M (ed) Animal cell culture, Cell engineering. Springer, Switzerland, pp 647–688
45. Schwamb S, Puskeiler R, Wiedemann P (2015) Monitoring of cell culture. In: Al-Rubeai M (ed) Animal cell culture, Cell engineering. Springer, Switzerland, pp 185–221
46. Kuystermans D, Mohd A, Al-Rubeai M (2012) Automated flow cytometry for monitoring CHO cell cultures. Methods 56(3):358–365
47. Bradley SA et al (2010) Fermentanomics: monitoring mammalian cell cultures with NMR spectroscopy. J Am Chem Soc 132(28):9531–9533
48. Read EK et al (2013) Fermentanomics informed amino acid supplementation of an antibody producing mammalian cell culture. Biotechnol Prog 29(3):745–753
49. Clavaud M et al (2013) Chemometrics and in-line near infrared spectroscopic monitoring of a biopharmaceutical Chinese hamster ovary cell culture: prediction of multiple cultivation variables. Talanta 111:28–38

50. Lourenço ND et al (2012) Bioreactor monitoring with spectroscopy and chemometrics: a review. Anal Bioanal Chem 404(4):1211–1237
51. Bhushan N, Hadpe S, Rathore AS (2012) Chemometrics applications in biotech processes: assessing process comparability. Biotechnol Prog 28(1):121–128
52. Mercier SM et al (2014) Multivariate PAT solutions for biopharmaceutical cultivation: current progress and limitations. Trends Biotechnol 32(6):329–336
53. Rathore AS et al (2014) Chemometrics application in biotech processes: assessing comparability across processes and scales. J Chem Technol Biotechnol 89(9):1311–1316
54. Rathore AS, Mittal S, Pathak M, Arora A (2014) Guidance for performing multivariate data analysis of bioprocessing data: pitfalls and recommendations. Biotechnol Prog 30(4): 967–973
55. Rathore AS, Kumar V et al (2014) Mechanistic modeling of viral filtration. J Membr Sci 458: 96–103
56. Rathore AS, Bhushan N, Hadpe S (2011) Chemometrics applications in biotech processes: a review. Biotechnol Prog 27(2):307–315
57. Vaidyanathan S et al (2001) Assessment of near-infrared spectral information for rapid monitoring of bioprocess quality. Biotechnol Bioeng 74(5):376–388
58. Kirdar AO, Green KD, Rathore AS (2008) Application of multivariate data analysis for identification and successful resolution of a root cause for a bioprocessing application. Biotechnol Prog 24(3):720–726
59. Kirdar AO et al (2010) Application of near-infrared (NIR) spectroscopy for screening of raw materials used in the cell culture medium for the production of a recombinant therapeutic protein. Biotechnol Prog 26(2):527–531
60. Rathore AS et al (2012) Chemometrics applications in biotechnology processes: predicting column integrity and impurity clearance during reuse of chromatography resin. Biotechnol Prog 28(5):1308–1314
61. Pearson S (2015) Smart process development strategies. Genet Eng Biotechnol News, pp 1–2
62. Bhambure R, Kumar K, Rathore AS (2011) High-throughput process development for biopharmaceutical drug substances. Trends Biotechnol 29(3):127–135
63. Lattermann C, Büchs J (2014) Microscale and miniscale fermentation and screening. Curr Opin Biotechnol 35C:1–6
64. Pollard D (2014) Are automated disposable small-scale reactors set to dominate the future of pharmaceutical bioprocess development? Pharm Bioprocess 2(1):9–12
65. Rameez S et al (2014) High-throughput miniaturized bioreactors for cell culture process development: reproducibility, scalability, and control. Biotechnol Prog 30(3):718–727
66. Razinkov V et al (2015) Automation and high-throughput technologies in biopharmaceutical drug product development with QbD approaches. In: Jameel F et al (eds) Quality by design for biopharmaceutical drug product development SE - 20, AAPS advances in the pharmaceutical sciences series. Springer, New York, pp 475–510
67. Simon LL et al (2015) Assessment of recent process analytical technology (PAT) trends: a multiauthor review. Org Process Res Dev 19(1):3–62
68. Clark EDB (2001) Protein refolding for industrial processes. Curr Opin Biotechnol 12(2): 202–207
69. Bade PD, Kotu SP, Rathore AS (2012) Optimization of a refolding step for a therapeutic fusion protein in the quality by design (QbD) paradigm. J Sep Sci 35(22):3160–3169
70. Rathore AS et al (2013) Refolding of biotech therapeutic proteins expressed in bacteria: review. J Chem Technol Biotechnol 88(10):1794–1806
71. Calvo B, Zuñiga L (2012) The US approach to biosimilars. BioDrugs 26(6):357–361
72. O'Connor A, Rogge M (2013) Nonclinical development of a biosimilar: the current landscape. Bioanalysis 5(5):537–544
73. Tsiftsoglou A et al (2014) Demonstration of biosimilarity, extrapolation of indications and other challenges related to biosimilars in Europe. BioDrugs 28(6):479–486

74. Rathore AS, Singh SK (2014) Clinical and non-clinical aspects of biosimilar development. BioQuality 9:1–6
75. Kumar V, Bhalla A, Rathore AS (2014) Design of experiments applications in bioprocessing: concepts and approach. Biotechnol Prog 30(1):86–99
76. Saremirad P et al (2014) Multi-variable operational characteristic studies of on-column oxidative protein refolding at high loading concentrations. J Chromatogr A 1359:70–75
77. Przybycien TM, Pujar NS, Steele LM (2004) Alternative bioseparation operations: life beyond packed-bed chromatography. Curr Opin Biotechnol 15(5):469–478
78. Williams A, Frasca V (2001) Ion-exchange chromatography. In: Current protocols in protein science, vol 15(8.2), pp 8.2.1–8.2.30
79. Barth HG, Jackson C, Boyes BE (1994) Size exclusion chromatography. Anal Chem 66(12): 595R–620R
80. Roque ACA, Silva CSO, Taipa MA (2007) Affinity-based methodologies and ligands for antibody purification: advances and perspectives. J Chromatogr A 1160(1–2):44–55
81. Ochoa J-L (1978) Hydrophobic (interaction) chromatography. Biochimie 60(1):1–15
82. Bhambure R, Rathore AS (2013) Chromatography process development in the quality by design paradigm I: establishing a high-throughput process development platform as a tool for estimating 'characterization space' for an ion exchange chromatography step. Biotechnol Prog 29(2):403–414
83. Rathore AS, Kapoor G (2015) Application of process analytical technology for downstream purification of biotherapeutics. J Chem Technol Biotechnol 90(2):228–236
84. Chhatre S et al (2009) Use of PAT principles for the open-loop control of laboratory and pilot-scale chromatography columns. J Chem Technol Biotechnol 84(9):1314–1322
85. Kamga M-H et al (2013) Quantification of protein mixture in chromatographic separation using multi-wavelength UV spectra. Biotechnol Prog 29(3):664–671
86. Vyas V, Nellis D, Burnette A (2009) BIOT-396. In: Abstracts of papers, 238th ACS national meeting, Washington DC
87. Mendhe R et al (2015) Comparison of PAT based approaches for making real-time pooling decisions for process chromatography – use of feed forward control. J Chem Technol Biotechnol 90(2):341–348
88. Kaltenbrunner O et al (2012) Risk-benefit evaluation of on-line high-performance liquid chromatography analysis for pooling decisions in large-scale chromatography. J Chromatogr A 1241:37–45
89. Rathore AS et al (2008) Case study and application of process analytical technology (PAT) towards bioprocessing: II. Use of ultra-performance liquid chromatography (UPLC) for making real-time pooling decisions for process chromatography. Biotechnol Bioeng 101(6):1366–1374
90. Rathore AS, Shirke A (2011) Recent developments in membrane-based separations in biotechnology processes: review. Prep Biochem Biotechnol 41(4):398–421
91. Rathore AS, Sharma A, Chilin D (2006) Applying process analytical technology to biotech unit operations. BioPharm Int 19(8):48–57
92. Shukla AA, Thömmes J (2010) Recent advances in large-scale production of monoclonal antibodies and related proteins. Trends Biotechnol 28(5):253–261
93. Croughan MS, Konstantinov KB, Cooney C (2015) The future of industrial bioprocessing: batch or continuous? Biotechnol Bioeng 112(4):648–651
94. Warikoo V et al (2012) Integrated continuous production of recombinant therapeutic proteins. Biotechnol Bioeng 109(12):3018–3029
95. Van Reis R, Zydney A (2001) Membrane separations in biotechnology. Curr Opin Biotechnol 12(2):208–211
96. Łącki KM (2014) High throughput process development in biomanufacturing. Curr Opin Chem Eng 6:25–32
97. Muthukumar S, Rathore AS (2013) High throughput process development (HTPD) platform for membrane chromatography. J Membr Sci 442:245–253

98. Orr V et al (2013) Recent advances in bioprocessing application of membrane chromatography. Biotechnol Adv 31(4):450–465
99. Fröhlich H et al (2012) Membrane technology in bioprocess science. Chem Ing Tech 84(6): 905–917
100. Fraud N, Kuczewski M, Hirai M (2009) Hydrophobic membrane adsorbers for large-scale downstream processing. BioPharm Int 2009(suppl 22):24–27
101. Limonta M et al (2010) The purification of plasmid DNA for clinical trials using membrane chromatography. BioPharm Int 23(2):46–54
102. Riordan W et al (2009) Design of salt-tolerant membrane adsorbers for viral clearance. Biotechnol Bioeng 103(5):920–929
103. Yu D, Ghosh R (2010) Purification of PEGylated protein using membrane chromatography. J Pharm Sci 99(8):3326–3333
104. Anuraj N et al (2012) An all-aqueous route to polymer brush-modified membranes with remarkable permeabilites and protein capture rates. J Membr Sci 389:117–125
105. Kuczewski M et al (2010) Development of a polishing step using a hydrophobic interaction membrane adsorber with a PER. C6®-derived recombinant antibody. Biotechnol Bioeng 105(2):296–305
106. Bensch M, Schulze Wierling P, von Lieres E, Hubbuch J (2005) High throughput screening of chromatographic phases for rapid process development. Chem Eng Technol 28(11): 1274–1284
107. Coffman JL, Kramarczyk JF, Kelley BD (2008) High-throughput screening of chromatographic separations: I. Method development and column modeling. Biotechnol Bioeng 100(4):605–618
108. Kelley BD, Switzer M, Bastek P, Kramarczyk JF, Molnar K, Yu T, Coffman J (2008) High-throughput screening of chromatographic separations: IV. Ion-exchange. Biotechnol Bioeng 100(5):950–963
109. Kramarczyk JF, Kelley BD, Coffman JL (2008) High-throughput screening of chromatographic separations: II. Hydrophobic interaction. Biotechnol Bioeng 100(4):707–720
110. Strauss DM et al (2010) Strategies for developing design spaces for viral clearance by anion exchange chromatography during monoclonal antibody production. Biotechnol Prog 26(3):750–755
111. Wiendahl M et al (2009) A novel method to evaluate protein solubility using a high throughput screening approach. Chem Eng Sci 64(17):3778–3788
112. Kong S, Aucamp J, Titchener-Hooker NJ (2010) Studies on membrane sterile filtration of plasmid DNA using an automated multiwell technique. J Membr Sci 353(1):144–150
113. Ahmad SS, Dalby PA (2011) Thermodynamic parameters for salt-induced reversible protein precipitation from automated microscale experiments. Biotechnol Bioeng 108(2):322–332
114. Grant Y, Matejtschuk P, Dalby PA (2009) Rapid optimization of protein freeze-drying formulations using ultra scale-down and factorial design of experiment in microplates. Biotechnol Bioeng 104(5):957–964
115. Cramer SM, Holstein MA (2011) Downstream bioprocessing: recent advances and future promise. Curr Opin Chem Eng 1(1):27–37
116. Voitl A, Müller-Späth T, Morbidelli M (2010) Application of mixed mode resins for the purification of antibodies. J Chromatogr A 1217(37):5753–5760
117. Naik AD et al (2011) Performance of hexamer peptide ligands for affinity purification of immunoglobulin G from commercial cell culture media. J Chromatogr A 1218(13): 1691–1700
118. Rao G, Moreira A, Brorson K (2009) Disposable bioprocessing: the future has arrived. Biotechnol Bioeng 102(2):348–356
119. Chennamsetty N et al (2010) Prediction of aggregation prone regions of therapeutic proteins. J Phys Chem B 114(19):6614–6624
120. Jordan JL et al (2009) Structural understanding of stabilization patterns in engineered bispecific Ig-like antibody molecules. Proteins 77(4):832–841

121. Perchiacca JM, Bhattacharya M, Tessier PM (2011) Mutational analysis of domain antibodies reveals aggregation hotspots within and near the complementarity determining regions. Proteins 79(9):2637–2647
122. Berkowitz SA et al (2012) Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars. Nat Rev Drug Discov 11(7):527–540
123. Eon-Duval A, Broly H, Gleixner R (2012) Quality attributes of recombinant therapeutic proteins: an assessment of impact on safety and efficacy as part of a quality by design development approach. Biotechnol Prog 28(3):608–622
124. Joubert MK et al (2012) Highly aggregated antibody therapeutics can enhance the in vitro innate and late-stage T-cell immune responses. J Biol Chem 287(30):25266–25279
125. Hou Y et al (2011) Improved process analytical technology for protein a chromatography using predictive principal component analysis tools. Biotechnol Bioeng 108(1):59–68
126. Tescione L et al (2015) Application of bioreactor design principles and multivariate analysis for development of cell culture scale down models. Biotechnol Bioeng 112(1):84–97
127. Kirdar AO et al (2007) Application of multivariate analysis toward biotech processes: case study of a cell-culture unit operation. Biotechnol Prog 23(1):61–67
128. Barbosa MDFS et al (2012) Biosimilars and biobetters as tools for understanding and mitigating the immunogenicity of biotherapeutics. Drug Discov Today 17(23):1282–1288
129. Kenett RS, Kenett DA (2008) Quality by design applications in biosimilar pharmaceutical products. Accred Qual Assur 13(12):681–690
130. Sampath K (2013) Drug product development of biosimilars: QbD based approach and strategies. In: Abstracts of papers of the American Chemical Society. American Chemical Society, Washington, DC
131. Zalai D, Dietzsch C, Herwig C (2013) Risk-based process development of biosimilars as part of the quality by design paradigm. PDA J Pharm Sci Technol 67:569–580
132. Velayudhan A (2014) Overview of integrated models for bioprocess engineering. Curr Opin Chem Eng 6:83–89

Top Med Chem (2017) 21: 69–122
DOI: 10.1007/7355_2015_5005

Published online: 15 March 2016

Characterization of Therapeutic Proteins

E.B. Struble, N. Kirschbaum, J. Liu, E. Marszal, and M. Shapiro

Abstract Therapeutic proteins are large biological molecules with complex structures and functions produced through complex manufacturing processes, which include multiple unit operations with finely tuned control parameters. The characterization of therapeutic protein products during development, manufacturing and at product release requires the development and qualification of appropriate analytical methods that measure physicochemical properties and biological activities. Analytical testing during product development forms the basis for identifying the critical quality attributes for the protein therapeutic product, establishing release and stability specifications, and developing an analytical comparability program that ensures safety and efficacy throughout the product life cycle. In this chapter we discuss analytical characterization in the context of the regulation of therapeutic proteins. We focus on polyclonal immune globulins, proteins for hemostasis, monoclonal antibodies, and other therapeutic proteins and emphasize the

Disclaimer: The findings and conclusions in this article have not been formally disseminated by the Food and Drug Administration and should not be construed to represent any agency determination or policy.

E.B. Struble (✉), N. Kirschbaum, and E. Marszal
Division of Hematology Research and Review, Office of Blood Research and Review, Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA
e-mail: evi.struble@fda.hhs.gov

J. Liu
Division of Biotechnology Review and Research II, Office of Biotechnology Products, Office of Pharmaceutical Quality, Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA

M. Shapiro
Division of Biotechnology Review and Research I, Office of Biotechnology Products, Office of Pharmaceutical Quality, Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA

commonalities and also highlight differences in the application of laws, regulations, and guidance.

Keywords Analytical characterization, Biologics, Therapeutic proteins

Contents

1 Introduction

Therapeutic proteins are large molecules with complex structure and function. They comprise a sequence of amino acids folded into complex secondary, tertiary, and, sometimes, quaternary structures. They may be further modified either naturally by attachment of other molecular entities such as glycans or by designing biological or nonbiological modifiers such as polyethylene glycol molecules to extend product in vivo half-life or reduce immunogenicity. Furthermore, therapeutic proteins are often composed of a series of structural domains, which may be linked to separate functions directing interactions with any number of other molecules in the environment in order to confer their therapeutic benefit, including the duration of the activity in vivo. It follows, therefore, that molecular characterization of a therapeutic protein requires the development and application of a larger number of suitable analytical methods that measure an extended catalog of physicochemical properties as well as the protein's biological and, in certain instances, immunochemical activities. By comparison, small-molecule drugs often lack this structural and functional complexity (Fig. 1). Their small size and somewhat simpler three-dimensional structures make small-molecule drugs amenable to complete characterization by physicochemical methods, including absolute structural determination by X-ray diffraction.

The manufacture of therapeutic proteins is also complex. Therapeutic proteins may be purified from complex biological starting materials or may be produced in living cells through biotechnology. Starting materials contain the target protein

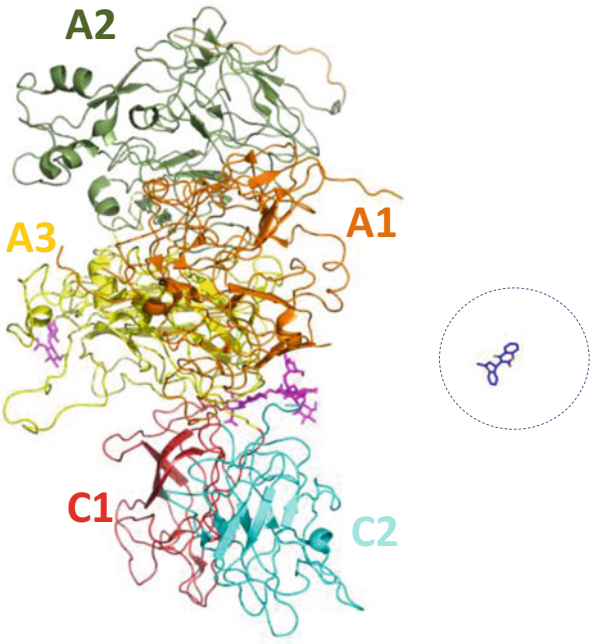


Fig. 1 Three-dimensional structures of the active ingredients in two different drugs that affect coagulation, drawn to scale. Warfarin (Coumadin) is shown in stick representation in *dark blue* and is *circled* for emphasis. Coagulation Factor VIII is shown in cartoon representation with different domains *colored* and *marked*; glycans are shown in stick representation in *magenta*. Created with PyMOL using coordinates from DOI: 10.2210/pdb2bxd/pdb and DOI: 10.2210/pdb2r7e/pdb for Warfarin and Factor VIII, respectively

among a heterogeneous mixture of endogenous proteins, added chemicals or reagents, and potential microbial contaminants. Purification of therapeutic proteins, thus, requires multiple unit operations with finely tuned control parameters. Certain processes such as elution from an affinity resin or viral inactivation expose the protein to harsh conditions. Therefore, analytical testing schemes must be designed to ensure retention of the protein's molecular integrity and native conformation as well as demonstrate the removal of critical impurities throughout the manufacturing process. Finally, since modifications to a commercial manufacturing process are likely to occur throughout a product's lifetime, analytical testing plans must implement methods suitably sensitive to detect any consequent negative impact on product quality or safety.

The regulation of therapeutic proteins as biological products aims to ensure that new protein therapeutics entering the market are proven to be safe and efficacious and consistently manufactured to a high-quality standard and that they remain so for the entire life cycle of the product.

2 Regulatory Framework for Ensuring Quality Throughout the Product Life Cycle

Therapeutic proteins are regulated in the USA as biological products. The regulatory requirement for biological product quality is defined in statute and expanded in the regulations under Title 21 of the Code of Federal Regulations (21 CFR). The Public Health Service Act of 1944 (42 U.S.C.262) requires that biological products introduced into the market be licensed based on a demonstration that the product is safe, pure, and potent. Biological products are also subject to applicable laws in the Federal Food, Drug, and Cosmetic Act, which prohibits the marketing of adulterated drugs, including those not manufactured under current Good Manufacturing Practice (cGMP). Regulations governing cGMP for finished pharmaceuticals are further codified under 21 CFR part 211 (§211). To ensure drug product conformance to a consistent quality standard, cGMP regulations under §211.165 and General Biological Products Standards under §610 require release testing of every lot for specified quality attributes, which must include tests for identity, purity, potency, and sterility. In-process testing is further required to ensure batch uniformity and integrity of drug products (§211.110). Finally, to ensure that product quality attributes remain within appropriate limits throughout the product shelf life, §211.166 requires stability testing, identifying the appropriate storage conditions and establishing product expiration. Regulations under §211.165 also require that analytical methods used for release testing be validated for accuracy, sensitivity, specificity, and reproducibility.

Robust biopharmaceutical development programs that incorporate modern principles of Quality by Design (QbD) promote product quality throughout the product life cycle. QbD is the systematic approach to pharmaceutical development that begins with predefined objectives, emphasizes product and process understanding and control, and is based on sound science and quality risk management. International Conference on Harmonisation (ICH) Q8, Q9, Q10, and Q11 guidelines [1–4] describe enhanced approaches to product and process development based on QbD principles. The definition of a quality target product profile (QTPP) at product inception focuses product characterization and process design on meeting patient needs and clinical performance requirements at each phase of the product life cycle. Integrating quality risk management with comprehensive analytical characterization using state-of-the-art methods allows a confident definition of a product's critical quality attributes upon which the design of the manufacturing process can be based. Knowledge management of analytical data from product and process development is leveraged to define an appropriate control strategy, support process validation at the licensing stage, and enable continuous process verification post-approval.

In this chapter, we will discuss analytical characterization and testing paradigms in the context of the regulation of therapeutic proteins, focusing on recombinant therapeutics, monoclonal antibodies, and plasma-derived proteins. The experience of several regulating divisions will be discussed, emphasizing the commonalities but also highlighting differences in the application of laws, regulations, and guidance based on product class characteristics and regulatory experience. Monoclonal

antibodies and related products, most Fc-fusion proteins, and recombinant protein therapeutics are regulated by the CDER/OPQ/OBP,[1] while plasma-derived proteins and their recombinant analogues are regulated by CBER/OBRR.[2]

3 Analytical Characterization to Support Biopharmaceutical Development and Life Cycle Management

Analytical characterization serves three main purposes in biopharmaceutical development and regulatory approval: (1) definition of critical quality attributes, (2) establishment of release and stability specifications, and (3) development of an analytical comparability program. Analytical testing should be included in any regulatory application, and depending on the phase of development can be extensive. At the time of approval, there should be a history of all lots tested including purpose and summary results. Ideally, analytical characterization will also include results from stress stability studies. General strategies for generating adequate analytical data are described in ICH guidelines such as Q6B (specifications) [5], Q5E (comparability) [6], Q1A-Q1E [7–11], and Q5C (stability) [12]. Characterization includes determination of physicochemical, biological, and immunological properties and the impurity profiles. To support the advancements in the analytical field and selection of the best methods for particular purpose, guidelines do not recommend specific analytical technologies, but focus on the type of quality information to be collected. To ensure sufficient characterization and due to method limitations (e.g., limited resolution and sensitivity), the use of orthogonal techniques based on different physicochemical or biological principles may be recommended. Analytical methods used for final container testing should be validated, while methods used for characterization should be qualified (ICH Q6B [5], ICH Q2 [13]).

3.1 Critical Quality Attributes

According to the ICH Q8(R2) [2] guideline, a critical quality attribute is "a physical, chemical, biological or microbiological property or characteristic that should be within an appropriate limit, range or distribution to ensure the desired product quality." Each unique product development program directs the

[1] Center for Drug Evaluation and Research/Office of Pharmaceutical Quality/Office of Biotechnology Products

[2] Center for Biologics Evaluation and Research/Office of Blood Research and Review

characterization of product-specific attributes within the regulatory framework of ensuring identity, quality, purity, potency, and safety.

3.1.1 Identity

As required by 21 CFR §610.14, "The identity test shall be specific for each product in a manner that will adequately identify it as the product designated on final container and package labels and circulars, and distinguish it from any other product being processed in the same laboratory."

3.1.2 Quality

The quality of a product is a complex property that includes characteristics of the biotherapeutic molecule, excipients, and solution properties. Molecular integrity of the active ingredient is ensured through extensive physicochemical characterization of primary structure, posttranslational modifications, higher-order structure, and biological tests, which provide evidence of the active conformation of the protein. General product quality tests may include pH, osmolality, appearance, color, or clarity of the reconstituted powder or liquid drug product. Robust quality systems and manufacture under cGMP provide continuous assurance that specified product quality is maintained throughout a product's shelf life.

3.1.3 Purity

According to 21 CFR §600.3(r), "*Purity* means relative freedom from extraneous matter in the finished product, whether or not harmful to the recipient or deleterious to the product." Impurities in a therapeutic protein starting material should be removed to the extent possible, and the remaining impurities should be defined and controlled to specified, justified limits. Impurities can be defined as product related or process related. Product-related impurities include molecular variants and degradants, such as aggregates, glycoform variants, or other degradation products of the active ingredient. Examples of process-related impurities include host-cell impurities or culture additives from cell cultivation processes, reaction by-products, leachates from resins used for chromatography purification, solvent and detergent used for viral inactivation, or components from buffers added to reach the desired pH or ionic strength during purification. Characterizing impurities to the greatest possible extent enables a manufacturer to design, validate, and monitor a purification process to ensure continued product purity and safety. Specific recommendations for mitigation immunogenicity risk related to the presence of impurities can be found in FDA Guidance for Industry: Immunogenicity Assessment for Therapeutic Protein Products (August 2014 [14]).

3.1.4 Potency

A suitable potency assay is crucial to manufacturing control and clinical use of protein therapeutics and is required by statute and regulation. According to 21 CFR §600.3(s), "The word *potency* is interpreted to mean the specific ability or capacity of the product, as indicated by appropriate laboratory tests or by adequately controlled clinical data obtained through the administration of the product in the manner intended, to effect a given result." Potency assays may be in vitro or in vivo tests, which provide a *quantitative* measure of the product attribute linked to its primary mechanism of action.

The ICH Q6B guideline [5] defines three categories of potency assays: (1) in vivo animal-based, (2) in vitro cell-based, or (3) biochemical assays using purified systems. The potency, which is defined in units of activity that are product specific, is determined for each product by comparing it to a suitably characterized, product-specific reference standard, which can be international, national, or in-house.

The use of a voluntary consensus standard such as an international standard or a regional standard as a primary standard is recommended whenever such standards are available. These standards are established in collaborative studies based on the need, e.g., when several products in a class are present on the market. Various CBER, WHO, and NIBSC standards have been established and are available for use by biologics manufacturers. It is recommended that in-house potency reference standards (secondary standards) be established using well-characterized product lots, whose potency assignments are traceable to the consensus standard. Reference standards should be established early in development and updated to reflect changes in manufacturing process. The potency assay should be validated as suitable by the time of pivotal trials.

At present, no international reference standards are available for monoclonal antibodies and most other therapeutic proteins. In such cases, ICH Q6B [5] recommends that an in-house primary reference standard be prepared from materials representative of production and clinical materials. The potency of the product (measured in potency assays) should be compared to data from the in-house reference standard and reported in "in-house units." As product evolves during process development, so do the in-house reference standards, so care should be taken to qualify the new standard against the previous iteration. It is crucial that criteria for potency have a narrow acceptable range to ensure control over product drift. By the time of a BLA submission, a two-tier system should be implemented where the primary reference standard should be representative of the materials used in the pivotal clinical studies and that a working reference standard be appropriately qualified against the primary reference standard.

3.2 Safety

3.2.1 Adventitious Agent Safety

For all biologics, orthogonal procedures aimed at drastically reducing the risk of pathogen contamination of the product are in place. Final products intended for parenteral administration are required to be sterile (21 CFR §610.12). Endotoxin levels should also meet USP limits specific for the route of administration (USP < 85 >). For parenteral products, the limit is no more than 5 EU/kg/h. For emerging pathogenic agents, such as prions associated with TSE, guidelines have been developed that address the novel concerns with these agents [15].

To ensure that products are free of adventitious viruses, manufacturers of licensed protein therapeutics implement the viral safety tripod recommended in the ICH Q5A guideline: (1) careful selection and extensive characterization of starting materials, other raw materials, and excipients, (2) viral clearance validation of targeted steps in the manufacturing process, and (3) in-process testing for adventitious viruses, as applicable.

For plasma protein therapeutics, there is a real potential for viral contamination of plasma material. Therefore, a multilayered approach to blood donation safety is implemented at each blood establishment to enhance viral safety. Overlapping safeguards include donor screening, donor deferral, infectious disease testing, quarantine of donated plasma until tested and demonstrated to be free of infectious agents, and compliance with cGMP for deviation investigation and corrective and preventive action. Various guidelines have been published by the agency to ensure implementation of this system.[3] For products that are made through biotechnology, viral safety is enhanced by establishment of a highly controlled, tiered cell banking system and extensive testing of each cell bank for relevant endogenous or adventitious viruses.

For mAbs and other products that are amenable to platform manufacturing processes, generic and modular virus clearance studies may be performed. A "generic" clearance study is one in which virus removal and inactivation is demonstrated for several steps in the purification process of a model antibody. These data may then be extrapolated to other antibodies following the same purification and virus removal or inactivation scheme as the model antibody. A modular clearance study is one that demonstrates virus removal or inactivation of individual steps during the purification process (column chromatography, filtration, pasteurization, solvent/detergent, low pH, etc.). Each module in the purification scheme may be studied independently of the other modules. Different model mAbs may be used to demonstrate viral clearance in different modules, if necessary. If the purification process of a product mAb differs at any of the virus removal or

[3] For FDA guidance documents for blood and plasma products, refer to http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/Blood/default.htm

inactivation modules from the model mAb, this module must be studied independently from the model. The other, identical modules in the procedure may be extrapolated to the product mAb.

For therapeutic proteins not manufactured using a platform technology, a clearance study demonstrating the effectiveness of viral removal or inactivation for select unit operations in the purification process is necessary to validate the viral inactivation capability of the process. The conditions and operating parameters, including column dimensions, buffer, temperature, time, etc., in the study should closely mimic the worst-case scenario of the manufacturing scale production condition [16].

3.2.2 Immunogenicity

While biologics provide remarkable therapeutic benefits, they can also trigger the development of anti-drug antibodies (ADA) that could result in loss of efficacy or adverse events. The immune responses may vary in clinical relevance, ranging from antibody responses with no apparent clinical manifestations to life-threatening and catastrophic reactions [14, 17]. The immunogenicity risk of a therapeutic protein should therefore be considered at the earliest stage and reevaluated at each stage of product development. The evaluation should consider factors such as product origin and characteristics, the immune responses of concern, the target disease indication, and the proposed patient population. Although certain quality attributes of biologics have been related to immunogenicity [17], the ability of a molecule to elicit an immune response cannot be reliably predicted by physicochemical characterization or by studies using animal models. Immunogenicity of a protein product is therefore assessed during clinical trials using binding and neutralizing ADA assays to evaluate the production of ADAs. In most cases, binding and neutralizing ADA results obtained with validated assays are required to support licensure. For high-risk products, such as protein products of nonredundant endogenous cytokines, critical enzymes, and growth factors, implementation of preliminary validated assays in early trials may be necessary to obtain real-time data and to minimize the risk of neutralization of the endogenous protein counterpart and subsequent loss of its physiological function [18].

3.3 Specifications

The ICH Q6B [5] guideline defines a specification as, “a list of tests, references to analytical procedures and appropriate acceptance criteria, which are numerical limits, ranges or other criteria for the tests described.” Specifications are a subset of product quality attributes chosen to confirm product quality as it relates to safety and efficacy. Tests and acceptance criteria should be based on levels shown to be safe and effective in clinical trials supporting licensure, in the context of

manufacturing process capabilities. Analytical methods used to measure specifications should be validated. The approach to setting specifications may be different for each product class and should be justified in the BLA.

3.4 Stability

Manufacturers should demonstrate acceptable product quality not only at the time of product release but also throughout its shelf life (21 CFR §211.137). The quality of the product changes with time and is affected by multiple factors including temperature, humidity, and light. Interactions of the product with the container closure system and resulting from stress occurring, for example, on the liquid/air interphase may affect the molecular integrity of the product. Stability testing is, thus, performed to establish adequate storage conditions and shelf life and is performed in the same container closure system in which the product is marketed (21 CFR §211.166). Changes to the primary, secondary, and higher-order structure during degradation processes may not only reduce the amount of active product but also affect product safety by increasing the risk of the immune response.

Relevant tests demonstrating stability should be selected, and they usually include characterizing appearance, protein concentration, activity, molecular integrity (fragmentation, aggregation), quantification of particulate matter, pH, water content for lyophilized products, pyrogenicity and/or bacterial endotoxin testing, and sterility. Stability testing starts during product development and upon approval of a BLA is performed annually to ensure that product characteristics remain within specifications. Enhanced stability testing is performed when manufacturing changes are introduced. Product stability profile is characterized under the normal storage conditions to establish the product shelf life and also at a higher temperature (s). Even with extensive characterization, changes to the complex product characteristics, which include the composition of the product and its physicochemical properties, may go unnoticed. Degradation profiles observed under accelerated or forced degradation conditions help early identification of changes that may potentially impact safety and efficacy during the product shelf life. Performing stability studies at elevated temperature(s) may help identify changes that affect product stability and to establish comparability of the product before and after manufacturing change.

3.5 Comparability

Extensive characterization that is performed during product development helps ensure that products maintain their quality, safety, and efficacy throughout life cycle. The properties of the product used in clinical trials serve as a critical comparator, i.e., the quality of the product after change has to be comparable or

better than of the product tested in pivotal clinical trials. Product comparability pre- and post-manufacturing change is established based on the analysis of data collected from product characterization, routine batch analyses, in-process control, process validation and evaluation, and product stability (ICH Q5E, [6]). Data collected after change are compared to predefined criteria established based on quality information accumulated before the change is made. For certain applications and when prospectively established, successful completion of such comparability protocols can impact regulatory pathway (and review timelines), for example, enabling a sponsor to file what would be a prior approval supplement (PAS) as a "changes being effected in 30 days" (CBE30) supplement.

4 The Development and Characterization of Different Classes of Protein Biotherapeutics

Below we describe approaches to analytical characterization for four different classes of proteins: (1) polyclonal immune globulins, (2) coagulation and hemostasis proteins, (3) monoclonal antibodies, and (4) other therapeutic proteins. Chronologically, these products span the entire history of the FDA ranging from the first human protein (an immune globulin preparation used for the prevention of measles) approved in 1934 [19] to the most recent recombinant proteins made in large quantities in bioreactors. As scientific knowledge and clinical experience with these products have progressed, so have the analytical methods used to characterize and monitor their quality attributes during development, manufacturing, and at release. cGMP manufacturing practices and regulatory expectations are that all products, even those first approved decades ago, be characterized to the extent possible using current analytical methods. Appropriate studies are used to bridge existing product information, including clinical safety and efficacy information with current analytical methodology and manufacturing practices. Thus, it follows that analytical methods in general are common for all product classes discussed here. What is different, as it will become apparent in the following sections, are the experiences in the product-specific regulatory divisions of FDA. These experiences, their relation to product mode of action, its manufacturing process, and the history of clinical use safety and efficacy have flavored the analytical emphasis in each division and provide the basis for the fit-for-purpose approach to characterizing therapeutic proteins described here.

4.1 Plasma-Derived Immunoglobulins

Polyclonal immune globulin products are biological therapies comprised of human or animal plasma or serum-derived products.[4] At present, there are 38 such products approved by the US FDA, manufactured by 15 different firms. Of these, DigiFab® and CroFab® contain the Fab part of the immunoglobulin G (IgG) molecule, whereas the active ingredient in Anascorp® and Botulism Antitoxin Heptavalent is the $(Fab')_2$. The remaining approved products are intact antibodies (Table 1).

4.1.1 Quality Polyclonal IgG Products Start with Quality Plasma

Plasma used for making human polyclonal IgG products is collected in licensed and inspected plasma or blood donation centers. The quality of the starting material is essential for the ultimate safety and quality of the end products. Thus, to ensure that donated plasma is safe from hazards to human health and to minimize the likelihood of transmitting infectious disease, multiple measures are in place. These include screening of donors by questionnaire, instituting good collection practices in the donation centers, and testing of donated plasma for blood-transmitted pathogens. Many of these mandatory practices are codified in appropriate subparts of 21 CFR, Part 640, and more specific details are set out in FDA and international guidelines [20–23]. Other safeguards aimed at lowering potential risk for contamination of US products with transmissible spongiform encephalopathies (TSEs), such as variant Creutzfeldt-Jakob disease (vCJD), are specified in a recent FDA guidance document [15]. Additional processes are incorporated at the IgG product manufacturing site to ensure that plasma pools collected in the US licensed donation centers (including all manufacturing intermediates and final products) are segregated from those not originating from centers not licensed by the FDA.

Polyclonal IgG biologics are often categorized as either normal or specific. The only difference between the two product categories lies in the specificity of the polyclonal antibodies they contain. Specific products have high titers of antibodies directed against particular pathogenic agents, such as viruses (e.g., hepatitis B, rabies, varicella, and vaccinia) or bacterial toxins (e.g., tetanus toxoid, botulinum toxin). Antilymphocyte and antithymocyte polyclonals are used to suppress the adaptive immune system in transplant settings. For specific IgG products sourced from human plasma, donors often receive vaccinations and are tested to ensure that their serum contains sufficient amounts of the specific antibodies of interest. Animal plasma is obtained from healthy animals that have high titers of desired antibodies achieved through immunization and boosting with the specific antigen(s). Animals are almost always used to make products directed against toxins, venoms,

[4] Regulatory oversight for these products is the responsibility of the Laboratory of Plasma Derivatives, Division of Hematology Research and Review in the Office of Blood Research and Review.

Table 1 IgG products currently marketed in the USA

	Product name	Manufacturer/sponsor	Proprietary name
Normal IgG	Immune globulin intravenous (human), 10% liquid	Biotest Pharmaceuticals Corporation	Bivigam®
	Immune globulin intravenous (human)	CSL Behring AG	Carimune® NF, Nanofiltered
	Immune globulin intravenous (human)	Instituto Grifols, S.A.	Flebogamma® 5% DIF
	Immune globulin intravenous (human)	Instituto Grifols, S.A.	Flebogamma® 10% DIF
	Immune globulin (human)	Grifols Therapeutics Inc.	GamaSTAN® S/D
	Immune globulin infusion (human), 10%	Baxalta (formerly Baxter Healthcare Corporation)	Gammagard Liquid®
	Immune globulin intravenous (human)	Baxalta (formerly Baxter Healthcare Corporation)	Gammagard® S/D
	Immune globulin intravenous (human), 5% liquid	Bio Products Laboratory Limited	Gammaplex®
	Immune globulin injection (human), 10% caprylate/chromatography purified	Grifols Therapeutics Inc.	Gamunex® C
	Immune globulin injection (human), 10%	Grifols Therapeutics Inc./Kedrion	Gamaked®
	Immune globulin subcutaneous (human), 20% liquid	CSL Behring AG	Hizentra®
	Immune globulin infusion 10% (human) with recombinant human hyaluronidase	Baxalta (formerly Baxter Healthcare Corporation)	HyQvia®
	Immune globulin intravenous (Human) 5% liquid	Octapharma Pharmazeutika Produktionsgesellschaft m.b.H. (Vienna) and Octapharma AB (Sweden)	Octagam® 5% liquid preparation
	Immune globulin intravenous (human) 10% liquid	Octapharma Pharmazeutika Produktionsgesellschaft m.b.H. (Vienna) and Octapharma AB (Sweden)	Octagam® 10% liquid preparation
	Immune globulin intravenous (human), 10% liquid	CSL Behring AG	Privigen®
Specific IgG	Centruroides (scorpion) immune $F(ab')_2$ (equine) injection	Rare Disease Therapeutics Inc. (RDT)	Anascorp®
	Anthrax IGIV	Emergent BioSolutions	Anthrasil®
	Antivenin (*Latrodectus mactans*) (black widow spider), equine origin	Merck	Antivenin
	Botulism immune globulin intravenous (human)	California Department of Public Health (CDPH)	BabyBIG®

(continued)

Table 1 (continued)

	Product name	Manufacturer/sponsor	Proprietary name
	Botulism Antitoxin Heptavalent (A, B, C, D, E, F, G) (equine)	Emergent BioSolutions	BAT®
	Antivenin (*Micrurus fulvius*) (North American coral snake), equine origin	Pfizer	Wyeth® Antivenin
	Crotalidae [pit viper] polyvalent immune fab (ovine)	BTG International Inc.	CroFab®
	Cytomegalovirus immune globulin intravenous (human) (CMV-IGIV)	CSL Behring AG	Cytogam®
	Digoxin immune fab (ovine)	BTG International Inc.	DigiFab®
	Hepatitis B immune globulin (human)	Biotest Pharmaceuticals Corporation	Nabi-HB®
	Hepatitis B immune globulin (human)	Emergent BioSolutions	HepaGam B®
	Hepatitis B immune globulin (Human)	Grifols Therapeutics Inc.	HyperHEP® B S/D
	Rabies immune globulin (human)	Grifols Therapeutics Inc.	HyperRAB® S/D
	Rh_o(D) immune globulin (human)	Grifols Therapeutics Inc.	HyperRHO® S/D (Full Dose)
			HyperRHO® S/D (Mini-Dose)
	Tetanus immune globulin (human)	Grifols Therapeutics Inc.	HyperTET® S/D
	Rabies immune globulin (human) USP, heat treated	Sanofi Pasteur SA	Imogam® Rabies HT
	Rho(D) immune globulin (human)	Kedrion Biopharma Inc.	RhoGam® Ultra-Filtered PLUS
			MICRhoGAM® Ultra-Filtered PLUS
	Rh_o(D) immune globulin intravenous (human)	CSL Behring AG	Rhophylac®
	Varicella zoster immune globulin (human)	Emergent BioSolutions	VARIZIG®
	Vaccinia immune globulin intravenous (human)	Emergent BioSolutions	VIGIV
	Rh_o (D) immune globulin intravenous (human)	Emergent BioSolutions	WinRho® SDF
Antithymocyte IgG	Equine thymocyte immune globulin injection, solution	Pfizer	Atgam®
	Antithymocyte globulin (rabbit)	Genzyme Polyclonals S.A.S.	Thymoglobulin®

or other highly potent antigens which may not be safe for immunization of human donors. By contrast, normal immunoglobulins represent a diverse antibody repertoire and are made from pooled human plasma from donors not receiving any specifically targeted vaccinations. As specified in 21 CFR, §640.102(d), each lot of normal human IgG represents a pool of material obtained from not less than 1,000 donors. In manufacturing practice, many more individual donations are frequently pooled to manufacture large batches of IgG. It is recommended that the size of plasma pool be limited to no more than 60,000 donations [24–26].

4.1.2 Application of Good Practices in Polyclonal IgG Product Manufacture

Manufacturing processes and analytical techniques used for isolating and evaluating pure and potent immune products obtained from human and animal plasma have evolved considerably in recent years. While the impetus on the manufacturers' side has been to increase yield to satisfy rising demand, the focus on the regulatory side has remained on ensuring safety while fulfilling the public health need for these products. Presently there are several general purification schemes used for licensed products; details on these can be found in recent reviews [27–30]. Many of them are derived from the original methods discovered and refined by Cohn and his collaborators [31–33] representing a series of finely tuned and closely controlled manufacturing steps to ensure the quality attributes of marketed products. The need for strict controls of the manufacturing process parameters cannot be overemphasized, given that seemingly minor changes or deviations can alter product characteristics, sometimes resulting in unanticipated and severe adverse outcomes [34–36].

As for all protein biologics, a series of analytical tests designed to monitor product quality and quantity are performed during the manufacturing process of IgG products. As for other protein therapeutics, these in-process tests, implemented at different key manufacturing steps, provide the data for the historical database that will be accumulated during the lifetime of the IgG product. This database is used to ensure the process is performing according to expectations and to establish comparability, discussed in detail earlier in this chapter. Many of the in-process testing methods used in manufacturing IgG products are validated. Those that are not should meet the bar of being scientifically sound and reflect current laboratory standards and methodology.

4.1.3 Parameters Measured to Ensure Polyclonal IgG Product Quality

The analytical measurements used to characterize polyclonal IgG preparations at release can roughly be divided into the following categories: those that measure (1) active ingredient indicative parameters, (2) impurities, (3) final product solution properties, (4) excipients, and (5) parameters related to viral safety.

Active Ingredient Indicative Parameters

For IgG products, key indicative parameters, i.e., those that measure active ingredient quantity, purity, and structural integrity in the final product, typically include total protein content, absolute and relative amount of IgG proteins, molecular size distribution, protein banding pattern, and other IgG biophysical properties. Unlike other protein biologics, there are a few product-specific release tests explicitly required for normal IgG products derived from human plasma. Potency tests for antibodies against measles, diphtheria, and poliomyelitis are codified in 21 CFR §640, Subpart J.[5] When performed by bioassay, these measurements provide information about biological function of the IgG lot at release and throughout its shelf life. Additionally, these parameters are an indicator of the manufacturing process consistency. For these reasons, binding only assays are not, as a general rule, acceptable methods for determining potency. Cell culture or in vivo neutralization assays are the methods accepted by the FDA for measuring these specificities.

The same potency assays are not generally required for specific IgG products, but potency tests measuring the neutralizing activity of the specific IgGs are. Due to methodological limitations, there have been instances when binding assays have been used in lieu of bioassays for some specific IgG products. This has been acceptable in cases, such as hepatitis B-specific IgG products, where the protective target levels of anti-Hbs were well known from epidemiologic or clinical studies [37].

Other antibody specificities can be measured, especially when characterizing a new product or as part of a comparability protocol. Examples include pneumococcal, B19, H. influenza B, or hepatitis A-specific antibodies.

Impurities

As for all protein biologics, impurities in polyclonal IgG products are categorized as either product related or process related.

Immunoglobulin-related impurities usually characterized in IgG products include IgM, IgE, IgA, or undesirable types or forms of IgG molecules. The latter include IgG aggregates and IgG fragments, which in addition to being measured at product release, also serve as stability indicators. High aggregate levels can cause clinical adverse events [38], whereas fragments measure product integrity during storage. Their increase over time can provide an early indication of pending potency loss. Anticomplementary activity (ACA) is measured as a functional property of IgG aggregates, and high levels are potential indicator of infusion-related adverse events. Antibodies against blood-group antigens A, B, and Rho(D) are

[5] 21 CFR §601.104 specifies that the minimum levels for diphtheria antibodies should be 2 international units per mL, and those for measles and poliomyelitis type 1, type 2, or type 3 should be compared to the levels found in CBER reference.

also measured and controlled in IgG products since high levels of these antibodies are associated with clinically significant hemolysis in recipients of these blood groups [39].

A second category of impurities is other plasma proteins that may co-purify with IgG at clinically significant levels. Several of these entities, i.e., pre-kallikrein activator (PKA) and activated coagulation Factor XI (FXIa), have been implicated in clinical adverse reactions [38, 40] and as such are measured in final release tests and controlled by release specifications. Depending on the manufacturing process, certain products may have additional release specifications to control product-specific protein impurities. Additional product qualities are measured when characterizing a new product or as part of a comparability protocol. Examples include other coagulation factors (fibrinogen, Factor IX, FXII), other plasma proteins (i.e., albumin, fibronectin, apolipoprotein A1, antithrombin III, C1-esterase inhibitor, α2-macroglobulin, transferrin), and lipids.

Process-related impurities include process-specific residues such as solvent or detergent from the viral deactivation steps, leachates from chromatographic column resins, filters as well as their housing, and buffers or other chemicals used at different steps in the process. Examples include polysorbate 80 (PS80, Tween 80), tri(*n*-butyl)phosphate (TnBP), and heavy metals. This class of impurities is controlled by designing a robust purification process that can perform well under "worst-case" conditions. Their levels are monitored by in-process and release tests; action limits and release specifications that ensure these impurities remain within acceptable levels are instituted. If these (or any impurities) are not removed to achieve sufficiently low levels, toxicity assessments and, in some cases, toxicology studies are performed to assess potential risk to patients under "worst-case scenario" conditions. If warranted, stricter process controls or additional mitigation steps are incorporated during the manufacturing to ensure adequate safety margins in clinical use, including for susceptible populations.

Final Product Solution Properties and Excipients

Final product solution properties often measured for polyclonal IgG products include pH, osmolarity, particulates (visible and subvisible), and appearance. The latter includes color and clarity for liquid products and, for lyophilized products, appearance and reconstitution time for the dried cake.

Polyclonal IgG products are formulated at a very high nominal protein concentration, ranging from 50–200 g/L. To ensure that native, biologically active conformation is preserved, excipients are added as stabilizers in the final formulation of these products. Most of the excipients fall under two classes of chemicals: sugars (i.e., maltose, sucrose, and glucose) or amino acids (i.e., glycine and proline). Both the choice and concentration of the excipient are critical parameters that help limit IgG aggregation and ensure product quality throughout the life cycle. Thus, these properties are measured and controlled at lot release.

Parameters Related to Microbial Growth and Pathogen Safety

Another group of analytical techniques are those used to measure parameters such as sterility and pathogen safety. Some of these concepts (i.e., sterility and viral clearance) are similar among all the different protein products discussed in this chapter and have already been mentioned. Others, such as bio-burden levels at upstream manufacturing steps, are monitored and controlled, and pyrogenic bacterial constituents are measured in final product using rabbit pyrogen test or LAL. Monocyte activation test has emerged as a possible method for measuring pyrogen activity in IgG products. Validation studies should be submitted to support any changes to existing release tests. Other parameters such as the level of antibodies directed against hepatitis B surface antigen (anti-HBsAg) that have been found to be important as markers for the stability of the preparation and provide an additional measure of viral safety are also measured.

4.1.4 Emerging Areas in Analytical Measurement Development

With the ever-increasing use of IgG products in auto-immune disorders, especially those with neurological manifestations [41], there is a lot of attention being given from the regulators, industry, and scientific community to defining the unique product attributes which underlie efficacy in these disease states and establishing methods to measure them. There is growing evidence that efficacy of polyclonal IgG products in auto-immune indications is Fc mediated [42]. Thus, discovering accurate and precise methods to measure Fc-receptor binding activity which correlates with effector function(s) has become a priority. An optimal such measure would be a bioassay connecting Fc-receptor binding of the IgG preparation with a biological readout such as activation or inhibition of a functional measurement, for example, cytokine release. The appropriate readout should also be related to the pathophysiology of the disease. Given that the underlying mechanisms of many auto-immune conditions are not only complex but also incompletely understood, picking the "ideal" assay represents a challenge and an area that would benefit from further research.

Other areas of regulatory interest include improving the methodology for measuring IgG aggregates in the nanometer range and developing neutralizing assays for specific IgG products where such assays are not available. For the latter, challenges remain in validating neutralizing assays for several viral agents, such as CMV, HBV, and HCV, and continued research would be beneficial.

4.2 Recombinant Therapeutic Proteins for Hemostasis

Proteins of the hemostatic system span a broad range of sizes, shapes, structures, and quality characteristics, which direct multifaceted interactions. Complex interactions, however, among these widely varying hemostatic proteins drive the

common goal of hemostasis, the system for maintaining vessel patency in response to injury.

4.2.1 The Hemostatic System

Hemostasis in response to vessel injury is a complex process involving coordinated interplay among subendothelial matrix components, the endothelial lining, blood cells, and finally the components of the blood coagulation cascade leading to the formation of a fibrin clot (Fig. 2). The physiological balance of procoagulant, anticoagulant, and fibrinolytic activities among the players of the blood coagulation cascade is designed to tailor clot location, size, and stability in accordance with the need at the injury site. The clot-forming cascade is propagated through a series of multicomponent, proteolytic complexes operating on the surfaces of activated cells, such as platelets, white blood cells, and endothelial cells. These cell surface

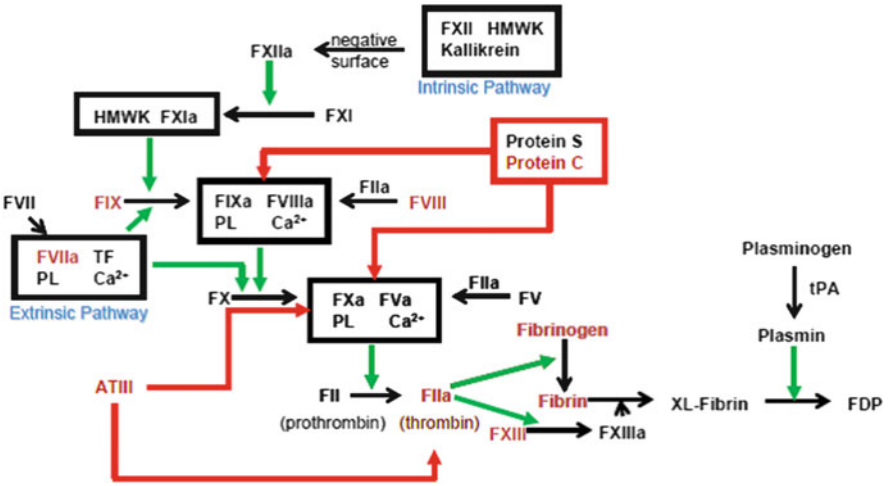


Fig. 2 The hemostatic system: The ability to form and maintain a blood clot in the right place, at the right time, involves proper, physiological coordination of procoagulant, anticoagulant, and fibrinolytic pathways. As depicted, each pathway executes a tightly controlled cascade of sequential proteolytic activation reactions, operated by specific proteolytic complexes, which are designed to amplify each system while responding to feedback mechanisms. Blood coagulation proteins circulate as inactive zymogens or cofactors designated as "factors" with assigned roman numerals. Inactive zymogens and cofactors are sequentially activated at the site of injury, through the cascade depicted in the figure. As illustrated, extrinsic and intrinsic pathways operate through proteolytic complexes (in *black boxes*) comprising an activated proteolytic enzyme, an activated cofactor, calcium, and phospholipid which serve to activate each zymogen in succession. Progression of the blood coagulation cascade is subject to modulation (indicated through *red arrows*) by the protease inhibitor, antithrombin III, and the anticoagulant complex comprising activated forms of protein C and protein S. Disorders of hemostasis may be treated by replacement therapy with hemostasis protein concentrates purified from plasma or produced through recombinant DNA technology. Currently marketed therapeutic concentrates are indicated in *red*.

assembled complexes accomplish the sequential activation of their serine proteases so that thrombin can be generated locally to convert fibrinogen to an insoluble, cross-linked fibrin clot at the specific site of injury. The importance of physiologically functional hemostasis is evidenced by the occurrence of pathologic bleeding or thrombotic disease caused by its disruption. Inherited disorders of hemostasis may result from a missing or nonfunctional coagulation factor, anticoagulant, or fibrinolytic protein. In addition, acquired bleeding or thrombotic disease may be a consequence of environmental stimuli [43]. The current US market offers a number of licensed, safe, and efficacious plasma protein therapeutics indicated for disorders of hemostasis. Licensed products include purified coagulation factor or anticoagulant protein concentrates, covering many proteins required for proper hemostasis. Hemostasis protein therapeutics may be purified from human or animal plasma or expressed through recombinant DNA technology. Figure 2 indicates in "red" those proteins for which there are currently licensed concentrates. The majority of licensed products are indicated as replacement proteins to correct disorders of hemostasis through on-demand treatment, routine prophylaxis, or perioperative management of bleeding. Thrombin and fibrinogen are components of fibrin sealant products, which mimic the final step in blood coagulation, to stop minor surgical bleeding when standard surgical practices are ineffective or impractical.

4.2.2 Analytical Characterization of Recombinant Factor VIII and Factor IX

Successful analytical programs provide conclusive evidence that demonstrates preservation of native protein structure, retention of all functional properties, and control of impurities. Advances in manufacturing and analytical techniques have allowed comparable approaches for complete characterization of the vast array of hemostasis proteins whether purified from plasma or recombinant DNA (rDNA)-derived cell culture. This section will focus on analytical paradigms for two recombinant coagulation factors: Factor VIII and Factor IX. Factor VIII is the missing protein in hemophilia A, likely the first documented bleeding disorder (Talmud, circa 200 CE) [44]. Factor IX is the missing protein in hemophilia B, a bleeding disorder prevalent in European royal families in the nineteenth and twentieth centuries [45]. Although highly divergent in structure and quality characteristics, Factor VIII and Factor IX are connected through their interdependence as components of the proteolytic complex responsible for activating Factor X. Factor VIII and Factor IX products currently occupy the majority of the commercial market for hemostasis protein concentrates. Furthermore, both proteins have been the targets of significant engineering efforts designed to enhance their clinical performance.

Primary Structure

Coagulation Factor VIII is a large, glycosylated protein, which undergoes multiple, directed proteolytic cleavages in order to create the biologically active molecule. The ~330 kDa protein contains 2,332 amino acids arranged in domain structure: A1–A2–B–A3–C1–C2. For the full-length molecule, several posttranslational, proteolytic cleavages in the B domain produce a series of divalent cation-linked, two-chain molecules. The N-terminal, heavy chain, comprising A1, A2, and cleaved B domains, displays heterogeneity upon analysis by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), exhibiting several bands possessing molecular weights (MW) between 210 kDa and 90 kDa. The light chain comprising A3, C1, and C2 domains has an MW of ~80 kDa [46]. There are four commercially available recombinant full-length Factor VIII concentrates. Thrombin cleavage of Factor VIII followed by subunit analysis using SDS-PAGE or HPLC methods is particularly useful for confirming Factor VIII domain structure. During development of expression systems for recombinant Factor VIII, it was discovered that removing the B domain increased expression and decreased molecular heterogeneity through removal of the majority of glycosylation sites and heavy chain size heterogeneity while retaining in vivo procoagulant function. SDS-PAGE analysis of B domain-deleted Factor VIII yields a simplified banding pattern comprising a single-chain molecule with MW ~170 kDa, a ~90 kDa heavy chain, and an ~80 kDa light chain [47]. Four commercially available Factor VIII concentrates are based on rDNA constructs that delete the B domain.

Factor IX circulates as a single chain, ~56 kDa molecule possessing five structural domains in the following order from amino to carboxyl terminus: Gla-EGF1-EGF2-activation peptide-protease domain. Similar to Factor VIII, directed proteolytic processing and unique posttranslational modifications are required in order to generate the biologically active molecule [48]. Four recombinant Factor IX concentrates are commercially available.

Posttranslational Modifications Important for Pharmacologic Action

The design of expression systems to produce biologically active recombinant Factor VIII or Factor IX must consider the fidelity of posttranslational proteolytic processing, glycosylation, sulfation, and γ-glutamyl carboxylation, all of which have reported impact on pharmacokinetic or pharmacodynamic properties. Recombinant expression of mature, fully processed Factor IX requires co-expression of a signal peptide-processing enzyme. Critical to the function of Factor IX and other vitamin K-dependent hemostasis proteins is the posttranslational addition of carboxyl groups to specific glutamyl residues contained within the N-terminal (Gla) domain. Factor IX contains 12 sites for γ-carboxylation. A high level of γ-carboxylation must be achieved in order for recombinant Factor IX to properly interact with activated cell surfaces as part of its procoagulant mechanism of action. Anion exchange chromatographic separation followed by peptide mapping

strategies enhanced by sensitive mass spectrometric techniques may be used to accurately quantitate the degree of γ-carboxylation. Tyrosine sulfation has been described for Factor VIII and Factor IX. Particularly important for Factor VIII function are sites on acidic peptides, which direct thrombin activation of Factor VIII. A critical site at Tyrosine1680 mediates Factor VIII binding to von Willebrand factor, a large multimeric protein that stabilizes Factor VIII in circulation [43, 46, 48]. Since heterogeneity in glycosylation gives rise to structural diversity and impacts clinical pharmacology, complete characterization and tight control of protein glycoforms are critical to product quality and manufacturing consistency. Factor VIII and Factor IX proteins produced in nonhuman cell lines have been characterized to contain minimal quantities of *N*-glycolylneuraminic acid (NGNA), a sialic acid substituted for the more common *N*-acetylneuraminic acid (NANA) and galactose-α-1,3-galactose (α-Gal), with recent advances in mass spectrometric techniques allowing more sensitive α-Gal detection [49]. Although no direct adverse clinical consequences have been reported, NGNA and α-Gal levels are tightly controlled in all Factor VIII and Factor IX concentrates.

Purity and Impurities

Purity is defined as the relative freedom from extraneous matter whether or not harmful to the recipient or deleterious to the product. It is an expectation that impurities in Factor VIII or Factor IX therapeutic concentrates be removed to the extent possible and that the remaining impurities be defined and controlled to specified, justified limits.

Product-Related Impurities

As for other recombinant proteins, Factor VIII and Factor IX concentrates are controlled for aggregates, polymeric forms, truncated forms, fragments, and oxidized molecular variants, as applicable, using peptide mapping, HPLC, and SDS-PAGE methods. Control of aggregates in the final product, most commonly by size-exclusion chromatography (SEC), is particularly critical due to their potential for inducing an immunogenic response. Nonspecific adsorption of proteins on the matrix of SEC columns can confound the analysis and give results that do not accurately represent the aggregate content present in the product. Nonspecific adsorption can be minimized by preconditioning the column through multiple injections of the protein prior to sample analysis, allowing a layer of the protein to coat binding sites on the column matrix. However, such preconditioning may also reduce the pore size of the matrix and change the performance characteristics of the column by reducing peak resolution and separation range. In addition, because of the dynamic equilibrium between the adsorbed protein and the protein in solution, the adsorbed protein may slowly leach into the eluate giving aberrant results. Therefore, orthogonal methods to analyze aggregates, such as dynamic light scattering or analytical ultracentrifugation, are essential in validating an SEC

method for its intended purpose. Factor VIII and Factor IX circulate as inactive precursors until their activation is initiated in vivo, in response to injury. Since premature activation is undesirable, purification procedures and quality control tests must ensure retention of the protein in its unactivated state. Therefore, product-related impurities in Factor VIII or Factor IX products with the greatest potential negative impact to product quality and safety constitute the activated forms. Of particular concern, activated Factor IX (FIXa) impurity has been associated with reports of thromboembolic complications and is routinely characterized and controlled in all Factor IX products [50].

Process-Related Impurities

Strategies described in the ICH Q6B guideline [5] for characterization and control of process-related impurities associated with recombinant cell culture systems and downstream purification steps are also routinely applied to characterization and control of recombinant Factor VIII or Factor IX concentrates. It is an expectation that host-cell protein (HCP) be characterized and controlled using an in-house assay specific to the product cell line and validated for adequate detection of all relevant proteins. Two-dimensional SDS-PAGE and immunoblotting techniques have been applied to the development and validation of suitable HCP enzyme-linked immunosorbent assays (ELISA). Affinity chromatography strategies are often employed during Factor VIII or Factor IX purification schemes. Sensitive in-house assays specific for affinity ligands should be developed and validated to control for affinity column leaching. Unique to Factor IX concentrates is the control of the signal peptide processing enzyme impurity.

Viral Safety

Viral safety is of utmost concern for a class of products with a history of virus transmission. Plasma-derived fibrinogen concentrate, first marketed in 1947, was removed from commercial distribution in 1977 due to transmission of hepatitis [51]. Then came the AIDS crisis in the early 1980s, which deeply impacted the hemophilia community. By the mid-1980s, methods for inactivation of blood-borne viruses had been implemented in manufacturing processes followed by the first recombinant Factor VIII and Factor IX products in the 1990s. Manufacturers of all Factor VIII and Factor IX concentrates are required to validate two orthogonal virus reduction steps, of which one step must be an inactivation method, such as solvent/detergent or heat treatment.

Biological Activity

ICH Q6B [5] advises that in vitro assessment of biological properties is an essential addition to physicochemical analyses in establishing a therapeutic protein's higher-order structure. For Factor VIII and Factor IX, biological activity is defined by the

potency unit and further characterized to demonstrate conformational fidelity through in vitro functional assays, which measure the full complement of intermolecular interactions responsible for procoagulant activity.

Potency

Potency is the quantitative measure of the product attribute linked to its primary mechanism of action. Assaying functional activities of hemostasis proteins in the complex mixture of patient plasma originally led scientists to define an activity unit for each plasma protein, including Factor VIII and Factor IX, as that amount contained in 1 mL of normal human plasma. Wide variations in the normal human population and the lack of agreement among functional assays performed in different laboratories necessitated the development of an international standardization program, which has been in place for over 40 years, is sponsored by the World Health Organization (WHO), and is facilitated by the National Institute for Biological Standards and Control (NIBSC) in the UK. WHO international standards (IS) for Factor VIII and Factor IX concentrates are prepared by NIBSC, are calibrated through international collaborative studies involving recognized experts from industry and regulatory authorities, and are established by the Expert Committee on Biological Standardization of the WHO [52]. WHO IS are intended as primary reference standards against which each manufacturer should calibrate in-house potency reference standards from well-characterized product lots.

Potency of Factor VIII and Factor IX concentrates may be determined by either of two methods: a one-stage clotting assay or a two-stage chromogenic substrate assay. Both methods were designed to represent the in vivo mechanism of procoagulant function by measuring the outcome following assembly and activation of the complex responsible for activating Factor X, termed the "Tenase" complex. As illustrated in Fig. 2 above, the Tenase complex comprises activated Factor VIII (FVIIIa) as the cofactor, activated Factor IX (FIXa) as the proteolytic enzyme, and phospholipid and calcium for proper assembly on an activated cell surface. The clotting assay was designed to mimic the in vivo mechanism of action by using a plasma substrate deficient in Factor VIII or Factor IX (as applicable) and measuring the correction of a prolonged clotting time as a function of added factor concentration. Variability in the clotting method led to the development of a more purified system designed to directly measure enzymatic activation of Factor X as reflected in the cleavage of a chromogenic substrate. Because reported discrepancies between the two assays have ranged from ~10–50%, with the chromogenic substrate assay generally giving higher values for recombinant products, the most challenging aspect of developing a recombinant Factor VIII product has been the decision on which assay, clotting or chromogenic, is most appropriate for product potency labeling. Clinical practice in the USA supports the use of the clotting assay, while the chromogenic assay is mandated for potency labeling of products in Europe [53]. Current development programs should include both assays for product characterization and clinical development with a decision on which assay will ultimately be used for commercial release and stability dependent upon analysis of all development data [54]. Although the clotting assay is currently used for

potency assignment of all licensed Factor IX concentrates, comparative assessment of chromogenic and clotting activities is recommended for complete characterization of recombinant Factor IX products since assay discrepancies have been reported [55].

In Vitro Functional Tests

Many of the in vitro functional tests currently used to assess Factor VIII or Factor IX quality are based on reconstitution of the Tenase complex and monitoring the generation of Factor Xa detected by a chromogenic substrate. By varying component concentrations and experimental conditions and performing kinetic analyses, the affinity of molecular interactions and degree of in vitro procoagulant functionality can be characterized [56]. Factor VIII and Factor IX function also depend on binding to phospholipid, which can be measured by surface plasmon resonance or Factor Xa generation assays. Factor VIII binding to von Willebrand factor (critical for Factor VIII stability in circulation) can also be measured using surface plasmon resonance assays. The activation of Factor VIII or Factor IX is intrinsic to in vivo function. In vitro assays have been developed for evaluating kinetics of activation and subsequent inactivation of Factor VIII by thrombin or Factor Xa. Similar assays have been developed to characterize the kinetics of Factor IX activation by Factor XIa. Activated protein C confers its anticoagulant activity through direct cleavage of Factor VIII; therefore, in vitro assays designed to specifically characterize activated protein C inactivation kinetics for Factor VIII are included in complete in vitro functional characterization strategies [57]. Thrombin generation assays, which measure kinetics of thrombin generation in complex systems from whole blood to defined combinations of purified proteins, are gaining wide use as global assays for evaluating product quality and in vivo performance [58]. In general, the ability of in vitro functional tests to assess product quality relies on comparative testing of the product under development to a licensed comparator. In vitro functional tests are highly valuable in supporting manufacturing changes to licensed products by comparative testing of pre-change to post-change product.

Immunogenicity

Preservation of native protein conformation is key to maintaining product safety and efficacy. Failure to implement analytical programs to guarantee retention of protein conformation for the duration of a product's shelf life may result in loss of therapeutic benefit or increased risk of immunogenicity. The development of Factor VIII or Factor IX "inhibitors," antibodies against the Factor VIII or Factor IX molecule, remains the major negative clinical consequence of hemophilia treatment. Factor VIII inhibitors develop in ~30% of patients with severe hemophilia A, and Factor IX inhibitors, some with anaphylactoid consequences, develop in ~5% of patients with severe hemophilia B. Despite suggestions that recombinant products may confer higher immunogenic risk, there is no conclusive evidence, to date,

showing a difference between plasma-derived and recombinant product immunogenicity [59]. Reports in the early 1990s of increased Factor VIII inhibitor incidence resulting from Factor VIII products denatured through certain viral inactivation processes highlighted the importance of robust analytical programs designed to demonstrate retention of protein conformation and clinical programs suitably designed to monitor for inhibitor development [60].

4.2.3 New Generation Products

New generation Factor VIII and Factor IX proteins bioengineered for better clinical performance have been licensed, and others are in development. PEGylation, Fc fusion, and albumin fusion strategies have been successfully applied to create Factor VIII or Factor IX analogues with increased circulating half-lives [61]. Development and life cycle management of these novel proteins have warranted the implementation of equally novel and sensitive analytical methods to ensure retention of full functionality without increased risk of immunogenicity. Particularly challenging can be the development and validation of suitable functional assays for quality control of the non-coagulation fusion moieties. The agency encourages manufacturers to qualify novel, product-specific, and sensitive new technologies to enhance product knowledge and understanding, facilitating Quality by Design approaches to life cycle maintenance of product quality.

4.3 Monoclonal Antibodies

4.3.1 Introduction

The first therapeutic mAb, OKT3, was licensed in 1986 for the treatment of acute renal allograft rejection. Although it was an effective treatment, a majority of patients made anti-drug antibodies (ADA) against this murine mAb, which blocked its therapeutic effectiveness [62]. In general, murine mAbs are immunogenic, have a short half-life, and are inefficient at eliciting effector functions in humans [63]. The therapeutic potential of mAbs began to be realized in the late 1990s after the introduction of chimeric and humanized mAbs, which contain human Fc regions. Subsequently, mice engineered to express human antibodies and phage display technology introduced human mAbs as clinical candidates. Since the early 2000s most mAbs entering clinical development are humanized or human. Of the 23 mAbs approved since 2009, 14 are human, 5 are humanized, 3 are chimeric, and 1 is murine.

4.3.2 Selecting and Engineering MAbs for Specific Applications

The majority of mAbs in development continue to be intact mAbs, mostly IgG1, but IgG2 and IgG4 mAb are used when the mAb is designed to have little to no effector function. In addition to intact mAbs, the diversity of mAb and related products includes antibody-drug conjugates (small-molecule drugs and radioimmunoconjugates), other types of antibody conjugates (bacterial toxins, enzymes, cytokines), antibody fragments (Fab, sFv, single V domain constructs), bispecific antibodies (full length or based on sFv or single V domain fragments), antibody cocktails, and Fc-fusion proteins. To date, the agency has approved ten Fc-fusion proteins, three antibody-drug conjugates, two therapeutic radioimmunoconjugates, and one bispecific mAb.

Intact antibodies have two functional domains, the Fab region for binding to antigen and the Fc region, which imparts effector function. Knowledge of Fc-glycan structures and amino acid residues in the Fc region that play a role in effector functions, such as complement-dependent cytotoxicity (CDC), antibody-dependent cellular cytotoxicity (ADCC), and antibody-dependent cellular phagocytosis (ADCP), led to engineering approaches that can enhance or reduce specific effector functions [64–72]. Engineering either the Fab or Fc regions can enhance the pharmacokinetic properties of a mAb [73–76].

Although the IgG3 isotype elicits high levels of effector function, especially CDC, it is not commonly used for therapeutic mAb, because it is susceptible to proteolytic cleavage and has a shorter half-life compared with the other IgG isotypes. However, specific mutations of amino acid residues in the IgG3 hinge region are being used to take advantage of IgG3 effector function properties [77].

4.3.3 Common Manufacturing Considerations to Ensure MAb Quality and Consistency

Posttranslational Modifications (PTMs)

Each mAb is subject to posttranslational modifications (PTMs) and degradation over time. One advantage of intact mAbs is that many PTMs in conserved regions and degradation products are well understood and apply to all products of this class. Many of these PTMs and degradation products occur naturally in endogenous IgG [78]. The risks to patient safety of these common PTMs are generally understood. Therefore, the challenge for sponsors is to evaluate the risks of PTMs and degradation products that are unique to each product.

N-Terminal Pyroglutamic Acid and C-Terminal Lysine

Many mAb PTMs result in charge variants that have the potential to impact stability or biological activity. Two of the most common PTMs resulting in basic charge variants are due to N-terminal glutamine and C-terminal lysine. These amino acid residues are often processed from N-terminal glutamine to pyroglutamic acid (PE) or the C-terminal lysine is cleaved, resulting in more acidic species. These PTMs have no known impact on mAb function or pharmacokinetics (PK) [79, 80].

The conversion of N-terminal glutamine to PE occurs spontaneously and mostly in the bioreactor; however, buffer composition and temperature can also lead to this conversion [81]. N-terminal glutamate can also be converted to pyroglutamic acid but does not result in a charge difference. Therefore, other methods such as LC/MS/MS are needed to characterize this variant.

The conversion to PE by either pathway, as well as C-terminal lysine cleavage, occurs in vivo and reflects normal modifications of both therapeutic mAbs and endogenous IgG [78, 79, 82–84].

C-terminal lysine cleavage occurs rapidly upon administration to patients [85]. As a quality attribute, it can be a predominant source of charge heterogeneity, with 0, 1, or 2 C-terminal lysine residues per molecule. To assess the contribution of C-terminal lysine, as well as other PTMs, to charge heterogeneity, samples can be treated with carboxy peptidase B to remove the C-terminal lysine [84]. A comparison of the charge profile of untreated and treated samples reveals the charge heterogeneity due to the presence of C-terminal lysine or other PTMs.

Proline Amidation

A recently identified PTM resulting in basic species is the amidation of the penultimate proline residue after C-terminal lysine cleavage [86]. This modification is fairly common among mAbs of different IgG isotypes and is catalyzed by peptidylglycine alpha-amidating monooxygenase in cell substrates after C-terminal lysine cleavage [87]. The reaction is sensitive to levels of copper in the bioreactor [88]. This PTM does not appear to impact biological activity or PK [79, 86].

Deamidation

Asparagine deamidation leads to aspartic acid or isoaspartate, which are found in acidic species of mAbs [79, 89]. The rates of conversion vary for each asparagine and are influenced by the surrounding sequence. Some conserved asparagine residues in the Fc region of IgG1 and IgG1 can be deamidated, with Asp384, being the most susceptible [90]. Endogenous antibodies are also deamidated at this site [78, 90]. The levels of deamidation at these constant region sites are not known to have an impact on biological function. Deamidation in either heavy chain

or light chain CDRs has also been reported, and the impact on binding to antigen can vary from little to significant impact on biological activity [89–95].

Oxidation

Oxidation occurs predominantly at methionine residues but can also occur at tryptophan, cysteine, histidine, and cysteine residues. Oxidation of methionine in the heavy chain constant region (Met256 and Met432) can reduce binding to protein A and G [96] and FcγRIIa 131H and FcRn [97, 98]. However, the levels of oxidation that reduce binding to these receptors are generally higher than typical levels due to manufacture and product shelf life and do not have an impact on PK of a mAb when compared with an unmodified IgG [99].

There are examples of oxidized residues in heavy chain or light chain CDRs that decrease binding to the antigen. Oxidized methionine and cysteine in the VH region of OKT3 impact binding to CD3 [91]. Oxidation of tryptophan, but not methionine, in an anti-RSV mAb resulted in loss of activity [100].

Glycation

Glycation occurs on lysine residues during manufacture or storage when reducing sugars are present in culture media or formulations. Because a mAb generally contains many lysine residues, glycated forms, which are acidic species, can be extremely heterogeneous. Glycation also occurs in vivo on both endogenous antibodies, as well as therapeutic mAbs, and does not appear to affect PK or effector function [79, 101]. For example, mAbs with up to 10 or 17% glycation in a heavy chain CDR [79, 102] or near complete glycation through forced glycation studies [103] did not have reduced biological activity. Forced glycation studies also showed that high levels of glycation (~42%) did not decrease binding to FcγRIIIa, FcRn, or protein A [101].

Glycosylation

Antibodies have a single N-linked glycosylation site at Asn297 in the Fc CH2 domain. The biantennary glycan structure contains the core heptasaccharide, and additional sugars attached to the core heptasaccharide result in heterogeneous structures. The predominant forms in unengineered cell lines are generally fucosylated, and galactose may or may not be present (G0, G1, and G2). Terminal sialic acid; afucosylated G0, G1, and G2; and high-mannose forms are generally present at low levels but can usually be quantitated by mass spectrometry, capillary gel electrophoresis, and HPLC with fluorescence detection. Other structures can be detected but usually are present at such low levels that they cannot be quantitated.

Although all species, including plants, synthesize the same core heptasaccharide with the same additional monosaccharides, there are species-specific differences [104, 105]. For example, there are species-specific differences in terminal sialylation and galactosylation [104], and relative to endogenous human mAbs, therapeutic mAbs produced in mammalian cell lines are hypogalactosylated [106]. Another difference is that human glycan structures can have a bisecting GlcNAc, which inhibits the addition of fucose. CHO cells do not express *N*-acetylglucosaminyltransferase-III, which is the enzyme that carries out this reaction [104].

Specific glycan structures play a role in antibody effector function. The complete removal of the glycan results in the loss of both CDC and ADCC activity [107]. Galactose is associated with CDC activity such that its removal results in a significant, but not complete reduction of activity [107, 108].

The presence of fucose inhibits ADCC activity [109, 110] specifically when mediated through FcγRIIIa. IgG and FcγRIIIa interact through a rare carbohydrate to carbohydrate interaction, with which fucose on the Ig glycan interferes [111, 112]. However, a mAb where the glycan is fucosylated on one constant region chain and is afucosylated on the other has enhanced ADCC activity relative to a fully fucosylated mAb but still has about 50% of the activity of a fully afucosylated mAb [113]. For mAbs where ADCC is a desired mechanism of action, CHO cell lines have been engineered to knock out the fucosyltransferase gene [114–116] or by adding genes that express *N*-acetylglucosaminyltransferase-III and α-mannosidase II, which block the addition of fucose by promoting structures containing a bisecting GlcNAc [117].

Heavy and light chain variable (V) regions can also contain N-linked glycan structures [118], which may enhance or interfere with antigen binding [119]. Some germ line gene VH and VL sequences contain the consensus glycan attachment Asn-X-Ser/Thr sequence, but others can be generated through the somatic mutation process [120]. V region glycosylation has been associated with various diseases, for example, 79% of VH sequences in follicular lymphoma patients had novel glycan attachment sites in CDRs [121], and V region glycosylation of autoantibodies can influence antigen binding and possibly the pathogenic nature of these autoantibodies [122]. For therapeutic mAbs that have potential V region glycosylation sites, it is important to determine if the Fab is glycosylated and, if present, the structures of the Fab glycan.

Cetuximab, first approved in 2004 for patients with metastatic colorectal cancer and subsequently for head and neck cancer, is glycosylated on the VH region. The glycan structure contains a galactose-α-1,3-galactose (α-gal) linkage, which is produced by the murine cell substrate. This structure was shown to induce anaphylaxis in patients who had preexisting IgE anti-α-gal antibodies prior to treatment [123].

4.3.4 Product-Related Impurities

Antibody fragments and aggregates are critical quality attributes of intact mAbs that can impact the potency of the product. V region-containing fragments generally maintain their ability to bind antigen. However, if effector function is part of the mechanism of action, both V region and Fc region fragments will have reduced potency. Aggregates often have reduced antigen-binding capacity but may demonstrate enhanced potency if presented as immune complexes to Fcγ receptors. Therefore, it is necessary to understand the causes of these degradation pathways in order to improve the manufacturing process to further remove fragments and aggregates from the drug substance and to optimize the formulation to reduce their formation over the shelf life of the product.

Fragments

The hinge region of IgG is susceptible to cleavage by proteases and nonenzymatic cleavage. Fragmentation of mAbs at the hinge region can occur in vivo by exposure to endogenous or bacterial proteases [124]. Residual proteases in host-cell proteins can also lead to fragmentation [125]. However, most of the fragmentation of mAbs is nonenzymatic and can occur during manufacturing or under normal storage conditions, although temperature and pH stress enhance the chemical reactions. Several studies investigated chemical reactions of nonenzymatic cleavage, such as peptide bond hydrolysis or β-elimination. These reactions typically have sequence specificity [126]; see [127] for a review). Different pH and temperature conditions can lead to fragmentation, but the presence or absence of the Fc-glycan structure does not [128, 129]. The rates of fragmentation were shown to be similar for five different IgG1 mAbs, suggesting that fragmentation is largely dependent on the primary sequence of the hinge region [129].

The presence of copper or iron in the presence of histidine can also result in nonenzymatic cleavage. Cleavage of alemtuzumab at 37°C was due to the presence of copper and was further elevated at high pH. This fragmentation was also temperature dependent; no fragmentation occurred below 20°C but increased at temperatures above 37°C [130]. Interestingly, in IgG mAbs containing a lambda light chains, cleavage was observed at a different site within the hinge region, and there was also cleavage of the lambda light chain. These reactions were due to the presence of iron in the drug substance. However, iron alone did not induce the cleavage but required the histidine in the formulation buffer for this reaction. This cleavage reaction was not seen in IgG mAbs with kappa light chains [131].

Aggregates

Antibody aggregates can form during upstream or downstream manufacturing, as well as during the shelf life of the product [132, 133]. Thermal, freeze-thawing, pH, light, and mechanical stress are known to induce aggregation of mAbs. Aggregation can also occur in the presence of human plasma, which is dependent on the diluent used for preparation of the infusion for intravenous infusion [133]. Formulation optimization studies are crucial for the long-term stability of mAbs.

Antibodies naturally contain aggregation-prone motifs, which may play a role in their ability to bind antigen, FcγR, or protein A [134]. However, these motifs play a role in aggregation of mAbs over their shelf life. Specific aggregation-prone amino acids in either the V region or C region can be engineered to maintain antigen binding and other functional properties of the mAb while reducing aggregation [134, 135].

Several studies applied different forms of stress to mAbs to understand the factors that contribute to aggregation. In a study by Joubert et al. [136], different forms of freeze-thaw, pH, mechanical, chemical, and thermal stress were applied to IgG2 mAbs, human and mouse IgG1 mAbs, and IVIG. Aggregates were characterized by orthogonal methods and divided into seven classes based on the biophysical properties of the aggregates, such as total aggregation, size, morphology, etc. Similar classes of aggregates were formed in each of the samples when treated by the same stress conditions. This suggests that specific types of aggregates are formed by specific types of stress, which is consistent with an earlier study [137] that showed the Fab fragment is more sensitive to heat stress, while the Fc region is more susceptible to low pH stress. On the other hand, in a study by Nicoud et al. [138], stressed IgG1 and IgG2 mAbs under identical conditions showed differences in the kinetic mechanisms of aggregate formation.

The simplest form of aggregates is non-covalent reversible dimers, which are seen under normal storage conditions. However, some dimers are covalent, and while some maintain activity, others lose activity. A study on epratuzumab showed that 70% of the dimers were covalent with three different forms; Fab/Fab dimers, Fab/Fc dimers. and Fc/Fc dimers. The Fab/Fc dimers were the predominant form [139], but all forms maintained binding activity. A different study [140] treated an IgG1 with three different forms of stress that all resulted in Fab/Fab dimers. Dimers generated by two of the stress conditions had reduced potency in a cell-based assay, but the dimers from all three stress conditions had enhanced binding to FcγRIIIa by surface plasmon resonance. A study by Luo et al [141] also showed enhanced binding of dimers to FcγRI, FcγRII, and FcγRIII.

4.3.5 Mechanism of Action

Understanding the mechanism(s) of action (MOA) of a mAb is a continuous process and may depend on the indication, new knowledge of the target biology, and antibody-Fc-receptor interactions. Antigen binding is a critical quality attribute of all mAbs. Whether antibody effector function plays a role in the mechanism of

action for a given mAb depends on many factors including whether the target is expressed on a cell surface or is soluble; the presence of an Fc region; the IgG isotype; or specific engineering of the Fc region to enhance or reduce effector function. Aspects of antigen binding and the potential to bind FcγR and complement should be part of a thorough characterization of a mAb. Potency assays that reflect the predicted mechanism of action should be developed for release and stability testing. When effector function is expected, a cell-based potency assay with appropriate target and effector cells demonstrated both antigen binding and effector function. Methods that bridge binding to the target as well as FcγR can also be developed as characterization and release methods [142].

Antigen Binding

Binding to the antigen is an MOA for every mAb, which should be characterized for binding affinity and on and off rates. Binding to cell surface antigens on appropriate target cells should be demonstrated. Homology between the human target and primate or other species is important in order to determine the relevant animal models for preclinical studies.

Some mAbs against soluble ligands may work simply by blocking the binding of that ligand to its receptor, resulting in the inhibition of downstream signaling effects. The mechanism of the anti-CD25 mAb, daclizumab, first approved for the prophylaxis of kidney graft rejection, was believed to be due to blocking IL-2-mediated signaling of activated T cells. However, in multiple sclerosis, additional mechanisms resulting from IL-2 neutralization have become apparent, such as the activation and expansion of regulatory NK cells, blocking of dendritic cells from presenting IL-2 to primed T cells, and modulation of the development of innate lymphoid cells [143].

Other mAbs against soluble targets, such as the TNFα antagonists, are thought to work primarily through TNFα neutralization; however, these agents can also bind to membrane forms of TNFα which can lead to apoptosis and cytokine suppression through reverse signaling and may include antibody effector functions such as CDC and ADCC [144].

Effector Function

There is a hierarchy among IgG isotypes regarding their ability to carry out effector function. IgG1 and IgG3 isotypes are chosen when effector function is desirable and IgG2 and IgG4 isotypes are generally chosen when effector function is not. Jiang et al. [145] published a reasonable approach describing the characterization of potential effector function and developed a control strategy for mAbs that includes effector function as part of their MOA, which is dependent on whether the mAbs are anticipated to have high, intermediate, or low potential for effector function.

In general, there is an expectation that the characterization of mAbs with a low potential for effector function should be performed early in development. No additional studies would be needed unless new information became available suggesting that effector function may contribute to the MOA. For mAbs where effector function is expected to contribute to the MOA, there is an expectation that characterization will include relevant cell-based bioassays as well as bind to complement and FcγRs. Analysis of glycoforms (galactose and fucose) that impact effector function should also be included in the characterization. The control strategy should ideally include a bioassay representing the most relevant proposed MOA.

Until recently, most potency assays for mAbs with effector function were CDC assays, because this is generally a robust method that can meet requirements for appropriate validation and quality control. While CDC potency assays represent a relevant biological activity, it is thought that for many mAbs, ADCC and ADCP may be more important MOAs in patients, especially oncology patients [146–148].

ADCC assays have been included in characterization studies, but not as potency assays, because the effector cells were peripheral blood mononuclear cells (PBMCs) derived from human donors. Because of donor-to-donor variability, these ADCC assays were not amenable for quality control purposes. However, over the last 5–10 years, NK cell lines have been developed that replace PBMCs as donor cells in the ADCC assay [149], and when available to a sponsor, these assays are now incorporated as part of the control strategy. Reporter gene assays have also been developed as a surrogate for an ADCC assay [150, 151], but these methods are not a direct measure of ADCC activity and should be demonstrated to be representative of the mAb's ADCC activity.

ADCP assays for characterization are just beginning to be developed. These methods use human PMBCs differentiated into macrophages or murine macrophages [148, 152–154] and are not amenable as quality control methods. However, ADCP methods using cell lines as effector cells are under development, and we expect they will be used first for characterizing mAb effector function and eventually as release and stability methods.

Along with developing relevant antibody effector function characterization and potency assays, understanding the most relevant effector function in any given patient population is a challenge. Different effector cell types have varying levels of activity in vitro [151] and this may also reflect in vivo efficacy. Different effector cells are present at different sites of disease, and other than NK cells, which express only FcγRIIIa and FcγRIIc, other effector cells, such as monocytes, macrophages, dendritic cells, and neutrophils, express all the FcγR [155, 156].

Other aspects of the antigen/mAb interaction influence effector function activity. Higher levels of EGFR expression correlate with Fc-mediated effector function in vitro [157]. Of particular note, in vitro ADCC studies with cetuximab using NK cells from healthy donors and squamous cell carcinoma of the head and neck (SCCHN) and colorectal cancer cell lines correlated with the high affinity FcγRIIIa V/V polymorphism [158]. However, a more recent study using NK cells derived

from SCCHN patients showed that the FcγRIIIa genotype was not associated with the clinical outcome but rather had an innate cytotoxicity capability [159].

The intracellular domain of the antigen also appears to play a role in effector function mechanisms. Using intact CD19 and HER2 and chimeric constructs swapping the intercellular domains of the molecules, Tiroch et al. [160] showed that mononuclear cells could trigger ADCC with the appropriate mAb for all the constructs, while polymorphonuclear granulocytes were only effective against wild-type HER2 or the CD19/HER2 intracellular domain chimeric protein.

Overall, while relevant bioassays may be available for characterization and release of mAbs, there may not be a complete understanding of the MOA for specific indications. Since there are many effector cell types with different expressions of FcγRs, methods that assess binding of the mAb to all the FcγRs are important for characterization and comparability exercises. Focusing on afucosylated glycans and NK cell ADCC activity may not always be a reflection of the in vivo MOA.

4.3.6 IgG Isotype-Specific Characterization

IgG2 MAbs

The IgG2 isotype is often chosen when effector function is not intended as a mechanism of action for a therapeutic mAb. Unlike the other IgG isotypes, human IgG2 contains four cysteine residues in the hinge region, which result in different IgG2 disulfide isoforms, termed IgG2-A, IgG2-A/B, and IgG2-B [161, 162]. These disulfide isomers occur in endogenous and myeloma-expressed IgG2, as well as in therapeutic mAbs. However, for some therapeutic mAbs, the isoforms may have different potency in cell-based assays [163].

The disulfide isomers were subjected to site-directed mutagenesis, and it was shown that specific cysteine to serine mutations would reduce the disulfide heterogeneity while maintaining in vitro activity [164]. However, disulfide isomers may also be exploited for enhanced activity. A recent study showed that IgG2 mAb, in particular the IgG2-B disulfide isoform, provides superior FcγR-independent activity relative to other isotypes to immunostimulatory mAbs [165].

In addition to the disulfide isoforms, IgG2 can also form covalent dimers, which are thought to increase the avidity of the IgG2 response against bacterial capsular polysaccharides [166]. Therefore, therapeutic IgG2 mAbs should be thoroughly characterized for disulfide isomers and covalent dimers, which should be controlled if they demonstrate different in vitro potency. Capillary gel electrophoresis [167] and capillary zone electrophoresis [168] methods can distinguish the disulfide isoforms.

IgG4 MAbs

The IgG4 isotype is also chosen when effector function is not an intended mechanism of action. Although IgG4 has two interchain disulfide bonds in the hinge region, the specific sequence surrounding the two cysteine residues on each H chain results in an unstable hinge region, which leads to the formation of IgG4 half antibodies (or monomers). This property is associated with the presence of the serine 228 residue immediately prior to the second cysteine residue ([169] and references therein). The result of half-antibody formation is that the IgG4 monomer can undergo Fab arm exchange to form bispecific IgG4 heterodimers with other IgG4 antibodies present in serum [170]. Although additional sequences in the CH3 domain have also been demonstrated to play a role in Fab arm exchange [171], mAbs containing a serine to proline mutation at position 228 do not undergo Fab arm exchange in vivo [172, 173].

It has been proposed that IgG4 half antibodies may have advantages over mAb fragments when it comes to their half-life. To this end, specific mutations in the CH3 domain were identified that result in monovalent IgG4 that may be useful for clinical development [174].

IgG4 half mAbs, with or without the serine to proline mutation at position 228, should be characterized and controlled throughout product development. Western blot [175], chip-based SDS-PAGE [176], and capillary SDS [177] methods have been developed to detect IgG4 half antibodies. Methods that provide quantitative results are preferred.

4.3.7 Future Trends in MAb Development

Therapeutic mAbs are currently the largest class of biologics in development. This is due in part to the success of antibody engineering the Fc region to reduce immunogenicity, the ability to use platform manufacturing processes, and a good understanding of general quality attributes. Better understanding of disease pathways and the identification of new targets, in combination with continued engineering of Fc regions, glycan structures, and development of novel constructs, such as antibody conjugates, bispecific products, mAb cocktails, and Fc-fusion proteins, provide a robust pipeline of mAbs and related products. Finally, we anticipate that in the next few years, there will be a better understanding of how different effector cell types respond to mAb-based immune complexes through different FcR.

4.4 Other Therapeutic Proteins

Therapeutic proteins encompass a wide variety of proteins products including replacement enzymes, cytokines, hormones, and toxins. Many expression systems such as bacteria, yeast, plant and mammalian cells, transgenic animals, and even natural sources are used to manufacture these proteins. The purification processes

are generally tailored to characteristics of each product and are aimed at maximizing the capture of the target and removal of process- and product-related impurities. Due to the lack of a single robust affinity purification step such as the protein A column for monoclonal antibody purification, the purification process for the majority of therapeutic proteins varies depending on the property of each protein. In addition, structural modifications or variations for each protein are specific to the protein, and the impact by these modifications on product quality can vary from product to product. There is no "platform" approach toward characterization of all therapeutic protein products. Here, we discuss some general guidelines for characterizing these products and how appropriate controls can be put in place throughout product life cycle to ensure their safety and efficacy.

4.4.1 Manufacturing Controls to Ensure Therapeutic Protein Product Quality

Identity

Historically various methods have been used as an identity test. The current regulatory expectation is that an HPLC chromatography method interrogating digested peptide patterns be used at drug substance release. The landmark peptide peaks should be identified by mass spectrometry analysis during product development. A robust test that provides an unambiguous identification is particularly important for products manufactured in a multiproduct manufacturing facility to prevent potential misidentification of products.

Purity

Because therapeutic proteins are heterogeneous, purity of a product is generally measured by multiple methods. The impurities should be interrogated by size, charge, and hydrophobicity. Each minor peak or band representing impurities, degradation products, or product variants should be identified and quantified based on manufacturing history.

Potency

In biological product testing, in vitro potency assays are considered a surrogate to clinical efficacy and are therefore an integral component of product quality testing as well as to monitor the consistency of the product throughout the life cycle of the product. Although by convention, potency assays are interpreted as a measurement of the biological activity of the product, such as the enzymatic activity of an enzyme, a well-designed potency assay should encompass all aspects of biological characteristics that collectively define the mechanism of action. Depending on the

class of products, the scope of potency assays also varies depending of the factors involved in the product to act. Therefore, in addition to biological activity assays, potency assays should also include testing for affinity to ligands, such as receptors and substrates as well as tests for attributes that impact circulation half-life.

Enzymatic Assays

Enzyme activities are impacted by two integral factors: the affinity of the enzyme to its substrate and the ability of the enzyme to catalyze the conversion of its substrate to the final product. Therefore, enzyme activity assays should cover both these aspects of any given enzyme product. Generally, this can be accomplished with an enzyme kinetic assay. In addition, to better represent the efficacy of the enzyme for the indication, the substrate(s) should resemble the property and structure of the natural substrate relevant in the indicated disease(s).

For the majority of currently approved enzyme products, enzyme activity assays are performed using a small-molecule synthetic substrate conjugated to a colorimetric function group or a fluorochrome. The assays are commonly conducted at saturating levels of substrate that generally do not represent levels of the natural substrate under the indicated disease conditions. Furthermore, the small molecules generally only represent the linkage or group specificity but lack the structural representation of natural substrates of the enzymes; the results therefore may not be representative of the in vivo enzymatic activity. Numerous studies have demonstrated that the K_M and catalytic efficiency (k_{cat}/K_M) obtained from the synthetic substrates can differ significantly from these obtained using natural substrates [178, 179]. Whereas enzymatic assay measuring the end point product of enzymatic reaction using a synthetic substrate may be appropriate for a rapid control for process intermediates, an enzymatic kinetics assay using a substrate that resembles the structure and linkage specificity is expected for release and stability testing for both drug substance and final drug product. Enzyme kinetic assays are also expected in post-licensure comparability and similarity studies.

Cell-Based and In Vivo Potency Assays (Bioassays)

For growth hormones and cytokines, the potency bioassays are aimed at measuring biological activity of the product in vitro using mammalian or other relevant cell culture systems, as well as in vivo in animals. For certain products, in vitro cell-based potency assays alone are insufficient to demonstrate clinical efficacy as they do not take into account product attributes affecting pharmacokinetics, organ distribution, or metabolism. For example, in addition to the ability to stimulate red blood cell production, clinical performance of recombinant erythropoietin relies heavily on the glycan structure on the molecule [180]. As a result, the potency assay for many erythropoietin products is still performed as mouse-based in vivo reticulocyte proliferation assays. However, in cases where biological activity and attributes

affecting half-life are well known, the in vivo potency assay can be replaced by a combination of assays independent testing these attributes. For example, a combination of in vitro cell-based potency assay and glycan analysis has been shown to be adequate as a control for erythropoietin products in lieu of the reticulocyte proliferation in vivo assay [181].

Receptor/Ligand Binding Affinity Assays

Enzymes, cytokines, and hormones generally act in targeted tissues or organs through binding to their corresponding receptors on the surface of cells. Proper control of the binding affinity of these proteins to their respective receptors is critical to ensure clinical performance of these protein products. The affinity of these products to their receptors should be well characterized during product development and tested at product release and during stability testing. Whereas cell-based biological activity assays provide some information for the binding of the products to their receptors, due to the inherent variability of these types of assays, they are generally suboptimal in monitoring alterations of binding affinity that may occur after manufacturing changes. An alternative method that directly assesses the affinity of the products to their respective receptors is necessary to measure this important aspect. This can generally be accomplished by in vitro binding assays where the equilibrium dissociation constant (K_D) of the product and its receptor/ligand is measured. Recent advances in surface plasmon resonance (SPR) technology make this analytical technique among the most commonly used method for measuring protein affinity to their targets [182, 183]. These assays are expected to be part of release and stability testing at the time of licensure. Sponsors of biological products are encouraged to develop them early in product development cycle and ideally implement a validated receptor/ligand binding assay for phase III clinical materials.

Assays Measuring Attributes Affecting Half-Life

There are many factors acting either independently or collectively to affect serum circulating half-life of a biologic product, many of which have been exploited by manufacturers to achieve better therapeutic effects. These include managing cell culture conditions to control for glycosylation, covalent conjugation of polyethylene glycol (PEG) molecule(s) to proteins, and fusion of various protein tags to proteins. This section discusses expected testing to confirm the consistency of these attributes.

Sialylation

The majority of therapeutic proteins and monoclonal antibodies expressed in mammalian cell lines are glycosylated to variable extent, at select Asn residues. The terminal monosaccharide of these glycans may be capped by sialic acid to various degrees. Numerous studies have shown that the total sialic acid content and, more profoundly, the structure of sialic acid-capped *N*-glycans significantly affect glycoprotein absorption to and clearance from serum [184–186]. Characterization and quantitation of sialylation are, therefore, crucial in maintenance of product consistency. Because the majority of therapeutic proteins contain multiple *N*-glycosylation sites, and each site is generally glycosylated at different degrees, the overall glycan and sialic acid content on protein molecules can be extremely heterogeneous. More importantly, levels of sialylation can be affected by many cell culture parameters such as pH, level and composition of nutrients, cell culture additives, cell growth rate, dissolved oxygen, and temperature [187]. Due to all these factors, from a product life cycle management perspective, analysis of sialic acid and sialylated glycan structures should be implemented as early as possible in product development. The tests for sialylation generally include relative sialic acid content (i.e., mol sialic acid/mol of protein) and quantitation of sialylated glycan structures. Due to sialylated glycans being negatively charged (with the net charge proportional to the number of sialic acid on each glycan), the latter often consists of quantitation of neutral, mono-, di-, tri-, and tetra-sialylated glycans cleaved from the protein molecules. Glycans are generally separated by capillary electrophoresis or anion exchange HPLC followed by mass spectrometry identification of the contents of each individual peak. Extensive charge profiling is expected as part of product characterization, whereas quantitation of critical glycan group(s) and sialic content should be part of release testing. However, for release testing, manufacturers can opt for a method that is highly reproducible, precise, and easy to operate in a QC environment.

PEGylation

For small proteins and proteins expressed from prokaryotic expression systems, conjugation of one or more polyethylene glycol (PEG) molecules can effectively increase half-life and reduce immunogenicity of the products [188]. However, PEGylation, especially multiple covalent attachments of PEG molecules to primary amines on a protein molecule, can generate undesirable effects including reduction in biological activity, reduction in affinity to receptors, and lot-to-lot variations for some products. To minimize these undesirable effects, the majority of manufacturers have moved away from random PEGylation to site-specific mono-PEGylation [189]. Owing to continuous efforts in optimization in PEG production, linker development, and PEGylation reaction conditions, PEGylation of therapeutic proteins has evolved into a very robust and reliable process, and the resulting PEGylated products are fairly homogenous and stable. The characterization

generally includes conjugation site identification by mass spectrometry and quantitation of non-, mono-, and di/multi-PEGylated species by a column chromatography.

Another source of variability in PEGylated products derives from the heterogeneous composition of the PEG starting material, so particular focus early in the development is placed on the qualification of a consistent PEG supply. As a critical raw material for PEGylated products, the qualification of PEG should be comprehensive and should include tests for purity, functional group activity, polydispersity, and stability testing. In some cases where PEGs from different suppliers are used, in addition of comparison of PEG manufacturing process and lot release data, the qualification of each PEG may require some nonclinical study comparing protein PEGylated with the PEGs from different sources.

In many instances, conjugation of even one PEG molecule interferes with critical attributes such as enzymatic activity or receptor binding affinity. Manufacturers are expected to fully characterize and control for site of PEGylation by mass spectrometry and quantitate single-, di-, multi-, and non-PEGylated proteins through SEC- or EX-HPLC. The manufacturers are also expected to characterize hydrolysis of PEG moiety from the protein under relevant storage conditions. These quantitative assays should be included in release and stability programs.

Polypeptide Tags

Another alternative used by manufacturers to extend serum half-life of protein biologics has been to add another protein or polypeptide tag at the amino- or carboxy- terminus or both termini. Such tags include carboxyl terminal peptide [190], human serum albumin [191], and XTEN [192]. They generally fold as a stand-alone unit separated structurally and functionally from the therapeutic protein, often not contributing or contributing little to biophysical characteristics of the chimeric construct other than size. The major control for these products focuses on the integrity of the tags in manufacturing process and during storage as the junction of the two separately folded structures tends to be targeted for degradation.

4.4.2 Posttranslation Modifications (PTMs) on Protein Products

Glycosylation occurs as part of protein biosynthesis, and, for the majority of protein products, e.g., products manufactured using eukaryotic cell expression systems, is considered the most important of the PTMs since differences in glycosylation have been shown to impact critical attributes including circulation half-life, affinity to receptors, and effector functions. The glycans collectively affect the structure and function of the protein, and even relatively small variations in overall glycosylation profile can have drastic effects on the performance of therapeutic proteins [193]. Due to the high degree of heterogeneity, the function of each individual glycan is hard to measure. However, mapping of overall glycan profile can provide some

details on glycosylation for the purpose of maintaining product consistency during process development. Such analyses include monosaccharide composition, overall glycan profiling, and site-specific glycan analysis and are expected to be included as part of characterization studies and incorporated as part of comparability studies to support post-licensure manufacturing changes. Recent advances in mass spectrometry have enabled identification of microheterogeneity and relative abundance of glycans on specific glycopeptides [194, 195].

Among the various glycan forms, glycans bearing mono-mannose-6-phosphate (M6P) or di-M6P (bis-M6P) are of particular importance for products used as enzyme replacement therapy (ERT) for lysosomal storage diseases [196]. M6P- and bis-M6P-containing glycans serve as primary mechanism in targeting these enzymes to lysosomes through binding to cation-dependent and cation-independent M6P receptors [197, 198]. The presence of M6P residues on the *N*-glycans of these recombinant therapeutic proteins is a critical quality attribute since increasing levels of M6P on ERT products has been observed to positively correlate with more efficient uptake and subsequent treatment efficacy for lysosomal storage diseases both in vitro and in vivo [199, 200]. Despite a decade of efforts, characterization of mono- and especially bis-M6P-bearing glycans still poses a considerable analytical challenge. Many of the methods used so far are too complex and require highly purified material and extensive post-testing data analysis, which has significantly limited their utility in upstream process optimization and monitoring. A recent study employed a combination of ultra-performance liquid chromatography (UPLC) and capillary electrophoresis with laser-induced fluorescence detection (CE-LIF) for a rapid separation of *N*-glycans. The method presents a novel yet relatively simple approach for the qualitative and semiquantitative structural characterization of M6P-containing oligosaccharides on therapeutic enzymes [197]. Regardless which method is chosen, for all enzyme replacement therapy products for lysosomal storage diseases, the total M6P content, relative bis-M6P content, and the K_D of the product to MPRs are expected to be included as release specifications as these attributes correlate directly with clinical efficacy and can vary by even subtle changes in manufacturing process and levels of product-related impurities.

Modifications on Amino Acid Residues

Certain amino acid residues in protein products are susceptible to undergoing some form of modification during manufacturing process or upon long-term storage. These include oxidation, deamidation, and atypical disulfide formation. Many factors, including pH, temperature, and ionic strength of buffers, can contribute to the formation of these modifications. In many instances, due to heterogeneity of biological products, it is hard to assess the impact of each individual modification on product safety and potency. It is therefore critical to characterize the degree of each modification early in product development and to optimize and control for condition to minimize change of certain modified forms over product development

and after licensure. Currently, some of these modifications are monitored at release and on stability by various chromatography methods. However, such modifications generally only slightly change the elution profile on chromatograms, severely impact the method's ability to resolute, and accurately quantify the impurity. Depending on how the area under the curve is calculated, these methods tend to over- or underestimate the actual content of the impurities. A more robust and accurate analysis of these impurities would be to focus on how to better resolve the impurity peaks away from the main product. It is expected that the manufacturers identify potential amino acid modifications that occur during manufacturing process and upon long-term storage conditions. For practical reasons, identification of amino acid residual modifications can be demonstrated by mass spectrometry analyses of digested peptides of products stored under relevant stressed or accelerated storage conditions. Peptides containing modifications observed in these analyses should be controlled for at release and on stability using a chromatography method or methods where the peptides are identified incontrovertibly and quantitated relative to their unmodified counterparts.

Dimers, Oligomers, and Aggregates

Due to their difference in size, the formation of dimers, oligomers, and aggregates can be easily identified by size-exclusion chromatography (SEC). It is important to stress that, due to heterogeneity of these species, especially for aggregates, an orthogonal method would be needed during product characterization to confirm that the SEC method is suitable to quantitate these impurities.

Some products on the market contain a protein stabilizer, such as HSA, in their final formulation. The presence of the stabilizing protein limits the use of SEC as a method to quantitate these impurities. In such cases, it is expected that such testing has been performed prior to formulation and a separate method to detect and quantitate product-specific oligomerization in product release and on stability. Due to the interference of the stabilizer, a semi-quantitation method, such as Western blot analyses both under reduced and non-reduced conditions, may be acceptable.

Truncated Forms

Endogenous proteases may cleave the amino or carboxyl terminus of a protein in a site-specific or random manner during cell culture process. Additionally, co-purifying proteases may further cleave protein products during manufacturing process when intermediates are held between unit operations. Whereas the integrity of the termini can be demonstrated by amino or carboxyl terminal sequencing, this method does not provide much quantitative information on the differentially processed forms. In cases where the differentially processed forms contribute to safety and potency, each form should be quantitated through a method that

indisputably resolves each form. Depending on the complexity of the product, the method can vary significantly. The manufacturer is expected to identify the impact of the variants on product quality and develop an appropriate product-specific method to quantify each form.

Charge Variant

The net charge of a protein is a fundamental physical property of any protein [201]. In addition to affecting intrinsic structure and thus solubility and stability of a protein, the charges also influence biological activity as well as their binding affinity to receptors or ligands. For therapeutic proteins, besides charges from the amino acid residues, charged glycans also account for a significant portion of the overall charge status of a protein. Because sialylated and phosphorylated glycans contribute significantly to the negative charges, a well-characterized charge profile not only ensures consistency of solubility and stability of a product but also provides controls for critical attributes affecting half-life and receptor binding. Generally, a quantitative method, such as capillary isoelectric focusing (cIEF) or ion-exchange HPLC should be implemented to provide quantitative limits for each peak or each group of peaks.

Non-monoclonal antibody protein therapeutics, as a product class, represent a very diversified group of proteins, each of which bears a unique series of critical attributes that collectively define a specific mechanism of action. The variety of expression hosts and combinations of chromatography steps also introduce a different profile of process-related, product-related impurities and product variants for each product. Consequently, besides several known attributes common to the majority protein products, characteristics to evaluate stability of each product can also vary. The characteristics discussed in this section summarize studies from a large list of currently approved non-monoclonal antibody protein products; they do not necessarily represent critical attributes defining the potency, purity, and safety profile of a particular product. A comprehensive characterization of a new product relies on the understanding of the product and accumulation of knowledge on how the manufacturing process impact the critical attributes throughout product development.

References

1. FDA (2006) Guidance for industry: Q9 quality risk management. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073511.pdf
2. FDA (2009) Guidance for industry: Q8(R2) pharmaceutical development. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073507.pdf
3. FDA (2009) Guidance for industry: Q10 pharmaceutical quality system. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073517.pdf

4. FDA (2012) Guidance for industry: Q11 development and manufacture of drug substances. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM261078.pdf
5. FDA (1999) Guidance for industry Q6B specifications: test procedures and acceptance criteria for biotechnological/biological products. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073488.pdf
6. FDA (2005) Guidance for industry: Q5E comparability of biotechnological/biological products subject to changes in their manufacturing process. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073476.pdf
7. FDA (1996) Guidance for industry: Q1B photostability testing of new drug substances and products. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073373.pdf
8. FDA (1997) Guidance for industry: Q1C stability testing for new dosage forms. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073374.pdf
9. FDA (2003) Guidance for industry: Q1A(R2) stability testing of new drug substances and products. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073369.pdf
10. FDA (2003) Guidance for industry: Q1D bracketing and matrixing designs for stability testing of new drug substances and products. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073379.pdf
11. FDA (2004) Guidance for industry: Q1E evaluation of stability data. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073380.pdf
12. FDA (1996) Guidance for industry: Q5C quality of biotechnological products: stability testing of biotechnological/biological products. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073466.pdf
13. FDA (2005) Guidance for industry: Q2(R1) validation of analytical procedures: text and methodology. http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm265700.htm
14. FDA (2014) Guidance for industry: immunogenicity assessment for therapeutic protein products. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM338856.pdf
15. FDA (2002) Guidance for industry: revised preventive measures to reduce the possible risk of transmission of Creutzfeldt-Jakob disease (CJD) and variant Creutzfeldt-Jakob disease (vCJD) by blood and blood products. http://www.fda.gov/downloads/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/Blood/ucm079711.pdf
16. Farshid M et al (2005) The clearance of viruses and transmissible spongiform encephalopathy agents from biologicals. Curr Opin Biotechnol 16(5):561–567
17. Sathish JG et al (2013) Challenges and approaches for the development of safer immunomodulatory biologics. Nat Rev Drug Discov 12(4):306–324
18. FDA (2009) Guidance for industry: assay development for immunogenicity testing of therapeutic proteins. Draft guidance. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM192750.pdf
19. FDA (2002) Science and the regulation of biological products: from a rich history to a challenging future. http://www.fda.gov/downloads/aboutfda/whatwedo/history/productregulation/100yearsofbiologicsregulation/ucm070313.pdf
20. FDA (1995) Guideline for quality assurance in blood establishments. http://www.fda.gov/downloads/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/Blood/UCM164981.pdf
21. FDA (2010) Guidance for industry: recommendations for blood establishments: training of back-up personnel, assessment of blood donor suitability and reporting certain changes to an approved application. http://www.fda.gov/downloads/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/Blood/UCM190373.pdf

22. WHO (2011) WHO Expert Committee on specifications for pharmaceutical preparations. World Health Organ Tech Rep Ser 961:148–214
23. FDA (2006) Guidance for industry: implementation of an acceptable full-length and abbreviated donor history questionnaires and accompanying materials for use in screening donors of source plasma. http://www.fda.gov/downloads/BiologicsBloodVaccines/BloodBloodProducts/ApprovedProducts/LicensedProductsBLAs/BloodDonorScreening/UCM341088.pdf
24. Orange JS et al (2006) Use of intravenous immunoglobulin in human disease: a review of evidence by members of the Primary Immunodeficiency Committee of the American Academy of Allergy, Asthma and Immunology. J Allergy Clin Immunol 117(4 Suppl):S525–S553
25. Lynch TJ et al (1996) Considerations of pool size in the manufacture of plasma derivatives. Transfusion 36(9):770–775
26. Zoon KC (1997) Safety implications of fractionated blood products, testimony before the subcommittee on human resources and intergovernmental affairs house committee on government reform and oversight
27. Burnouf T (2007) Modern plasma fractionation. Transfus Med Rev 21(2):101–117
28. Farrugia A, Quinti I (2014) Manufacture of immunoglobulin products for patients with primary antibody deficiencies – the effect of processing conditions on product safety and efficacy. Front Immunol 5:665
29. Hooper JA (2008) Intravenous immunoglobulins: evolution of commercial IVIG preparations. Immunol Allergy Clin North Am 28(4):765–778, viii
30. Radosevich M, Burnouf T (2010) Intravenous immunoglobulin G: trends in production methods, quality control and quality assurance. Vox Sang 98(1):12–28
31. Cohn EJ, Strong LE (1946) Preparation and properties of serum and plasma proteins; a system for the separation into fractions of the protein and lipoprotein components of biological tissues and fluids. J Am Chem Soc 68:459–475
32. Kistler P, Nitschmann H (1962) Large scale production of human plasma fractions. Eight years experience with the alcohol fractionation procedure of Nitschmann, Kistler and Lergier. Vox Sang 7:414–424
33. Oncley JL et al (1949) The separation of the antibodies, isoagglutinins, prothrombin, plasminogen and beta1-lipoprotein into subfractions of human plasma. J Am Chem Soc 71(2): 541–550
34. Etscheid M et al (2011) Identification of kallikrein and FXIa as impurities in therapeutic immunoglobulins: implications for the safety and control of intravenous blood products. Vox Sang 102:40–46
35. (2010) Octagam withdrawal, l, URGENT: voluntary market withdrawal, September 23, 2010 Octagam [immune globulin intravenous (Human)] 5% liquid preparation
36. Spath PJ et al (2015) On the dark side of therapies with immunoglobulin concentrates: the adverse events. Front Immunol 6:11
37. Mast EE et al (2006) A comprehensive immunization strategy to eliminate transmission of hepatitis B virus infection in the United States: recommendations of the Advisory Committee on Immunization Practices (ACIP) part II: immunization of adults. MMWR Recomm Rep 55((RR-16)):1–33
38. Finlayson JS et al (1980) Immunoglobulins: characteristics and uses of intravenous preparations. DHHS publication; no (FDA) 80–9005 [Bethesda, MD]: U.S. Department of Health and Human Services, Public Health Service, Food and Drug Administration; For sale by the Supt. of Docs., US GPO x, 245.
39. Bellac CL et al (2015) The role of isoagglutinins in intravenous immunoglobulin-related hemolysis. Transfusion 55(Suppl 2):S13–S22
40. Menis M et al (2013) Hyperimmune globulins and same-day thrombotic adverse events as recorded in a large healthcare database during 2008–2011. Am J Hematol 88(12):1035–1040
41. Hartung HP (2008) Advances in the understanding of the mechanism of action of IVIg. J Neurol 255(Suppl 3):3–6

42. Schwab I, Nimmerjahn F (2013) Intravenous immunoglobulin therapy: how does IgG modulate the immune system? Nat Rev Immunol 13(3):176–189
43. Marder VJ et al. (eds) (2013) Hemostasis and thrombosis: basic principles and clinical practice, 6th edn. Lippincott Williams & Wilkins, Philadelphia
44. Rosner F (1994) Hemophilia in classic rabbinic texts. J Hist Med Allied Sci 49(2):240–250
45. Lannoy N, Hermans C (2010) The 'royal disease' – haemophilia A or B? A haematological mystery is finally solved. Haemophilia 16(6):843–847
46. Chtourou S (2013) Production and clinical profile of human plasma coagulation factor VIII. In: Bertoni J, Goss N, Curling J (eds) Production of plasma proteins for therapeutic use. Wiley, Hoboken
47. Sandberg H et al (2001) Structural and functional characterization of B-domain deleted recombinant factor VIII. Semin Hematol 38(2 Suppl 4):4–12
48. Grancha S et al (2013) Factor IX. In Bertoni J, Goss N, Curling J (eds) Production of plasma proteins for therapeutic use. Wiley, Hoboken
49. Kumar SR (2015) Industrial production of clotting factors: challenges of expression, and choice of host cells. Biotechnol J 10(7):995–1004
50. Gray E et al (1995) Measurement of activated factor IX in factor IX concentrates: correlation with in vivo thrombogenicity. Thromb Haemost 73(4):675–679
51. Aronson DL, Finlayson JS (1980) Historical and future therapeutic plasma derivatives (epilogue). Semin Thromb Hemost 6(2):18
52. Raut S, Hubbard AR (2010) International reference standards in coagulation. Biologicals 38(4):423–429
53. Hubbard AR (2015) Potency labeling of novel factor VIII and factor IX concentrates: past experience and current strategy. Semin Thromb Hemost 41(8):849–854
54. Hubbard AR et al (2013) Recommendations on the potency labelling of factor VIII and factor IX concentrates. J Thromb Haemost 11(5):988–989
55. Wilmot HV, Hogwood J, Gray E (2014) Recombinant factor IX: discrepancies between one-stage clotting and chromogenic assays. Haemophilia 20(6):891–897
56. Panteleev MA et al (2004) Kinetics of factor X activation by the membrane-bound complex of factor IXa and factor VIIIa. Biochem J 381(Pt 3):779–794
57. Kaufman RJ (1992) Expression and structure-function properties of recombinant factor VIII. Transfus Med Rev 6(4):235–246
58. Brummel-Ziedins KE, Wolberg AS (2014) Global assays of hemostasis. Curr Opin Hematol 21(5):395–403
59. Osooli M, Berntorp E (2015) Inhibitors in haemophilia: what have we learned from registries? A systematic review. J Intern Med 277(1):1–15
60. Rosendaal FR et al (1993) A sudden increase in factor VIII inhibitor development in multitransfused hemophilia a patients in The Netherlands. Dutch hemophilia study group. Blood 81(8):2180–2186
61. Mannucci PM (2015) Half-life extension technologies for haemostatic agents. Thromb Haemost 113(1):165–176
62. Jaffers GJ et al (1986) Monoclonal antibody therapy. Anti-idiotypic and non-anti-idiotypic antibodies to OKT3 arising despite intense immunosuppression. Transplantation 41(5): 572–578
63. Glennie MJ, Johnson PW (2000) Clinical trials of antibody therapy. Immunol Today 21(8): 403–410
64. Beck A, Reichert JM (2012) Marketing approval of mogamulizumab: a triumph for glycoengineering. MAbs 4(4):419–425
65. Chu SY et al (2008) Inhibition of B cell receptor-mediated activation of primary human B cells by coengagement of CD19 and FcgammaRIIb with Fc-engineered antibodies. Mol Immunol 45(15):3926–3933
66. Desjarlais JR, Lazar GA (2011) Modulation of antibody effector function. Exp Cell Res 317(9):1278–1285

67. Lazar GA et al (2007) A molecular immunology approach to antibody humanization and functional optimization. Mol Immunol 44(8):1986–1998
68. Richards JO et al (2008) Optimization of antibody binding to FcgammaRIIa enhances macrophage phagocytosis of tumor cells. Mol Cancer Ther 7(8):2517–2527
69. Dall'Acqua WF et al (2006) Modulation of the effector functions of a human IgG1 through engineering of its hinge region. J Immunol 177(2):1129–1138
70. Mimoto F et al (2013) Novel asymmetrically engineered antibody Fc variant with superior FcgammaR binding affinity and specificity compared with afucosylated Fc variant. MAbs 5(2):229–236
71. Kubota T et al (2009) Engineered therapeutic antibodies with improved effector functions. Cancer Sci 100(9):1566–1572
72. Vafa O et al (2014) An engineered Fc variant of an IgG eliminates all immune effector functions via structural perturbations. Methods 65(1):114–126
73. Igawa T et al (2010) Antibody recycling by engineered pH-dependent antigen binding improves the duration of antigen neutralization. Nat Biotechnol 28(11):1203–1207
74. Li B et al (2014) Framework selection can influence pharmacokinetics of a humanized therapeutic antibody through differences in molecule charge. MAbs 6(5):1255–1264
75. Monnet C et al (2015) Selection of IgG variants with increased FcRn binding using random and directed mutagenesis: impact on effector functions. Front Immunol 6:39
76. Yeung YA et al (2009) Engineering human IgG1 affinity to human neonatal Fc receptor: impact of affinity improvement on pharmacokinetics in primates. J Immunol 182(12): 7663–7671
77. Natsume A et al (2008) Engineered antibodies of IgG1/IgG3 mixed isotype with enhanced cytotoxic activities. Cancer Res 68(10):3863–3872
78. Liu H et al (2014) In vitro and in vivo modifications of recombinant and human IgG antibodies. MAbs 6(5):1145–1154
79. Khawli LA et al (2010) Charge variants in IgG1: isolation, characterization, in vitro binding properties and pharmacokinetics in rats. MAbs 2(6):613–624
80. Vlasak J, Ionescu R (2008) Heterogeneity of monoclonal antibodies revealed by charge-sensitive methods. Curr Pharm Biotechnol 9(6):468–481
81. Dick LW Jr et al (2007) Determination of the origin of the N-terminal pyro-glutamate variation in monoclonal antibodies using model peptides. Biotechnol Bioeng 97(3):544–553
82. Yin S et al (2013) Characterization of therapeutic monoclonal antibodies reveals differences between in vitro and in vivo time-course studies. Pharm Res 30(1):167–178
83. Liu YD et al (2011) N-terminal glutamate to pyroglutamate conversion in vivo for human IgG2 antibodies. J Biol Chem 286(13):11211–11217
84. Harris RJ (2005) Heterogeneity of recombinant antibodies: linking structure to function. Dev Biol (Basel) 122:117–127
85. Cai B, Pan H, Flynn GC (2011) C-terminal lysine processing of human immunoglobulin G2 heavy chain in vivo. Biotechnol Bioeng 108(2):404–412
86. Johnson KA et al (2007) Cation exchange-HPLC and mass spectrometry reveal C-terminal amidation of an IgG1 heavy chain. Anal Biochem 360(1):75–83
87. Tsubaki M et al (2013) C-terminal modification of monoclonal antibody drugs: amidated species as a general product-related substance. Int J Biol Macromol 52:139–147
88. Kaschak T et al (2011) Characterization of the basic charge variants of a human IgG1: effect of copper concentration in cell culture media. MAbs 3(6):577–583
89. Harris RJ et al (2001) Identification of multiple sources of charge heterogeneity in a recombinant antibody. J Chromatogr B Biomed Sci Appl 752(2):233–245
90. Liu YD, van Enk JZ, Flynn GC (2009) Human antibody Fc deamidation in vivo. Biologicals 37(5):313–322
91. Kroon DJ, Baldwin-Ferro A, Lalan P (1992) Identification of sites of degradation in a therapeutic monoclonal antibody by peptide mapping. Pharm Res 9(11):1386–1393

92. Wakankar AA, Borchardt RT (2006) Formulation considerations for proteins susceptible to asparagine deamidation and aspartate isomerization. J Pharm Sci 95(11):2321–2336
93. Huang L et al (2005) In vivo deamidation characterization of monoclonal antibody by LC/MS/MS. Anal Chem 77(5):1432–1439
94. Vlasak J et al (2009) Identification and characterization of asparagine deamidation in the light chain CDR1 of a humanized IgG1 antibody. Anal Biochem 392(2):145–154
95. Haberger M et al (2014) Assessment of chemical modifications of sites in the CDRs of recombinant antibodies: susceptibility vs. functionality of critical quality attributes. MAbs 6(2):327–339
96. Gaza-Bulseco G et al (2008) Effect of methionine oxidation of a recombinant monoclonal antibody on the binding affinity to protein A and protein G. J Chromatogr B Analyt Technol Biomed Life Sci 870(1):55–62
97. Bertolotti-Ciarlet A et al (2009) Impact of methionine oxidation on the binding of human IgG1 to Fc Rn and Fc gamma receptors. Mol Immunol 46(8–9):1878–1882
98. Wang W et al (2011) Impact of methionine oxidation in human IgG1 Fc on serum half-life of monoclonal antibodies. Mol Immunol 48(6–7):860–866
99. Stracke J et al (2014) A novel approach to investigate the effect of methionine oxidation on pharmacokinetic properties of therapeutic antibodies. MAbs 6(5):1229–1242
100. Wei Z et al (2007) Identification of a single tryptophan residue as critical for binding activity in a humanized monoclonal antibody against respiratory syncytial virus. Anal Chem 79(7): 2797–2805
101. Goetze AM et al (2012) Rates and impact of human antibody glycation in vivo. Glycobiology 22(2):221–234
102. Quan C et al (2008) A study in glycation of a therapeutic recombinant humanized monoclonal antibody: where it is, how it got there, and how it affects charge-based behavior. Anal Biochem 373(2):179–191
103. Miller AK et al (2011) Characterization of site-specific glycation during process development of a human therapeutic monoclonal antibody. J Pharm Sci 100(7):2543–2550
104. Raju TS et al (2000) Species-specific variation in glycosylation of IgG: evidence for the species-specific sialylation and branch-specific galactosylation and importance for engineering recombinant glycoprotein therapeutics. Glycobiology 10(5):477–486
105. Gomord V et al (2010) Plant-specific glycosylation patterns in the context of therapeutic protein production. Plant Biotechnol J 8(5):564–587
106. Jefferis R (2012) Isotype and glycoform selection for antibody therapeutics. Arch Biochem Biophys 526(2):159–166
107. Boyd PN, Lines AC, Patel AK (1995) The effect of the removal of sialic acid, galactose and total carbohydrate on the functional activity of Campath-1H. Mol Immunol 32(17–18): 1311–1318
108. Hodoniczky J, Zheng YZ, James DC (2005) Control of recombinant monoclonal antibody effector functions by Fc N-glycan remodeling in vitro. Biotechnol Prog 21(6):1644–1652
109. Shields RL et al (2002) Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity. J Biol Chem 277(30):26733–26740
110. Shinkawa T et al (2003) The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. J Biol Chem 278(5):3466–3473
111. Ferrara C et al (2011) Unique carbohydrate-carbohydrate interactions are required for high affinity binding between FcgammaRIII and antibodies lacking core fucose. Proc Natl Acad Sci U S A 108(31):12669–12674
112. Shibata-Koyama M et al (2009) Nonfucosylated rituximab potentiates human neutrophil phagocytosis through its high binding for FcgammaRIIIb and MHC class II expression on the phagocytotic neutrophils. Exp Hematol 37(3):309–321

113. Shatz W et al (2013) Knobs-into-holes antibody production in mammalian cell lines reveals that asymmetric afucosylation is sufficient for full antibody-dependent cellular cytotoxicity. MAbs 5(6):872–881
114. Mori K et al (2004) Engineering Chinese hamster ovary cells to maximize effector function of produced antibodies using FUT8 siRNA. Biotechnol Bioeng 88(7):901–908
115. Yamane-Ohnuki N et al (2004) Establishment of FUT8 knockout Chinese hamster ovary cells: an ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity. Biotechnol Bioeng 87(5):614–622
116. Malphettes L et al (2010) Highly efficient deletion of FUT8 in CHO cell lines using zinc-finger nucleases yields cells that produce completely nonfucosylated antibodies. Biotechnol Bioeng 106(5):774–783
117. Ferrara C et al (2006) Modulation of therapeutic antibody effector functions by glycosylation engineering: influence of Golgi enzyme localization domain and co-expression of heterologous beta1, 4-N-acetylglucosaminyltransferase III and Golgi alpha-mannosidase II. Biotechnol Bioeng 93(5):851–861
118. Wright A, Morrison SL (1997) Effect of glycosylation on antibody function: implications for genetic engineering. Trends Biotechnol 15(1):26–32
119. Gala FA, Morrison SL (2004) V region carbohydrate and antibody expression. J Immunol 172(9):5489–5494
120. Dunn-Walters D, Boursier L, Spencer J (2000) Effect of somatic hypermutation on potential N-glycosylation sites in human immunoglobulin heavy chain variable regions. Mol Immunol 37(3–4):107–113
121. Zhu D et al (2002) Acquisition of potential N-glycosylation sites in the immunoglobulin variable region by somatic mutation is a distinctive feature of follicular lymphoma. Blood 99 (7):2562–2568
122. Xu PC et al (2012) Influence of variable domain glycosylation on anti-neutrophil cytoplasmic autoantibodies and anti-glomerular basement membrane autoantibodies. BMC Immunol 13:10
123. Chung CH et al (2008) Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose. N Engl J Med 358(11):1109–1117
124. Brezski RJ, Jordan RE (2010) Cleavage of IgGs by proteases associated with invasive diseases: an evasion tactic against host immunity? MAbs 2(3):212–220
125. Gao SX et al (2011) Fragmentation of a highly purified monoclonal antibody attributed to residual CHO cell protease activity. Biotechnol Bioeng 108(4):977–982
126. Cordoba AJ et al (2005) Non-enzymatic hinge region fragmentation of antibodies in solution. J Chromatogr B Analyt Technol Biomed Life Sci 818(2):115–121
127. Vlasak J, Ionescu R (2011) Fragmentation of monoclonal antibodies. MAbs 3(3):253–263
128. Gaza-Bulseco G, Liu H (2008) Fragmentation of a recombinant monoclonal antibody at various pH. Pharm Res 25(8):1881–1890
129. Kamerzell TJ et al (2011) The relative rate of immunoglobulin gamma 1 fragmentation. J Pharm Sci 100(4):1341–1349
130. Smith MA et al (1996) Specific cleavage of immunoglobulin G by copper ions. Int J Pept Protein Res 48(1):48–55
131. Ouellette D et al (2009) Elevated cleavage of human immunoglobulin gamma molecules containing a lambda light chain mediated by iron and histidine. Anal Biochem 389(2): 107–117
132. Vazquez-Rey M, Lang DA (2011) Aggregates in monoclonal antibody manufacturing processes. Biotechnol Bioeng 108(7):1494–1508
133. Arvinte T et al (2013) Aggregation of biopharmaceuticals in human plasma and human serum: implications for drug research and development. MAbs 5(3):491–500
134. Chennamsetty N et al (2009) Design of therapeutic proteins with enhanced stability. Proc Natl Acad Sci U S A 106(29):11937–11942

135. Clark RH et al (2014) Remediating agitation-induced antibody aggregation by eradicating exposed hydrophobic motifs. MAbs 6(6):1540–1550
136. Joubert MK et al (2011) Classification and characterization of therapeutic antibody aggregates. J Biol Chem 286(28):25118–25133
137. Vermeer AW, Norde W (2000) The thermal stability of immunoglobulin: unfolding and aggregation of a multi-domain protein. Biophys J 78(1):394–404
138. Nicoud L et al (2014) Kinetic analysis of the multistep aggregation mechanism of monoclonal antibodies. J Phys Chem B 118(36):10595–10606
139. Remmele RL Jr et al (2006) Active dimer of epratuzumab provides insight into the complex nature of an antibody aggregate. J Pharm Sci 95(1):126–145
140. Paul R et al (2012) Structure and function of purified monoclonal antibody dimers induced by different stress conditions. Pharm Res 29(8):2047–2059
141. Luo Y et al (2009) Dimers and multimers of monoclonal IgG1 exhibit higher in vitro binding affinities to Fcgamma receptors. MAbs 1(5):491–504
142. da Silva AJ et al (2002) Alefacept, an immunomodulatory recombinant LFA-3/IgG1 fusion protein, induces CD16 signaling and CD2/CD16-dependent apoptosis of CD2(+) cells. J Immunol 168(9):4462–4471
143. Bielekova B (2013) Daclizumab therapy for multiple sclerosis. Neurotherapeutics 10(1): 55–67
144. Tracey D et al (2008) Tumor necrosis factor antagonist mechanisms of action: a comprehensive review. Pharmacol Ther 117(2):244–279
145. Jiang XR et al (2011) Advances in the assessment and control of the effector functions of therapeutic antibodies. Nat Rev Drug Discov 10(2):101–111
146. Weiskopf K, Weissman IL (2015) Macrophages are critical effectors of antibody therapies for cancer. MAbs 7(2):303–310
147. Braster R, O'Toole T, van Egmond M (2014) Myeloid cells as effector cells for monoclonal antibody therapy of cancer. Methods 65(1):28–37
148. Overdijk MB et al (2015) Antibody-mediated phagocytosis contributes to the anti-tumor activity of the therapeutic antibody daratumumab in lymphoma and multiple myeloma. MAbs 7(2):311–321
149. Schnueriger A et al (2011) Development of a quantitative, cell-line based assay to measure ADCC activity mediated by therapeutic antibodies. Mol Immunol 48(12–13):1512–1517
150. Parekh BS et al (2012) Development and validation of an antibody-dependent cell-mediated cytotoxicity-reporter gene assay. MAbs 4(3):310–318
151. Chung S et al (2014) Characterization of in vitro antibody-dependent cell-mediated cytotoxicity activity of therapeutic antibodies – impact of effector cells. J Immunol Methods 407: 63–75
152. Lim SH et al (2011) Fc gamma receptor IIb on target B cells promotes rituximab internalization and reduces clinical efficacy. Blood 118(9):2530–2540
153. Beers SA et al (2008) Type II (tositumomab) anti-CD20 monoclonal antibody out performs type I (rituximab-like) reagents in B-cell depletion regardless of complement activation. Blood 112(10):4170–4177
154. Beers SA et al (2010) Antigenic modulation limits the efficacy of anti-CD20 antibodies: implications for antibody selection. Blood 115(25):5191–5201
155. Griggs J, Zinkewich-Peotti K (2009) The state of the art: immune-mediated mechanisms of monoclonal antibodies in cancer therapy. Br J Cancer 101(11):1807–1812
156. Smith KG, Clatworthy MR (2010) FcgammaRIIB in autoimmunity and infection: evolutionary and therapeutic implications. Nat Rev Immunol 10(5):328–343
157. Derer S et al (2012) Impact of epidermal growth factor receptor (EGFR) cell surface expression levels on effector mechanisms of EGFR antibodies. J Immunol 189(11): 5230–5239

158. Taylor RJ et al (2009) FcgammaRIIIa polymorphisms and cetuximab induced cytotoxicity in squamous cell carcinoma of the head and neck. Cancer Immunol Immunother 58(7): 997–1006
159. Taylor RJ et al (2015) Ex vivo antibody-dependent cellular cytotoxicity inducibility predicts efficacy of cetuximab. Cancer Immunol Res 3(5):567–574
160. Tiroch K et al (2002) Intracellular domains of target antigens influence their capacity to trigger antibody-dependent cell-mediated cytotoxicity. J Immunol 168(7):3275–3282
161. Wypych J et al (2008) Human IgG2 antibodies display disulfide-mediated structural isoforms. J Biol Chem 283(23):16194–16205
162. Martinez T et al (2008) Disulfide connectivity of human immunoglobulin G2 structural isoforms. Biochemistry 47(28):7496–7508
163. Dillon TM et al (2008) Structural and functional characterization of disulfide isoforms of the human IgG2 subclass. J Biol Chem 283(23):16206–16215
164. Allen MJ et al (2009) Interchain disulfide bonding in human IgG2 antibodies probed by site-directed mutagenesis. Biochemistry 48(17):3755–3766
165. White AL et al (2015) Conformation of the human immunoglobulin G2 hinge imparts superagonistic properties to immunostimulatory anticancer antibodies. Cancer Cell 27(1): 138–148
166. Yoo EM et al (2003) Human IgG2 can form covalent dimers. J Immunol 170(6):3134–3138
167. Lacher NA et al (2010) Development of a capillary gel electrophoresis method for monitoring disulfide isomer heterogeneity in IgG2 antibodies. Electrophoresis 31(3):448–458
168. He Y et al (2010) Analysis of identity, charge variants, and disulfide isomers of monoclonal antibodies with capillary zone electrophoresis in an uncoated capillary column. Anal Chem 82(8):3222–3230
169. Correia IR (2010) Stability of IgG isotypes in serum. MAbs 2(3):221–232
170. van der Neut Kolfschoten M et al (2007) Anti-inflammatory activity of human IgG4 antibodies by dynamic Fab arm exchange. Science 317(5844):1554–1557
171. Davies AM et al (2013) Crystal structure of the human IgG4 C(H)3 dimer reveals the role of Arg409 in the mechanism of Fab-arm exchange. Mol Immunol 54(1):1–7
172. Lewis KB et al (2009) Comparison of the ability of wild type and stabilized human IgG(4) to undergo Fab arm exchange with endogenous IgG(4)in vitro and in vivo. Mol Immunol 46(16):3488–3494
173. Stubenrauch K et al (2010) Impact of molecular processing in the hinge region of therapeutic IgG4 antibodies on disposition profiles in cynomolgus monkeys. Drug Metab Dispos 38(1): 84–91
174. Wilkinson IC et al (2013) Monovalent IgG4 molecules: immunoglobulin Fc mutations that result in a monomeric structure. MAbs 5(3):406–417
175. Deng L et al (2004) Detection and quantification of the human IgG4 half-molecule, HL, from unpurified cell-culture supernatants. Biotechnol Appl Biochem 40(Pt 3):261–269
176. Forrer K, Hammer S, Helk B (2004) Chip-based gel electrophoresis method for the quantification of half-antibody species in IgG4 and their by- and degradation products. Anal Biochem 334(1):81–88
177. Zhu ZC et al (2013) Investigation of monoclonal antibody fragmentation artifacts in non-reducing SDS-PAGE. J Pharm Biomed Anal 83:89–95
178. Kuenzel EA, Krebs EG (1985) A synthetic peptide substrate specific for casein kinase II. Proc Natl Acad Sci U S A 82(3):737–741
179. Friedhoff P et al (1996) Kinetic analysis of the cleavage of natural and synthetic substrates by the Serratia nuclease. Eur J Biochem 241(2):572–580
180. Takeuchi M et al (1989) Relationship between sugar chain structure and biological activity of recombinant human erythropoietin produced in Chinese hamster ovary cells. Proc Natl Acad Sci U S A 86(20):7819–7822
181. Egrie JC et al (2003) Darbepoetin alfa has a longer circulating half-life and greater in vivo potency than recombinant human erythropoietin. Exp Hematol 31(4):290–299

182. Caucheteur C, Guo T, Albert J (2015) Review of plasmonic fiber optic biochemical sensors: improving the limit of detection. Anal Bioanal Chem 407(14):3883–3897
183. Olaru A et al (2015) Surface plasmon resonance (SPR) biosensors in pharmaceutical analysis. Crit Rev Anal Chem 45(2):97–105
184. Bork K, Horstkorte R, Weidemann W (2009) Increasing the sialylation of therapeutic glycoproteins: the potential of the sialic acid biosynthetic pathway. J Pharm Sci 98(10): 3499–3508
185. Elliott S et al (2004) Control of rHuEPO biological activity: the role of carbohydrate. Exp Hematol 32(12):1146–1155
186. Jongen SP et al (2007) N-glycans of recombinant human acid alpha-glucosidase expressed in the milk of transgenic rabbits. Glycobiology 17(6):600–619
187. Hossler P, Khattak SF, Li ZJ (2009) Optimal and consistent protein glycosylation in mammalian cell culture. Glycobiology 19(9):936–949
188. Fishburn CS (2008) The pharmacology of PEGylation: balancing PD with PK to generate novel therapeutics. J Pharm Sci 97(10):4167–4183
189. Veronese FM, Mero A (2008) The impact of PEGylation on biological therapies. BioDrugs 22(5):315–329
190. Fares F et al (2010) Designing a long-acting human growth hormone (hGH) by fusing the carboxyl-terminal peptide of human chorionic gonadotropin beta-subunit to the coding sequence of hGH. Endocrinology 151(9):4410–4417
191. Kenanova VE et al (2010) Tuning the serum persistence of human serum albumin domain III: diabody fusion proteins. Protein Eng Des Sel 23(10):789–798
192. Doknic M, Stojanovic M, Popovic V (2014) Novel long-acting GH preparations. Pediatr Endocrinol Rev 12(2):206–212
193. Venetz D et al (2015) Glycosylation profiles determine extravasation and disease-targeting properties of armed antibodies. Proc Natl Acad Sci U S A 112(7):2000–2005
194. West MB et al (2010) Analysis of site-specific glycosylation of renal and hepatic gamma-glutamyl transpeptidase from normal human tissue. J Biol Chem 285(38):29511–29524
195. Xu Y, Bailey UM, Schulz BL (2015) Automated measurement of site-specific N-glycosylation occupancy with SWATH-MS. Proteomics 15(13):2177–2186
196. Brady RO (2006) Enzyme replacement for lysosomal diseases. Annu Rev Med 57:283–296
197. Bones J et al (2011) Identification of N-glycans displaying mannose-6-phosphate and their site of attachment on therapeutic enzymes for lysosomal storage disorder treatment. Anal Chem 83(13):5344–5352
198. Ghosh P, Dahms NM, Kornfeld S (2003) Mannose 6-phosphate receptors: new twists in the tale. Nat Rev Mol Cell Biol 4(3):202–212
199. Parenti G (2009) Treating lysosomal storage diseases with pharmacological chaperones: from concept to clinics. EMBO Mol Med 1(5):268–279
200. Zhu Y et al (2009) Glycoengineered acid alpha-glucosidase with improved efficacy at correcting the metabolic aberrations and motor function deficits in a mouse model of Pompe disease. Mol Ther 17(6):954–963
201. Winzor DJ (2004) Determination of the net charge (valence) of a protein: a fundamental but elusive parameter. Anal Biochem 325(1):1–20

Top Med Chem (2017) 21: 123–160
DOI: 10.1007/7355_2016_1

Published online: 23 April 2017

Immunogenicity Lessons Learned from the Clinical Development of Vatreptacog Alfa, A Recombinant Activated Factor VII Analog, in Hemophilia with Inhibitors

Kasper Lamberth, Karin Nana Weldingh, Silke Ehrenforth, Mette Ribel Chéhadé, and Henrik Østergaard

Abstract Hemophilia A is a rare bleeding disorder characterized by defective blood clotting due to diminished levels or absence of coagulation Factor VIII (FVIII). The preferred treatment option is FVIII replacement therapy. However, in 20–30% of the patients neutralizing (inhibitory) anti-FVIII antibodies develop rendering patients dependent on other treatment modalities such as the bypassing agent recombinant factor VIIa (rFVIIa). rFVIIa has a 20-year safety track record with no reports of immunogenicity in congenital hemophilia patients with inhibitors. To improve treatment efficacy of rFVIIa, the recombinant analog vatreptacog alpha was developed by Novo Nordisk A/S and taken into clinical development in 2006. Despite differing from rFVIIa by only three amino acid substitutions, results from the phase III trial demonstrated that some patients developed anti-drug antibodies. In this chapter, we give an introduction to hemophilia with focus on rFVIIa and the development of vatreptacog alfa. In addition, we summarize the findings from the clinical trials and characterization of the identified anti-drug antibodies. Finally, we show how various immunogenicity prediction tools have been used to investigate the immunogenicity risk of vatreptacog alfa leading to the identification of a potential new T-cell epitope that could contribute to the observed immunogenicity of the compound in humans.

Keywords Antidrug antibodies, FVIIa, Hemophilia, Immunogenicity, Inhibitors, Prediction, T-cell epitope, Vatreptacog alfa

K. Lamberth (✉), K.N. Weldingh, M.R. Chéhadé, and H. Østergaard
Novo Nordisk, Copenhagen, Denmark
e-mail: kplm@novonordisk.com

S. Ehrenforth
Novo Nordisk, Zurich, Switzerland

Contents

1 Introduction

Vatreptacog alfa (Novo Nordisk, Bagsværd, Denmark) and BAY 86-6150 (Bayer, Leverkusen, Germany) were both recently in clinical development as bioengineered coagulation proteins designed to address the need for new treatment options in hemophilia with inhibitors [1–4]. These novel variants of activated coagulation factor VII (FVIIa) represent the vanguard of over a dozen bioengineered therapeutic coagulation proteins currently (or recently) in development [5–8]. Both vatreptacog alfa and BAY 86-6150 were shown to be safe and well tolerated in early clinical studies [2, 4]. However, development of both agents was discontinued when antidrug antibodies (ADAs) were identified in some treated patients during clinical trials [1, 9].

The vatreptacog alfa and BAY 86-6150 stories offer valuable lessons about the potential immunogenicity risks associated with bioengineered, sequence-modified proteins. These bioengineered analogs both triggered ADA development despite very modest changes in the primary sequence: vatreptacog alfa had three amino acid substitutions [1, 3], while BAY 86-6150 carried six substitutions [4, 7]. This

chapter presents a detailed discussion of the vatreptacog alfa case, which also offers a unique and valuable insight into the tools and methods used to predict the immunogenicity of protein therapeutics.

2 What Is Hemophilia?

Hemophilia is a congenital, X-linked bleeding disorder caused by deficiency or absence of specific coagulation factors [10–13]. The two forms of the disease are hemophilia A (deficiency or lack of coagulation factor VIII [FVIII]) and hemophilia B (deficiency or lack of coagulation factor IX [FIX]) [11, 12, 14]. The clotting factor deficiencies in hemophilia A and B result from a variety of mutations in the respective clotting factor genes [10, 13].

Hemophilia occurs in approximately 1 in 5,000 (hemophilia A) and 1 in 30,000–50,000 (hemophilia B) live male births [10, 12, 14]. Both hemophilia A and B are characterized by a lifelong bleeding tendency [13, 14]; the severity of this bleeding phenotype generally corresponds to residual clotting factor levels, such that severe disease is defined by FVIII or FIX levels of $\leq$1% [12–14]. Patients with severe hemophilia A or B suffer spontaneous bleeds into joints and muscles without any identifiable trauma or injury [13, 14]. If these bleeds are inadequately treated, over time they produce chronic pain, joint destruction, and disability [11, 14].

Standard treatment for hemophilia A and B involves infusion of the missing or deficient clotting factor [5, 13]. FVIII or FIX concentrates used in replacement therapy can be derived from fractionated human plasma or developed through recombinant DNA technology and can be administered on demand when a bleed occurs or as prophylaxis to prevent bleeding [5, 11]. Prophylaxis is the preferred clinical management strategy in developed countries.

3 Immune Responses in Hemophilia

Autoimmune responses to endogenous coagulation proteins are very rare, occurring in approximately one per million individuals per year [15, 16], and result in acquired hemophilia. However, as FVIII and FIX are biotherapeutic proteins, they trigger an immune response characterized by antidrug antibodies (ADAs) much more frequently when used therapeutically in hemophilia patients who lack the endogenous counterparts [17, 18]. Development of ADAs is a T cell-dependent process [17–21] resulting from a lack of central T-cell tolerance. Tolerance to self-proteins is a vital part of immune system development [18, 20, 22]; proteins (antigens) encountered later in life are usually recognized as foreign and may elicit an immune response [18] culminating in antibody production. In hemophilia A and B, mutations in the underlying FVIII and FIX genes, respectively, mean that the resulting factor protein is either nonfunctional, incomplete, or missing altogether.

As a result, tolerance to the complete, functional factor protein cannot be established during immunological development. When factor proteins are administered therapeutically, part of the protein molecule may therefore be recognized as foreign and may trigger an immune response [17, 18, 22]. Certain types of mutation in the FVIII and FIX genes are associated with a greater risk of inhibitor development [23, 24]. Even with small mutations in the FVIII or FIX gene, infused factor protein may trigger antibody development as one or more portions of the protein sequence will be foreign to the immune system [18].

ADAs do not always have a clinical manifestation, and such binding antibodies toward FVIII have been reported in healthy subjects as well as hemophilia patients who are well treated [25–27]. In some cases, however, they neutralize the target binding and thus the clinical efficacy of the infused protein [25, 26]. Clinically relevant neutralizing ADAs are known as inhibitors.

3.1 Inhibitor Development

The development of inhibitors against FVIII or FIX is the most serious and challenging treatment-related complication of hemophilia [2, 12, 13, 27–30]. Inhibitors are immunoglobulin G (IgG) antibodies that neutralize the infused clotting factor, rendering standard replacement therapy with FVIII or FIX ineffective [13, 29, 31].

The development of inhibitors in hemophilia depends on a complex interaction of numerous patient-related, clinical, and product-related factors [12, 28, 32]. In addition to the type of causative FVIII/FIX mutation, patient-related factors include polymorphisms of immune response genes, such as interleukin (IL)-10, tumor necrosis factor-alfa (TNFα), cytotoxic T-lymphocyte-associated protein 4 (CTLA-4), and major histocompatibility complex class II (MHC-II) – which is referred to as human leukocyte antigen (HLA) class II in humans [12]. The binding of peptides to MHC-II molecules on the surface of antigen-presenting cells in the immune system is a key factor in generating an immune response [20, 33]. The patient's HLA class II genotype governs which peptides can bind to HLA class II molecules and therefore influences the development of ADAs (including inhibitors) [34]. Clinical factors that may affect inhibitor development include, but are not limited to, the mode and route of administration, the length and type of exposure to product, and the indication for which the product is administered (e.g., surgery, severe bleeds) [12]. Potential product-related impact factors include product formulation, the presence and type of posttranslational modifications, level of similarity to endogenous protein, and the presence of T-cell epitopes.

4 Recombinant Activated Factor VII

FVIII- and FIX-bypassing agents are the only treatment options for hemophilia with high-titer inhibitors [35]. Recombinant FVIIa (rFVIIa; NovoSeven®; Novo Nordisk, Bagsværd, Denmark) is one of only two bypassing agents currently available and is the only recombinant one [35].

4.1 Rationale Behind rFVIIa Development

Hemostasis is the process of blood clot formation at the site of vessel injury. According to the cell-based model of hemostasis, three integrated phases are required to generate sufficient thrombin to form a fibrin clot [31, 36]. The initiation phase begins when endothelial damage allows tissue factor (TF) on the surface of TF-bearing cells to be exposed to the bloodstream. TF then binds and activates FVII, which facilitates the generation of a small amount of thrombin. In the amplification phase, thrombin activates platelets, leading to subsequent activation of FV, FVIII, and FXI on platelet surfaces. During the propagation phase, the formation of FVIIIa–FIXa (tenase) complexes activates FX and drives a "burst" of thrombin generation on activated platelets [36]. This leads to cleavage of fibrinogen to fibrin, yielding a stable, cross-linked fibrin clot.

The absence of either FVIII or FIX prevents assembly of FVIIIa–FIXa complexes on activated platelets, which prevents both FX activation and a full thrombin burst [31, 36]. The end result is a fibrin clot that is unstable and easily dissolved. Although most hemophilia patients can be treated with replacement FVIII or FIX products, patients with inhibitors require therapies that bypass the need for these factors. High-dose rFVIIa accomplishes this via TF-independent activation of FX on activated platelets in the absence of FVIIIa–FIXa complexes. This restores the thrombin burst and allows the formation of a stable clot [36].

4.2 rFVIIa in Hemophilia

rFVIIa has been available since 1996 in Europe and since 1999 in the USA and Canada. It has an amino acid sequence identical to that of endogenous FVIIa and has provided high levels of hemostatic efficacy and safety during clinical use in hemophilia patients with inhibitors for more than 20 years [37]. Hemostasis following treatment with rFVIIa 90 μg/kg has been rated effective after a mean of 2.2 injections in 92% (566/614) of bleeds [38], and successful outcomes were reported in patients with joint bleeds who injected single 270 μg/kg doses [39–41]. Data from hemophilia registries report that rFVIIa can be safely used at home and support the efficacy and safety of rFVIIa doses >200 μg/kg [42]. To date, there

have been no confirmed cases of neutralizing antibodies (inhibitors) against rFVIIa in congenital hemophilia patients [43].

Both currently available bypassing agents used to treat inhibitor patients are perceived to be less effective than regular factor replacement therapy in non-inhibitor patients [35]. As a result attempts have been made to develop improved rFVIIa variants with enhanced intrinsic procoagulant activity; one of these variants is vatreptacog alfa [3, 44].

5 rFVIIa Analog Vatreptacog Alfa

Novo Nordisk developed vatreptacog alfa to provide a bypassing agent that offered more reliable and sustained resolution of acute bleeds in hemophilia patients with inhibitors [1, 2]. It is a genetically engineered variant of rFVIIa produced in a Chinese hamster ovary cell line cultured in a serum-free medium [1–3]. No raw materials or excipients of human or animal origin were used in its production.

5.1 Structure of Vatreptacog Alfa

Vatreptacog alfa has 99% amino acid identity to native FVIIa [1]. It is composed of a light and heavy chain connected by a disulfide bond (Fig. 1). The light chain includes one membrane-interactive domain and two epidermal growth factor-like domains; these domains are the same as in rFVIIa in order to maintain the clinically beneficial properties of rFVIIa [3]. The heavy chain comprises a single serine protease domain that is structurally similar to rFVIIa with the exception of three amino acid substitutions (V158D, E296V, and M298Q) [1–3]. These amino acid substitutions allow the N-terminal to dock efficiently and stabilize FVIIa in its active conformation without the need for TF [3, 45] and increase the molecule's TF-independent activity.

5.2 Functional Characterization of Vatreptacog Alfa

The procoagulant activity of vatreptacog alfa has been verified in various in vitro [46–48], ex vivo [49–52], and in vivo [53, 54] mice models of hemophilia. Together, these studies demonstrated that vatreptacog alfa provides hemostasis at lower concentrations than rFVIIa and that clot formation is faster and stronger, with increased stability against fibrinolytic degradation.

In a cell-based model of hemophilia, vatreptacog alfa generated 30-fold higher FXa levels than similar concentrations of rFVIIa [46]. It also resulted in four- to tenfold higher maximal thrombin generation rates than equal rFVIIa concentrations. Another cell-based model found that normalization of maximum thrombin

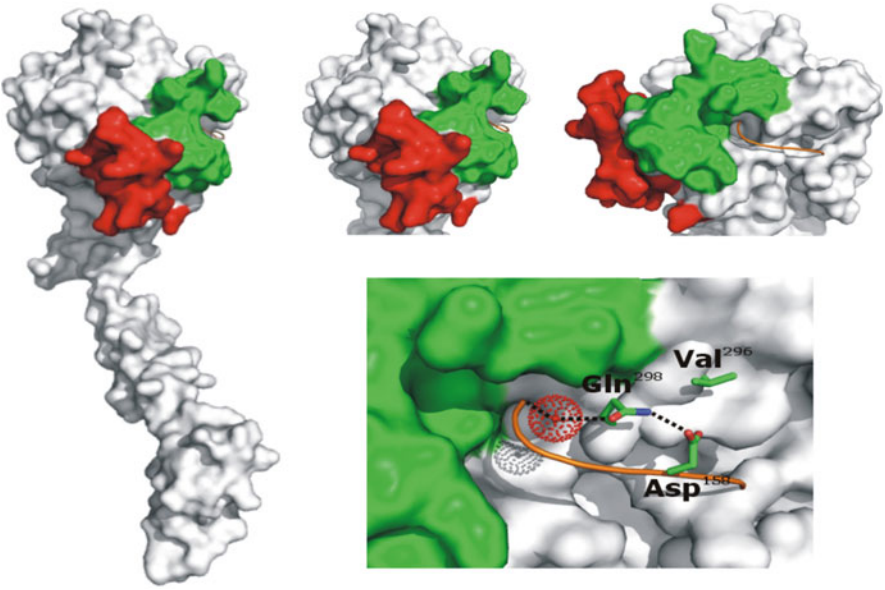


Fig. 1 The structure of vatreptacog alfa, showing the sites of the three amino acid substitutions *Left*: the entire molecule, with the protease domain at the *top*. The green area represents the activation domain, and the red area represents the tissue factor (TF)-interactive surface. *Top middle*: The protease domain in the same orientation as in the main figure to the *left*. *Top right*: The protease domain has been rotated 90° relative to the middle structure to clearly show the N-terminal tail (*orange*). *Bottom right*: Close-up of the region of the three amino acid substitutions. V158D is located in the tail, while E296V and M298Q are on the same β strand close to the activation domain. The *dotted lines* indicate hydrogen bonds that connect the N-terminus, a water molecule (*dotted red sphere*), M298Q and V158D. This hydrogen bond network does not exist in FVII. Reproduced from Persson et al. [3]

generation in severe hemophilia required 500 nM rFVIIa but only 25 nM vatreptacog alfa, while moderate/mild hemophilia required 25–100 nM rFVIIa but only 5 nM vatreptacog alfa [47]. Importantly, the generation of excess thrombin was confined to areas of damaged epithelium where hemostasis was required [48].

In assays using plasma from hemophilia patients, vatreptacog alfa shortened the clotting onset and increased the maximum rate of fibrin formation and fibrin network density in plasma clots at a concentration significantly lower than rFVIIa [52]. Vatreptacog alfa also resulted in a more pronounced, less variable hemostatic effect in blood samples from patients with varying degrees of hemophilia, producing a normalization of clotting parameters equivalent to values obtained in subjects with normal coagulation [50, 51]. In addition, increasing concentrations of vatreptacog alfa normalized platelet function, clot structure, and thrombin generation consistently in blood obtained from hemophilia patients with or without inhibitors [49].

The greater and faster thrombin burst for vatreptacog alfa relative to rFVIIa reported in the in vitro and ex vivo studies described above translated into greater efficacy and faster hemostatic effect in vivo. In murine models of hemophilia A,

vatreptacog alfa significantly shortened bleeding times and decreased blood loss in tail bleeding experiments, with a potency three to four times that of rFVIIa [53]. Vatreptacog alfa also showed greater efficacy and faster bleeding resolution than FVIII, rFVIIa, and pd-aPCC in a severe tail bleeding model [54]. No evidence of adverse events with vatreptacog alfa was found in these studies [53, 54].

5.3 *Mechanism of Action of Vatreptacog Alfa*

Since vatreptacog alfa and rFVIIa have the same membrane-binding domain and TF-interacting regions [3], the mechanism they employ to activate FX is the same, i.e., a TF-independent reaction localized to the platelet surface that bypasses the need for the FVIIIa–FIXa complex. The increased potency of vatreptacog alfa relative to rFVIIa can be explained by differing equilibria between the inactive and active conformations of each of the molecules: vatreptacog alfa exists predominantly in the active form, whereas rFVIIa exists predominantly in the inactive form [3]. As a result, a larger proportion of vatreptacog alfa molecules than rFVIIa molecules are active at any given time, resulting in an increased rate of vatreptacog alfa-mediated FX activation [55]. In addition, vatreptacog alfa binds to more sites on the surface of activated platelets than FVIIa, as assessed by flow cytometry (Fig. 2) [3, 36]. Together, these findings explain the larger and more rapid thrombin burst seen with vatreptacog alfa compared with rFVIIa (Fig. 3) [55].

Vatreptacog alfa, like rFVIIa, cannot directly activate platelets [3]. Its thrombin-generating effects require the presence of platelets that have already been activated and are therefore restricted to the injury site. As there is no reason to suspect that systemic thrombin generation occurs, there is consequently no reason to suspect an increased risk of thromboembolic events.

5.4 *Early Immunogenicity Risk Assessment*

Early assessment of the potential immunogenicity risk of vatreptacog alfa took place in the preclinical research phase using early in silico tools and in vivo animal models. At that time, in silico immunogenicity prediction tools that analyze peptide binding to MHC-II indicated a risk for generating a new potential T-cell epitope. However, these tools were still in their infancy and were not considered reliable. In rats neonatally tolerized to FVII, no difference in tolerance breakage was observed after administration of vatreptacog alfa compared with rFVIIa [56]. Furthermore, in immunization studies with rFVIIa and vatreptacog alfa in transgenic mice engineered to express human FVII, no significant difference in breaking of tolerance was found between the two molecules (unpublished data). Consequently, no difference in immunogenicity risk was found between wild-type FVIIa and vatreptacog alfa in these early animal studies.

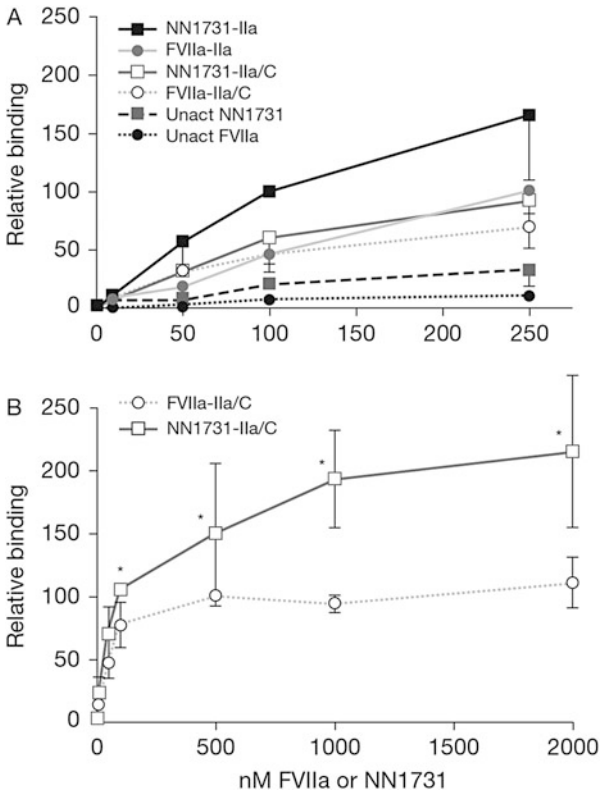


Fig. 2 Platelet binding of vatreptacog alfa (NN1731) and FVIIa. Unactivated (unact) platelets or platelets activated with thrombin (IIa) or thrombin + convulxin (IIa/C) were incubated with the indicated concentration of FVIIa or vatreptacog alfa and analyzed by flow cytometry. Binding was normalized to the binding of vatreptacog alfa at 100 nM in each experiment (=100%). *Upper panel*: Data shown are means and SD of data from four separate experiments on platelets from different donors. Binding of vatreptacog alfa to thrombin-activated platelets was slightly higher than for FVIIa ($p < 0.05$ by paired *t*-test at 250 nM), but binding of vatreptacog alfa to platelets activated by thrombin + convulxin was markedly higher than for FVIIa ($p < 0.01$ by paired *t*-test at 100 and 250 nM). *Lower panel*: Binding was characterized up to a concentration of 2 μM FVIIa or vatreptacog alfa on platelets activated by thrombin + convulxin. Data are from four experiments and were different from the experiments shown in the upper panel. Maximal binding of vatreptacog alfa to thrombin + convulxin-activated platelets was significantly greater than the maximal binding of FVIIa; therefore, vatreptacog alfa appears to bind to a greater number of sites per platelet. Binding of both molecules exhibited an apparent Kd of ~100 nM. *$p < 0.05$ by paired *t*-test. Reproduced from Hoffman et al. [97]

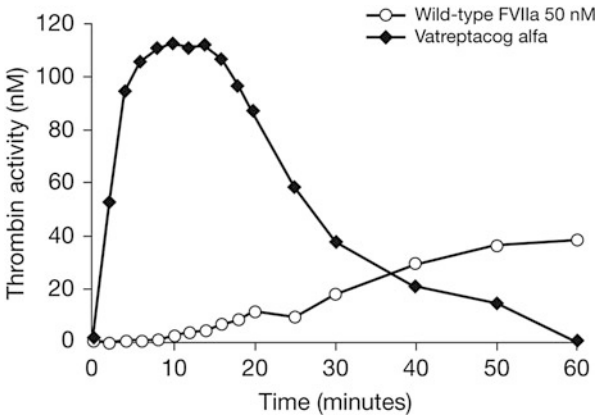


Fig. 3 Thrombin generation induced by vatreptacog alfa and other rFVIIa variants. Platelets were activated with the thrombin receptor agonist peptide SFLLRN, and then 50 nM of wild-type FVIIa (*open circle*) and vatreptacog alfa (*filled diamond*) were added along with protein mixture and calcium. Aliquots were removed and analyzed for thrombin amidolytic activity. Data shown are representative of three separate experiments. Reproduced from Persson et al. [55] Copyright (2001) National Academy of Sciences, USA

6 Clinical Trials with Vatreptacog Alfa: Design and Results

6.1 First Human Dose Trial [45]

The first clinical trial with vatreptacog alfa assessed the safety and pharmacokinetics (PK) of single doses in healthy male subjects. The trial had a randomized, double-blind, placebo-controlled, parallel-group, dose-escalation design; patients received vatreptacog alfa 5, 10, 20, and 30 μg/kg or placebo, with eight patients in each treatment group ($n = 32$). Progression to the next dose level depended on a blinded evaluation of safety data and PK.

Single doses of vatreptacog alfa $\leq$30 μg/kg were well tolerated. No serious or severe adverse events were reported, and there were no thromboembolic events. Three subjects (one in the placebo group) had a total of three five mild or moderate adverse events: headache [3], nausea, and vomiting. There were no clinically significant abnormal reports in any treatment group for laboratory safety values, vital signs, electrocardiogram, or physical examinations, and there were no local injection-site reactions. Based on FVIIa activity, vatreptacog alfa had a rapid initial distribution with a half-life of ~20 min, followed by a less rapid terminal phase with a half-life of ~3.1 h (Fig. 4). The total ($AUC_{0–24}$) and maximum (C_{5min}) exposure to vatreptacog alfa increased proportionally with dose. In addition, vatreptacog alfa was shown to be pharmacologically active based on coagulation-regulated parameters, suggesting the potential for a rapid induction of events leading to hemostasis and bleeding cessation.

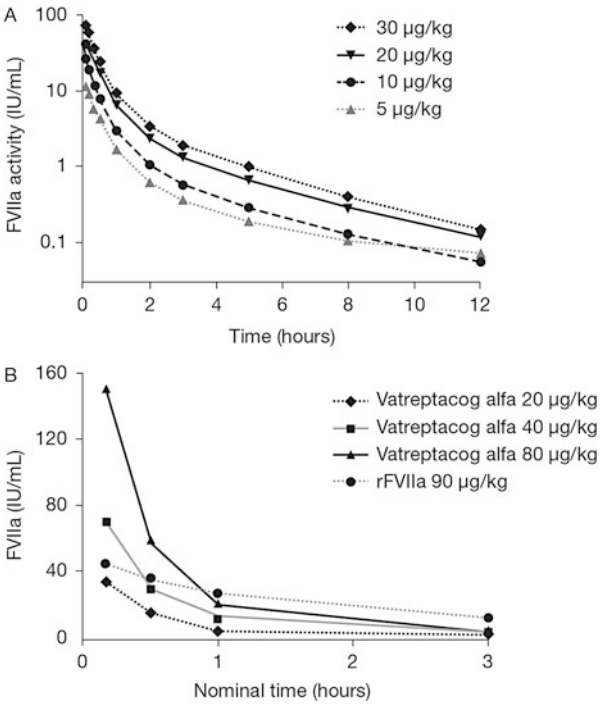


Fig. 4 Pharmacokinetic profiles (FVIIa activity) of vatreptacog alfa (**a**) mean single dose PK of vatreptacog alfa (5, 10, 20 or 30 µg/kg) in first human dose trial. Reproduced from Møss et al. [45]; (**b**) mean single dose PK of vatreptacog alfa (20, 40, or 80 µg/kg) and rFVIIa (90 µg/kg) in phase 2 trial. Reproduced from de Paula et al. [2]

Importantly, no antibodies to vatreptacog alfa were detected up to 3 months after a single dose in any of the vatreptacog alfa dose groups.

6.2 Phase 2 Dose-Escalation Trial (Adept™1) [2]

The phase 2, randomized, double-blind, active-controlled adept™1 trial evaluated the safety and preliminary efficacy of five escalating doses of vatreptacog alfa for the treatment of joint bleeds in hemophilia patients with inhibitors. Patients with at least two joint bleeds in the previous 6 months ($n = 51$) were randomized (4:1) to receive 1–3 doses per joint bleed (up to a maximum of five bleeds) of either vatreptacog alfa 5, 10, 20, 40, or 80 µg/kg or rFVIIa 90 µg/kg. The primary endpoint was frequency of adverse events; secondary endpoints included evaluations of immunogenicity, PK, and efficacy defined as the number of bleeds successfully controlled with a single treatment dose.

Results showed that vatreptacog alfa was well tolerated in the adept™1 trial, with a low frequency of adverse events in all dose groups. There were 15 serious adverse events (12 in the vatreptacog alfa groups) and one thrombotic event that were considered by investigators to be unrelated to treatment. The three adverse events considered to be related to vatreptacog alfa were mild. Laboratory

parameters revealed no safety concerns. The mean single-dose PK profiles of FVIIa activity versus time for vatreptacog alfa showed an exponential decline in activity following infusion (Fig. 4). The peak activity of vatreptacog alfa 80 μg/kg was three- to fourfold higher than that with rFVIIa 90 μg/kg, and vatreptacog alfa clearance was approximately three times faster than rFVIIa clearance. At 1 h post-dose, the mean plasma activities of all doses of vatreptacog alfa were below the level obtained following rFVIIa treatment.

A total of 95 joint bleeds were included in the analysis, 76 were treated with vatreptacog alfa and 19 with rFVIIa; 86 of 95 were controlled with 1–3 treatment doses. In a combined analysis of the vatreptacog alfa 20–80 μg/kg dose groups, 98% (41/42) of bleeds were controlled successfully with 1–3 doses, compared with 90% of bleeds treated with rFVIIa. The number of doses needed to control bleeding decreased with increasing dose of vatreptacog alfa, with 40% of bleeds effectively treated with a single 80 μg/kg dose.

The mean total number of vatreptacog alfa doses received during the trial was 3.6 and ranged from 1 to 15. Only one patient had >3 exposure days (EDs) to vatreptacog alfa. No antibody development was detected in any patients exposed to vatreptacog alfa in the trial for up to 28 days after the last treatment.

6.3 Phase 3 Adept™2 Trial [1]

The aim of the phase 3 adept™2 trial was to assess the efficacy and confirm the safety of vatreptacog alfa for treatment of bleeds in hemophilic patients with inhibitors. The randomized, double-blind, active-controlled, crossover trial enrolled 72 patients who had experienced at least five bleeds requiring treatment in the previous 12 months. Each bleeding episode was randomized to treatment with 1–3 doses of either vatreptacog alfa 80 μg/kg or rFVIIa 90 μg/kg, with 300 bleeds planned to be treated with vatreptacog alfa and 200 with rFVIIa. It was estimated that ≥15 patients had to have ≥10 days of exposure to vatreptacog alfa to sufficiently evaluate the potential risk of ADA development. The primary efficacy endpoint was effective bleeding control, and the main safety endpoint was immunogenicity.

Of 567 bleeds reported in the trial, 340 were treated with vatreptacog alfa. Both vatreptacog alfa and rFVIIa showed 93% efficacy in controlling bleeding with 1–3 doses at 12 h, including joint, mucocutaneous, muscle, soft tissue, and other bleeds. The mean number of doses administered to control bleeding within 9 h was significantly lower for vatreptacog alfa (2.42 doses) than for rFVIIa (2.52 doses; $p = 0.0304$).

Although the trial confirmed the efficacy of vatreptacog alfa, ADAs were detected in 8/72 patients (11%). The ADAs developed after <10 vatreptacog alfa EDs in seven of the eight patients and after 14–28 EDs in the remaining patient. This strong immunogenicity safety signal (discussed in detail in Sect. 7) was not

seen in previous clinical trials with vatreptacog alfa. Given the potential risks associated with ADAs, clinical development of vatreptacog alfa was discontinued.

7 Antidrug Antibody Development Observed in the Phase 3 Trial

In the phase 3 adept™2 trial, binding antibodies against vatreptacog alfa were detected in 8/72 (11%) exposed patients (Table 1). No antibody-positive patients had any concomitant medical conditions that would predispose to antibody development, and there were no differences in baseline characteristics between antibody-positive and antibody-negative patients [1]. Furthermore, when seven of the eight patients with anti-vatreptacog alfa antibodies were investigated by DNA analysis, none were found to have polymorphisms or mutations in the FVII gene (versus wild-type FVII), thus eliminating mutation in the endogenous protein as a causative factor underlying inhibitor development.

Four of the eight patients with ADAs against vatreptacog alfa, the ADAs developed low-titer cross-reactivity against rFVIIa; for one of these patients, the ADA also had in vitro neutralizing activity against vatreptacog alfa in a single blood sample taken at day 250 after first exposure [1]. There were no clinical manifestations of ADAs in any patient, including the patient who developed neutralizing antibodies, and all patients responded well to treatment with vatreptacog alfa and/or rFVIIa. Furthermore, all patients tested negative for neutralizing activity against endogenous FVIIa at all visits [1].

The immunogenicity findings were unexpected for two main reasons. First, vatreptacog alfa has 99% identity with rFVIIa, differing only by three amino acid substitutions, and rFVIIa is associated with a low risk of immunogenicity [1, 43]. Second, no immunogenic safety signals were detected in the animal immunogenicity studies or in previous clinical studies [1, 2, 45, 56]. Even so, most patients in the earlier phase 2 trial (adept™1) were only exposed to vatreptacog alfa once [2]; and consequently the overall exposure was markedly higher in the phase 3 adept™2 trial, though still limited to $\leq$10 EDs for most patients [1]. The longer exposure to vatreptacog alfa in adept™2 (1–28 EDs) versus adept™1 (1–5 EDs), together with the greater number of patients exposed in adept™2 ($n = 67$ versus $n = 46$ in adept™1), may explain why an immunogenicity signal was not apparent in the earlier trial [1].

Table 1 Overview of ADA development in the phase 3 adept™2 trial [1, 57]

Patient	Binding ADAs		Neutralizing ADAs		Total EDs in trial		Vatreptacog alfa EDs before ADA development[b]	Peak anti-vatreptacog alfa ADA titer
	Vatreptacog alfa	rFVIIa	Vatreptacog alfa	FVIIa	Vatreptacog alfa	rFVIIa[a]		
A	Positive	Negative	Negative	Negative	11[c]	3	1–3[c]	16
B	Positive	Negative	Negative	Negative	5	3	1–5	4
C	Positive	Negative	Negative	Negative	4	1	4	4
D	Positive	Negative	Negative	Negative	4	2	1–4	1
E	Positive	Positive	Negative	Negative	11	4	4–8	256
F	Positive	Positive	Negative	Negative	7	3	1–4	256
G	Positive	Positive	Negative	Negative	28	16	14–28	64
H	Positive	Positive	Positive	Negative	12[c]	5	5–8[c]	64

ADA antidrug antibody, *EDs* exposure days, *FVIIa* activated factor VII, *rFVIIa* recombinant activated factor VII

[a]Some patients were exposed to rFVIIa before the adept™2 trial and therefore had a higher number of cumulative EDs

[b]Number of vatreptacog alfa EDs before ADA development is presented, a range is given if there was no antibody assessment between EDs

[c]Patients A and H were previously exposed to vatreptacog alfa in the phase 2 trial and therefore had a higher number of cumulative EDs

7.1 Characteristics and Clinical Course of Anti-vatreptacog Alfa Antibodies

7.1.1 Assessment of ADAs

Blood samples for ADA testing were taken before first drug exposure, before dosing at every scheduled visit during the trial (i.e., at least every 3 months), and at least 1 month after last trial product administration [1].

To assess the presence of binding ADAs against vatreptacog alfa and rFVIIa, all blood samples were incubated with ^{125}I-labeled vatreptacog alfa or rFVIIa and analyzed in two validated radioimmunoassays (RIA) including antibody confirmation assays [1]. All samples that were confirmed positive for vatreptacog alfa- or rFVIIa-binding ADAs were further analyzed for neutralizing activity in two in vitro clotting assays that measured neutralization of vatreptacog alfa or endogenous human FVII including FVIIa [1]. To detect anti-vatreptacog alfa-neutralizing activity, vatreptacog alfa was added to all pretrial and trial blood samples; clot formation was initiated by adding soluble truncated recombinant TF, and time to clot formation was measured. Samples were analyzed in parallel for neutralizing activity against endogenous FVII by initiating clot formation with full-length TF, which activates endogenous FVII to FVIIa. Assay sensitivity was ~160 ng/mL and 1,500–3,000 ng/mL antibody for the binding and neutralizing assays, respectively, and all assays had a false-positive rate of 0.1% [1, 57].

Antibody specificity was analyzed in patients with rFVIIa cross-reactive antibodies. For this analysis, blood samples were incubated with an excess of the unlabeled antigen of interest in a competition RIA [1]. Antibody binding to radioactively labeled rFVIIa was competed with unlabeled vatreptacog alfa, while labeled vatreptacog alfa competed for binding with unlabeled single-mutation analogs [57]. A significant reduction of antibody binding in the competition assay indicated a shared antibody epitope between either rFVIIa and vatreptacog alfa or the single-mutation analog and vatreptacog alfa [57].

7.1.2 Characterization of ADAs

The development of ADAs over time for each of the eight ADA-positive patients is shown in Fig. 5. All but one of these patients (88%) developed ADAs to vatreptacog alfa after 1–8 EDs, while the remaining patient first tested positive for ADAs between 14 and 28 EDs (Table 1) [57]. Peak antibody titers ranged from 1 to 256 (Table 1) [1, 57]. There was no further increase in ADA titer following additional vatreptacog alfa exposure in two patients (patients A and E), while two further patients showed an increase in ADA levels following additional exposure (patients F and H), suggesting that the maximum antibody titer had not yet been reached [57].

The four patients with the highest anti-vatreptacog alfa ADA titers (patients E–H) also developed cross-reactivity to rFVIIa (Table 1) [1, 57]. All anti-rFVIIa

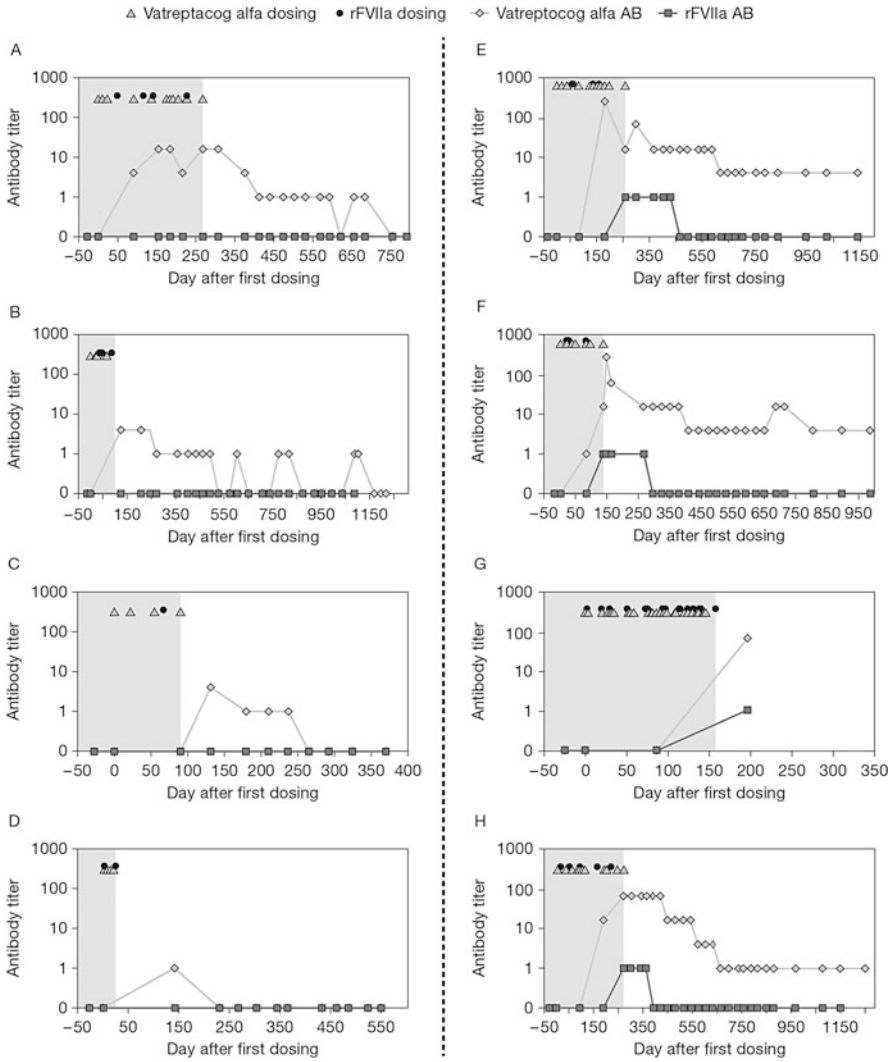


Fig. 5 The time course of ADA development in patients from the phase 3 adept™2 trial. (**a**–**d**) ADA development over time in four patients who did not develop rFVIIa cross-reactive antibodies. (**e**–**h**) ADA development over time in four patients who developed rFVIIa cross-reactive antibodies. Antibody titers for vatreptacog alfa and rFVIIa are shown during the trial dosing period (*gray-shaded area*) and during the post-dosing follow-up period (*white-unshaded area*). Treatment days on which patients received vatreptacog alfa and rFVIIa are also shown. rFVIIa, recombinant activated factor VII. Reproduced from Mahlangu et al. [57] with permission from John Wiley and Sons

antibody titers were low and developed later than anti-vatreptacog alfa antibodies, coinciding with the last vatreptacog alfa treatment in three patients (patients E, F, and H) and developing 1 month after the last vatreptacog alfa exposure in the fourth

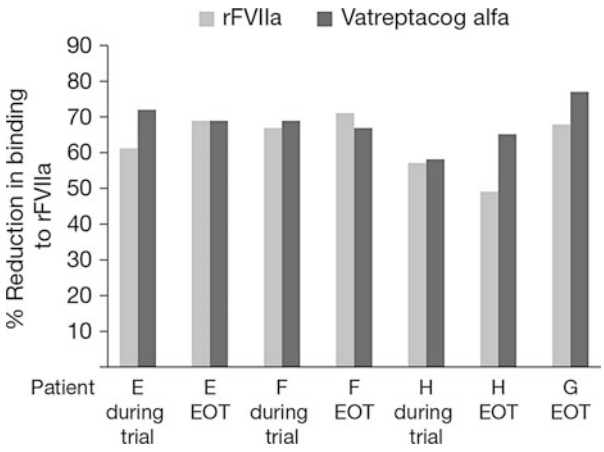


Fig. 6 Reduction of antibody binding to rFVIIa in a competition assay with vatreptacog alfa and rFVIIa in four patients with rFVIIa cross-reactive antibodies rFVIIa binding was competed with an excess of either unlabeled vatreptacog alfa or rFVIIa. The percentage reduction in binding to radiolabeled rFVIIa is presented. EOT, end of trial; rFVIIa, recombinant activated factor VII. Reproduced from Mahlangu et al. [57] with permission from John Wiley and Sons

(patient G). One reexposure to vatreptacog alfa after development of anti-rFVIIa ADAs in patients E, F, and H did not lead to an increase in rFVIIa antibody titer [57]. In all four patients with cross-reactive ADAs, the antibodies were reactive to structures shared between vatreptacog alfa and rFVIIa; there was no subpopulation of antibodies that were specific only to rFVIIa (Fig. 6) [57].

In vitro neutralizing activity against vatreptacog alfa was detected in one patient (patient H) in a single blood sample taken at day 250 after first vatreptacog alfa exposure [1, 57]. This sample neutralized 40% of the in vitro activity of 0.53 ng vatreptacog alfa present in the antibody assay which corresponds to neutralization of 0.2 ng vatreptacog alfa. The cut point for a positive neutralization result in this assay was 29% corresponding to neutralization of 0.15 ng vatreptacog alfa. However, all subsequent samples from this patient were negative for vatreptacog alfa-neutralizing antibodies [57].

Vatreptacog alfa ADAs were of the IgG subtype in seven of the eight patients, and isotyping results were inconclusive in the remaining patient (patient G) [57]. In patients F and H, the antibody maturation profile was consistent with the typical immune maturation pattern displayed by IgG antibodies: increased titer, cross-reactivity, and one sample of neutralizing in vitro activity. As binding ADAs can precede the development of neutralizing ADAs [26], it is possible that any further exposure to vatreptacog alfa in these patients may have triggered the development of high-titer antibodies that neutralized rFVIIa or endogenous FVIIa [57].

After completion of adept™2, 7/8 (88%) ADA-positive patients entered an ongoing follow-up study to monitor the course of the ADAs [57]. Blood samples are collected and tested on a monthly basis. Patients are followed up until two consecutive samples test negative for ADAs, or until further follow-up is deemed

unnecessary by the study sponsor or investigators. Anti-vatreptacog alfa ADA titer declined in all patients following discontinuation of vatreptacog alfa treatment (Fig. 5). At the time of writing, the ADA-positive patients have been followed for 3 years. Three of the seven patients are now negative for anti-vatreptacog alfa ADAs. One patient (patient B) has been fluctuating between being antibody negative and positive (titer 1). Cross-reactivity to rFVIIa disappeared in all three tested patients (E, F, and H) at 124, 156, and 208 days, respectively, after last vatreptacog alfa exposure, despite reexposure to FVII according to local practice [57]. One patient (patient G) did not participate in the follow-up trial, and so his ADA status remains unknown.

7.1.3 Pharmacokinetics of ADAs

To determine the potential impact of anti-vatreptacog alfa ADAs, PK profiles were assessed up to 8 h post-dose in four ADA-positive patients (patients A, E, F, and H). All PK measurements were taken after a single intravenous injection of vatreptacog alfa given to the patients in a non-bleeding state [57].

There was no indication of reduced recovery of FVIIa activity 10 min post-dose in any of the ADA-positive patients tested (Fig. 7). The reason for this lack of effect on in vivo recovery is not clear, but it could be attributable to the high sensitivity of the neutralizing antibody assay: a neutralization of 0.2 ng vatreptacog alfa may be masked at the activity levels evident in the patients tested (e.g., the post-dose FVIIa activity levels in patient H corresponded to a concentration of 596–253 ng/mL) [57].

Somewhat surprisingly, all four ADA-positive patients in whom PK was investigated showed prolonged elimination of vatreptacog alfa when compared with the mean PK profile of ADA-negative patients from the earlier phase 2 trial (adept™1 [2]) [57].

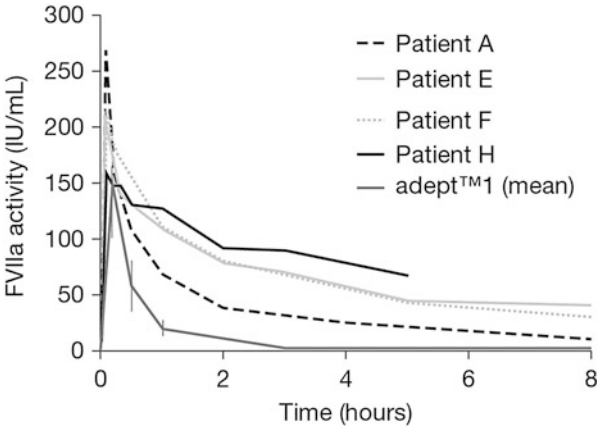


Fig. 7 PK profiles of four patients (patients A, E, F, H) with ADAs against vatreptacog alfa. Individual PK profiles are presented, together with the mean (±SD) FVIIa activity in ADA-negative patients in the phase 2 study (adept™1). Reproduced from Mahlangu et al. [57] with permission from John Wiley and Sons

8 Investigating the Immunogenicity of Vatreptacog Alfa Using Immunogenicity Prediction Tools

Since the only differences between rFVIIa and vatreptacog alfa are three amino acids, it is highly possible that the observed ADAs are due to the change in sequence. As the development of a mature ADA response is a T cell-dependent process [17–21], the assessment of vatreptacog alfa immunogenicity focused on the role of T-cell epitopes and their relationship with MHC-II molecules. Therefore, the role of T cells, T-cell epitopes, and MHC-II molecules in the generation of immune responses is briefly reviewed.

8.1 The Role of T Cells and T-Cell Epitopes in Immunogenicity

Several processes involving interaction between T-helper cells, B-cells, and antigen-presenting cells (APCs; e.g., dendritic cells) are necessary for antibody generation [18, 21]. If recognized by an APC, such as a dendritic cell, the infused therapeutic protein is taken inside the cell where it is processed and cleaved into peptides [18, 21]. These therapeutic protein-derived peptides then bind to MHC-II molecules, and the resulting peptide–MHC-II complexes are transported to the surface of the dendritic cell [18, 19, 21]. If antibodies are to be generated, the peptide must be presented as peptide–MHC-II complexes on dendritic cells and then be recognized by T-cell receptors (TCRs) on T-helper cells [18, 19, 21]. TCRs to self-proteins have generally been eliminated; thus, only peptides foreign to that individual will be recognized. Peptides that mediate a sustained association between dendritic cells and T cells are called T-cell epitopes [19], and T-cell epitope content is one of the major contributory factors to antigenicity [20]. The binding strength of a particular T-cell epitope to MHC-II molecules largely defines its immunogenicity, as epitopes with higher binding affinities are more likely to be presented on the surface of dendritic cells and thus recognized by TCRs [20]. Effective and sustained association between dendritic cells and T cells through formation of an MHC-II–peptide–TCR complex is necessary to induce stimulation and proliferation of T-helper cells [19, 21]; in turn, this leads to cytokine secretion and activation of B-cells that have also taken up the factor and presented the peptides on the same MHC-II molecules [19, 21]. A further series of steps leads to maturation of B-cells into factor-specific antibody-secreting plasma cells [18, 20, 21, 27].

Peptide binding to MHC-II molecules expressed by the host is one of the key determinants of whether the peptide will trigger an immune response [20, 33]. In this regard, the genotype of the highly polymorphic HLA determines the sequence of peptides that can be bound. As ADA development is thought to be influenced by a patient's HLA type [34], it is reasonable to expect that therapeutic proteins carrying promiscuous HLA class II-binding peptides (i.e., peptides that bind to

numerous HLA class II alleles with high affinity [58, 59]) have the potential to elicit an antibody response. These observations suggest that the affinity of mutant peptides for an individual's HLA class II repertoire may predict immunogenicity risk in that individual [60].

Immunogenicity prediction assessment of vatreptacog alfa sought to determine whether ADA development in adept™2 could have been elicited by presence of high-affinity mutant peptide sequences in vatreptacog alfa for individual patients' HLA class II profiles [98].

8.2 *Tools for Predicting T Cell-Dependent Immunogenicity*

Immunogenicity prediction typically involves more than one approach, as all methods have strengths and limitations [17] and possess different levels of predictive strength. There are currently no validated methods for predicting immunogenicity of therapeutic proteins in the clinic. However, in silico (i.e., using computers or computer simulations) tools and in vitro assays using recombinant proteins and living cells have been used to evaluate the T-cell epitope content, peptide–MHC-II interactions, and immunogenicity potential of therapeutic proteins [17, 22, 61]. Many of these tools have been developed and optimized relatively recently. As discussed in Sect. 5.4, early immunogenicity risk evaluation was performed for vatreptacog alfa using the tools available at that time; however, the unique case of vatreptacog alfa represents a valuable opportunity to evaluate the predictive value of the newer and different immunogenicity tools. An immunogenicity prediction strategy was developed for vatreptacog alfa using in silico, in vitro, and ex vivo tools to address the key processes of antigen presentation: (1) binding of the mutant peptides within the vatreptacog alfa molecule to MHC-II (in silico and in vitro), (2) protein and peptide processing and display by MHC-II on dendritic cells (in vitro), and (3) TCR recognition of the MHC-II–peptide complex (ex vivo).

8.3 *In Silico and In Vitro Peptide–MHC-II Binding Experiments*

In silico analysis involves the creation of computational models or simulations in order to make binding affinity predictions for specific peptide – HLA allele pairs [62]. Sequence analysis by in silico screening facilitates identification of potential T-cell epitopes [63] and is often used as a first step in the immunogenicity screening of therapeutic proteins [17]. Peptide–MHC-II binding is largely determined by a core residue of 9–10 amino acids within T-cell epitope sequences [17]. Therefore, it is computationally possible to predict T-cell epitopes based on peptide amino acid sequences if there is sufficient information available for peptides that are known to

bind to a particular MHC-II variant [17]. Databases such as the Immune Epitope Database and Analysis Resource (www.iedb.org) provide the basis for developing in silico T-cell epitope prediction tools [17].

In silico methods for immunogenicity prediction were still in their infancy during the early stages of vatreptacog alfa development (see Sect. 5.4). Since that time, a number of computational approaches to T-cell epitope prediction and immunogenicity risk assessment have been developed and optimized [64–74]. Such approaches are now well accepted in the field of vaccine design [63] and widely used to identify key epitopes that trigger autoimmunity [75]. In silico screening has also demonstrated predictive accuracy in the immunogenicity testing of protein therapeutics by showing an association between peptide–MHC-II binding and T cell-dependent immunogenicity in the clinic [17, 20, 70, 74]. These studies contribute to the growing body of evidence suggesting that in silico analysis predicts peptide–MHC-II interactions, serves as a first-line method to evaluate immunogenicity risk, and can even aid in the design of therapeutic proteins.

After defining putative T-cell epitopes with MHC-II binding affinity using in silico analysis, the results can be confirmed and validated using peptide/MHC-II binding assays. A number of different MHC-II binding assays can be used, including competition binding assays, real-time kinetic measurements, and direct binding assays [17].

8.3.1 Strengths and Limitations of In Silico and In Vitro Binding Experiments

In silico and in vitro peptide–MHC-II binding experiments focus on the contribution of T cells to ADA development. They are based on the binding of T-cell epitopes to HLA class II alleles, are cost-effective and relatively easy to use, and have a short time course that fits well into research and development programs for protein biotherapeutics [76–78]. Additionally, the in silico tools available today can rapidly screen and compare many different protein sequences and HLA alleles for putative T-cell epitopes [20, 76, 77] and can thus markedly reduce downstream in vitro testing [17].

Current in silico analyses offer a reasonable level of accuracy when predicting peptide–HLA class II-binding affinities for a large number of HLA alleles. However, the major limitation of both in silico and in vitro peptide–MHC-II binding analyses is their tendency to overpredict the number of T-cell epitopes [77], as they do not account for other important biological factors that limit the number of true functional T-cell epitopes [76]. These additional factors include protein and peptide processing in the APC, T-cell phenotype, TCR affinity for the peptide–MHC-II complex, and induction of T-cell tolerance to non-germline peptides [17]. Furthermore, neither approach proves that a peptide predicted to bind to MHC-II molecules will actually initiate a T-cell response [77, 78]. However, the protein needs to contain T-cell epitopes in order to induce a high affinity ADA response. These epitopes can be

identified by in silico or in vitro peptide–MHC-II binding analyses with a certain level of false positives and a very low number of false negatives.

8.3.2 In Silico and In Vitro Assessment of Vatreptacog Alfa

For vatreptacog alfa, the aim of in silico peptide–MHC-II binding experiments was to identify potential neo-T-cell epitopes within the mutant peptides, using the wild-type FVIIa as a reference sequence to exclude endogenous, tolerized epitopes. The in silico predictions were made using two different algorithms: NetMHCIIpan v2.1 [79] and NetMHCII v2.2 [80]. While NetMHCIIpan v2.1 evaluates the binding affinity of peptides to all HLA-DRB1 alleles, NetMHCII v2.2 evaluates peptide binding affinity to the HLA-DRB1 alleles as well as the most frequent HLA-DP and HLA-DQ alleles. As peptide–MHC-II binding affinity is governed mainly by a core sequence of 9–10 amino acids within potential T-cell epitopes [17] and partly by the three flanking amino acid residues on each side of the binding core sequence, the in silico analyses were conducted using the default peptide length (15 amino acids with a nine amino-acid binding core sequence; [98]).

In silico analysis generated a heat map in which colors correspond to the binding affinity of a given peptide/HLA allele pair [98]. After subtracting the HLA class II-binding peptides that contained only wild-type FVIIa sequences, various neo-epitopes that were predicted to bind to HLA-DR, HLA-DP, and HLA-DQ alleles were identified for vatreptacog alfa. The heat map of predicted peptide–MHC-II affinities showed that peptides with the E296V and M298Q mutations bind to several HLA class II alleles with high affinity and therefore contain potential T-cell epitopes; in contrast, the E158D mutation does not give rise to HLA class II-binding peptides (Fig. 8).

The in silico findings were then confirmed using an in vitro direct peptide binding assay, which assessed the binding affinities of vatreptacog alfa peptides for ten HLA variants representing the major HLA class II molecules. The assay chosen for this purpose was the luminescent oxygen channeling immunoassay (LOCI), a two-bead assay system [81], performed by Professor Søren Buus at Copenhagen University. Donor beads were coated with streptavidin, a capture reagent that binds various biotinylated HLA-DRB1 alleles [81], and acceptor beads were coated with monoclonal antibodies specific to fully folded HLA-DRB1. Donor and acceptor beads formed pairs in the presence of vatreptacog alfa peptides that bind to the HLA class II variants, and photochemically triggered chemiluminescence allowed detection of these bead pairs [81–83]. This in vitro analysis validated the in silico findings.

While peptide–MHC-II binding is a necessary step in ADA development, it is not sufficient to induce an immune response on its own: recognition of the peptide–MHC-II complex on dendritic cells by TCRs, followed by T-cell activation, is also required. Therefore, the immunogenicity of vatreptacog alfa was tested further using human cell-based assays that assess peptide presentation and T-cell responses.

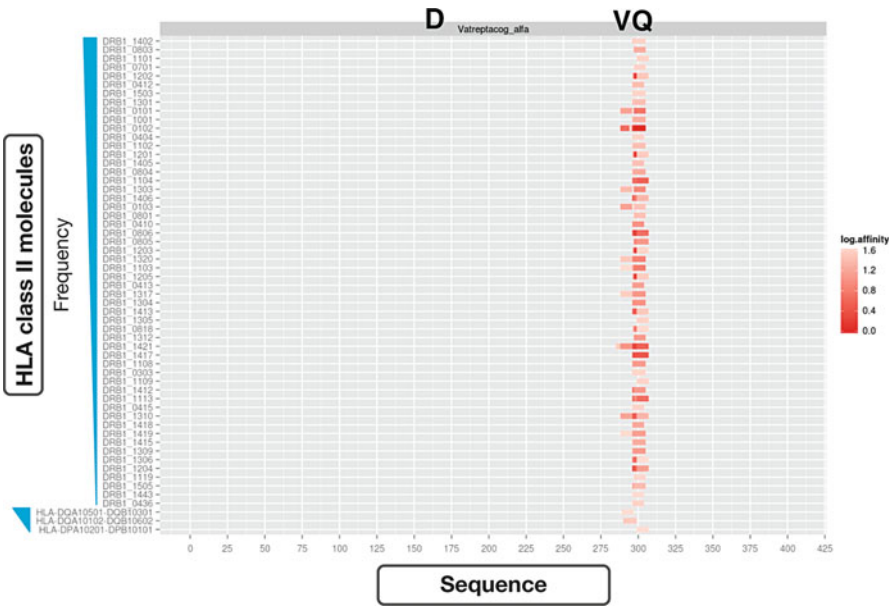


Fig. 8 In silico predicted HLA class II-binding peptides for vatreptacog alfa. Graphic representation of the predicted HLA class II-binding peptides in vatreptacog alfa and their binding affinities to the tested HLA-DR, HLA-DP, and HLA-DQ alleles. Each *red bar* represents a HLA class II-binding peptide with its position in the vatreptacog alfa sequence (*x* axis) and its binding affinity to a given HLA allele (*y* axis). Native FVIIa has been used as reference sequence, thus only novel peptides binding to HLA molecules are identified

8.4 Human Cell-Based Peptide Processing, Presentation, and Activation Assays

While in silico analysis can predict putative T-cell epitopes and reveal their binding affinities for MHC-II molecules, cell-based assays that use T cells and APCs derived from the peripheral blood (peripheral blood mononuclear cells [PBMCs]) of human donors address several processes relevant to T-cell activation: APC antigen uptake and processing, the stability of peptide–HLA class II binding, identification of a T-cell repertoire toward the peptide in question, and subsequent T-cell activation resulting in proliferation and cytokine release.

8.4.1 Strengths and Limitations of Human Cell-Based Assays

Although more technically demanding than in silico analysis, human cell-based assays provide more biologically relevant measurements of T-cell epitopes, predict the relative immunogenicity potential of the epitopes, and, in some cases, correlate

well to immunogenicity seen in the clinic [17, 77, 78]. As a result, they offer a greater level of predictive accuracy and reliability than in silico analyses in immunogenicity risk assessments.

However, cell-based assays are not without limitations. For example, to avoid overestimating the T-cell response to the peptides under investigation, PBMC-based T-cell activation assays measuring cytokine release must be fine-tuned to negate the effects of other cytokine-secreting cells present in the PBMC preparation [17]. Indeed, a major challenge with these assays is the need to find an appropriate balance between minimizing irrelevant immune responses and supporting the immune response of interest [17]. A large number of individual blood donors and large blood volumes are required to perform a dendritic cell presentation assay or a T-cell activation assay [17]. This can be very cost prohibitive for many laboratories and is also time-consuming.

Despite these challenges, however, and despite the improvements that are still needed, human cell-based assays effectively strengthen the predictive value of immunogenicity analyses by evaluating the final impact of the potential T-cell epitopes identified in silico [17].

8.4.2 In Vitro Dendritic Cell Presentation Assay

In vitro profiling of vatreptacog alfa- and FVIIa-derived peptides was performed to determine whether mutant peptides in vatreptacog alfa are presented on dendritic cells and therefore have the potential to induce T-cell activation. This was achieved using human cell-based antigen presentation assays, which investigate peptide processing in the dendritic cells and presentation of peptides by HLA class II molecules on the dendritic cell surface [76]. One of the aspects of peptide processing reflected in the assay is intracellular peptide trimming. MHC-II-binding peptides usually range between 12 and 25 amino acids in length and can be represented as linear sequences including the core binding sequence of 9–10 amino acids. This core binding sequence fits into a binding groove on the MHC-II molecule; however, additional amino acids flank the core binding sequence outside the groove at each end [84, 85]. The in vitro dendritic cell presentation assay addresses intracellular peptide trimming and allows clusters of peptides to be defined that share a common stretch but differ in the lengths of the amino acid sequences that lie outside the MHC-II binding groove.

The assay used human monocyte-derived dendritic cells, which are widely considered to be the major APC in human immune responses [86, 87]. Immature dendritic cells are present in peripheral tissues, where they effectively take up both foreign and “self”-proteins and process them intracellularly into small peptides for antigen presentation via MHC-II to the adaptive immune system [88–90]. When dendritic cells capture antigens in the presence of “danger signals” [87, 88], they switch to a mature state, following which they become very efficient in presenting

protein-derived peptides to naïve T-helper cells in the context of their MHC-II repertoire [67, 87, 88, 91].

Peripheral blood samples obtained from healthy donors of unknown HLA type were processed to generate human monocyte-derived dendritic cells. Human monocytes were incubated in vitro with IL-4 and granulocyte–macrophage colony-stimulating factor (GM-CSF) for 5 days, to obtain dendritic cells, which were then pulsed with antigen. Exposure of the dendritic cell to (and subsequent uptake of) vatreptacog alfa was followed by overnight lipopolysaccharide-induced maturation of the dendritic cell. Peptide processing and binding to HLA class II in the mature dendritic cell led to cell surface presentation of peptide–HLA class II complexes, after which the membrane proteins were solubilized using cell lysis. Peptide–HLA class II complexes were then immunocaptured using HLA-DR-coated beads. HLA class II-binding peptide sequences were eluted and characterized by mass spectrometry (MS). The eluted HLA class II-binding peptide sequences had undergone intracellular peptide trimming and thus represented clusters of peptides sharing a common stretch but differing in the lengths of their amino- and carboxy-terminal extensions.

The results showed that HLA-DR-displayed peptides could be grouped into distinct clusters distributed throughout the vatreptacog alfa heavy chain, with one cluster overlapping the predicted HLA class II-binding peptides from the initial in silico and in vitro analyses (data not shown). These results thus correlated with the in silico findings and demonstrate that the peptides predicted to bind HLA class II were in fact processed by the dendritic cells. Importantly, the data showed that peptides spanning positions 296 and 298 were presented on the cell surface of APCs. Therefore, this analysis substantiates the hypothesis that the mutations in these positions could be the trigger of the ADAs that developed in some patients.

8.4.3 Ex Vivo T-Cell Activation Assays

Ex vivo T-cell activation assays were conducted to explore potential T-cell responses to the mutations in the modified sequences of vatreptacog alfa. These assays are based on human PBMCs containing CD4+ T cells and APCs from healthy blood donors or the desired patient population [76, 77]; T-cell activation is typically evaluated by T-cell proliferation (as measured by incorporation of radiolabeled thymidine) and/or cytokine (e.g., IL-2) release [67, 77, 78, 92]. Immunogenic potential can be assessed based on the magnitude and frequency of donor T-cell responses to the investigated protein or peptide sequences [77, 78] and typically includes a panel of healthy donors representing various HLA types.

The ex vivo T-cell activation assays for vatreptacog alfa assessment were performed at Antitope Ltd (Cambridge, UK). All assays were performed using peripheral blood leucocytes from 50 healthy donors who had been HLA typed and

selected to provide a proportional representation of HLA alleles found in the world population. Donor cells were incubated with full-length FVIIa and vatreptacog alfa peptides; T-cell proliferation and IL-2 secretion – both markers of T-cell activation – were then estimated using a tritiated thymidine uptake assay and an IL-2 ELISpot assay, respectively [93]. The percentage of responders to the full-length proteins or peptides was then determined. According to Antitope, a frequency of <10% is associated with low or no risk for immunogenicity in the clinic [94].

Full-length wild-type FVIIa and vatreptacog alfa induced a T-cell response in <10% of the donor cohort; thus, there was no significant difference in responders between full-length vatreptacog alfa and full-length wild-type FVIIa. This suggests that ex vivo T-cell activation assays were not sensitive enough to predict the immunogenicity of vatreptacog alfa when full-length molecules were used.

In order to increase the sensitivity of the assay, short (15 amino acids) and long (23–25 amino acids) vatreptacog alfa peptides spanning the protein sequences of interest were used. This more diverse set of mutant peptides was designed to fit two different versions of the T-cell assay (a time course assay and a peptide assay) (Table 2). The long peptides were designed to allow the point of mutation to be located in any of the nine amino acid positions in the binding core sequence with a three-amino-acid flanking sequence at each end of the peptide. The long peptides were tested in the time course assay to allow for antigen uptake and processing of the peptides. This approach facilitates enhanced presentation of mutant peptides by HLA class II molecules when donor PBMCs are incubated with the wild-type or

Table 2 Short (15 amino acids) and long (23–25 amino acids) peptides used in T-cell activation assays

Peptide number	Peptide name	Amino acid sequence
Time course assay		
1	158_V	SKPQGRIVGGKVCPKGECPWQVL
2	158_D	SKPQGRIVGGKDCPKGECPWQVL
3	296_E	GQLLDRGATALELMVLNVPRLMT
4	296_V	GQLLDRGATALVLMVLNVPRLMT
5	298_M	LLDRGATALELMVLNVPRLMTQD
6	298_Q	LLDRGATALELQVLNVPRLMTQD
7	298_F	LLDRGATALELFVLNVPRLMTQD
8	296_E_298_M	GQLLDRGATALELMVLNVPRLMTQD
9	296_V_298_Q	GQLLDRGATALVLQVLNVPRLMTQD
Peptide assay		
10	296_E_298_M	ATALELMVLNVPRLM
11	296_V_298_M	ATALVLMVLNVPRLM
12	296_E_298_Q	ATALELQVLNVPRLM
13	296_E_298_F	ATALELFVLNVPRLM
14	296_V_298_Q	ATALVLQVLNVPRLM
15	HA307-319 (pos. ctl.)	PKYVKQNTLKLAT

mutant peptides. The 15-amino-acid peptides that were identified and based on the in silico algorithm were tested in the peptide assay where the peptides are loaded directly onto the HLA class II on the cell surface of the APCs. In turn, these approaches increase the density of the mutant peptide in question on the dendritic cell surface and therefore increase the chance and strength of T-cell activation if the peptide is immunogenic.

In the second analysis, three 15-amino-acid mutant peptides from vatreptacog alfa were found to trigger T-cell activation: the E296V, M298Q, and double E296V/M298Q mutants produced a higher T-cell response than wild-type FVIIa, while the E158D mutant did not (Fig. 9). This difference was more pronounced when longer peptides were used. These results are consistent with the heat map generated by in silico analysis (Fig. 8) and confirm that the E296V and M298Q mutant peptides bind with high affinity to several HLA class II alleles, while the E158D mutant does not. According to the in silico prediction analysis, peptides with the 298F mutation bind with an increased affinity compared to 298Q. Therefore peptides with the 298F mutation were added to the assays to explore how this high affinity would translate in the T-cell assays. Actually, the peptides with the 298F mutation had a slightly increased T-cell response compared to the peptides with the 298Q mutation.

8.5 HLA Typing of ADA-Positive Patients

Following discontinuation of the adept™2 trial, seven of the eight patients who developed ADAs against vatreptacog alfa and who participated in the follow-up study were HLA typed [57]. Saliva samples were obtained from all seven patients, and HLA typing was performed using polymerase chain reaction sequence–specific oligonucleotide probe (PCR–SSOP) to resolve major allele groups into four digits [57]. The purified genomic DNA was amplified using PCR and incubated with a panel of different oligonucleotide probes that have distinctive reactivity with different HLA types [57]. The HLA class II loci of interest were DRB1, DRB3/DRB4/DRB5, DPA1, DPB1, DQA1, and DQB1.

As expected, due to the promiscuity of the predicted T-cell epitopes, ADA-positive patients did not share a common HLA class II allele, suggesting that no single allele was responsible for ADA development (Table 3). However, all ADA-positive patients had at least one HLA class II allele that binds to a vatreptacog alfa peptide with high affinity. HLA class II alleles identified in six of the seven patients were represented in the healthy donor cohort used for the T-cell activation assays; also, as expected, HLA class II alleles that occur frequently in the population (e.g., DRB1*0701) also appeared in the ADA-positive patients. Interestingly, according to the in silico analysis, these alleles did not bind with high affinity to vatreptacog alfa. However, the double-mutant E296V/M298Q peptides from vatreptacog alfa bound with high affinity to at least one HLA class II

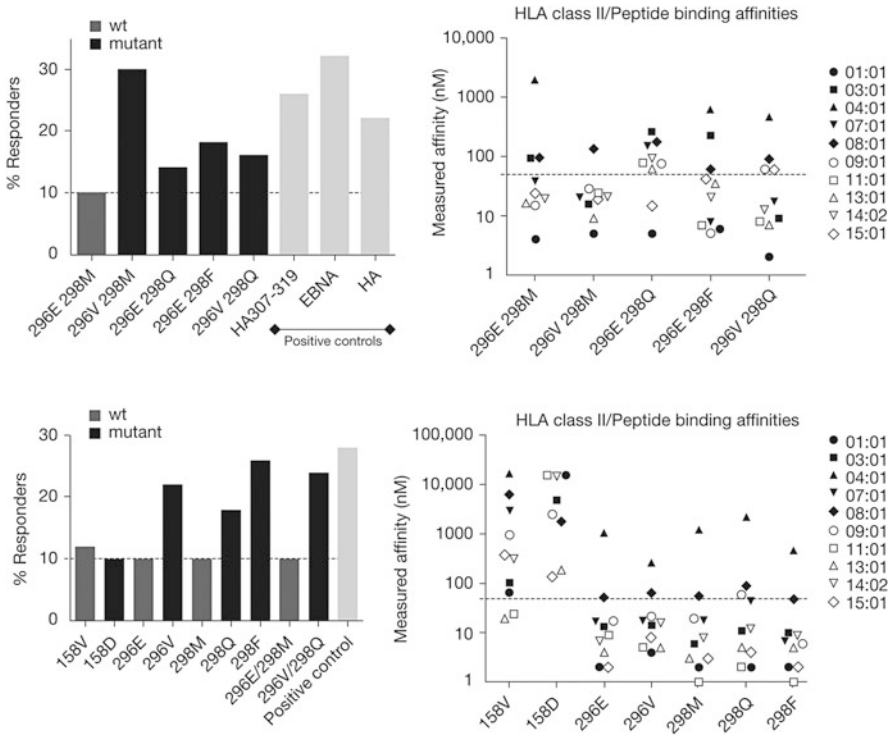


Fig. 9 Ex vivo T-cell activation assays using (**a**) 15-amino-acid peptides and (**b**) long peptides (23–25 amino acids). *Left panels*: Percentage of responders in a T-cell proliferation assay in a cohort of 50 healthy donors. Wild-type peptides all induced a response in ~10% of the cohort, and this was used as a threshold. The E296V, M298Q, and double E296V/M298Q mutants showed an increase in the percentage of responders (versus wild-type FVIIa), whereas the E158D mutant did not. This increase in responders was evident for 15-amino-acid mutant peptides (**a**) but much more pronounced for long peptides (**b**). Right panels: Peptide–HLA class II binding affinities. (**a**) The first peptide on the *x* axis is wild-type FVIIa; the others are different mutant peptides derived from vatreptacog alfa. Wild-type and all mutant peptides bind with similar affinity to the different HLA class II molecules. (**b**) On the *x* axis, peptides 158V, 296E, and 298M are wild-type FVIIa; the others are different mutant peptides derived from vatreptacog alfa. Wild type and mutants spanning the amino acid positions 296–298 bind with very high affinity to the different HLA class II molecules, while peptides spanning position 158 show a lack of HLA binding affinity

allele in all ADA-positive patients. One limitation of this analysis is that only ADA-positive patients from adept™2 were HLA typed. Therefore, it is not possible to conclude whether there is a set of HLA class II alleles that occur more frequently in ADA-positive (versus ADA-negative) patients.

Table 3 HLA typing of patients with anti-vatreptacog alfa ADAs in the adept™2 trial

Patient	HLA-DRB1	HLA-DRB3	HLA-DRB4	HLA-DRB5	HLA-DQB1	HLA-DPB1	HLA-DQA1	HLA-DPA1
A	*01:02/43/46	NT	*01:01/03/06/08	NT	*05:01/12/18	*01:01	*01:01/04/05	*01:03
	*04:04/23/108/118/120N/121/149	NT	NT	NT	*04:02	*18:01	*03:01/02/03	*02:01
B	*04:06/144	NT	*01:03	NT	*03:02P	*02:01P	*01:03/10	*01:03P
	*08:03P	NT	NT	NT	*06:01/43	*05:01:01G	*03:01:01G	*02:02/05
C	*11:04	*02:02/27/28/29N	*01:01/03/06	NT	*02:02	*04:01	*02:01	*01:03
	*07:01	NT	NT	NT	*03:01/27/28/29/35/42/47	*10:01	*05:05/09/11	*02:01
D	*13:01/105/112/117/148/153	*02:02/27/28/29N	*01:01/03/06	NT	*06:03/41	*01:01	*01:03/10	*02:01
	*07:01	NT	NT	NT	*02:02	*13:01/107:01	*02:01	NT
E	*13:01/105/109/112/117/138/143/146/147/148/153	*02:02/27/28/29N	NT	NT	*06:02	*01:01	*01:03/10	*02:01
	*08:04	NT	NT	NT	*04:02	*14:01	*04:01	*02:02/05
F	*03:01/50/68N/82/83/86	*01:01/11	NT	NT	*02:01	*01:01	*05:01/05/09/11	*01:03
	*11:01/95/97/100/117/133/134/140/141	*02:02/27/28/29N	NT	NT	*03:01/27/28/29/35/42/47	*04:01	NT	*02:01
H	*15:02/80N	*03:01	NT	*01:02/08N	*05:01/12/18	*05:01	*01:01/04/05	*02:01
	*12:02/37	NT	NT	NT	*03:01/27/28/29/35/42/47	*13:01/107:01	*06:01	*02:02/05

Reproduced from Mahlangu et al. [57] with permission from John Wiley and Sons

9 Vatreptacog Alfa and Immunogenicity: What Have We Learned?

ADAs developed against vatreptacog alfa in 8/72 (11%) treated patients in the phase 3 adept™2 trial. Of these eight patients, the antibodies in four developed cross-reactivity against rFVIIa, of which one developed in vitro neutralizing activity against vatreptacog alfa [1, 57]. No effects of the ADAs on clinical efficacy were manifest in any patients, and ADA titer declined in all cases following discontinuation of vatreptacog alfa treatment. Cross-reactivity to rFVIIa also disappeared after last exposure to vatreptacog alfa, despite reexposure to rFVIIa according to local standard care. Vatreptacog alfa ADAs were of the IgG subtype in all patients who had conclusive immunoglobulin isotype determination, and all ADAs were specific for vatreptacog alfa rather than rFVIIa [57]. PK assessment of ADA-positive patients suggested prolonged elimination of vatreptacog alfa but gave no indication of reduced FVIIa levels [57].

Vatreptacog alfa ADAs did not appear to be caused by FVII polymorphisms or mutations.

The binding of peptides to HLA class II molecules is a key factor in determining whether the peptides will trigger an immune response [20, 33]; ADA development is therefore influenced by an individual's HLA class II repertoire [34], and immunogenicity risk may be predicted by peptide–HLA class II binding affinity [60, 98]. Post hoc immunogenicity prediction studies using in silico, in vitro, and ex vivo tools were undertaken to investigate whether ADA development in the eight patients from adept™2 could have been triggered by the affinity of the mutant peptide sequences in vatreptacog alfa for individual patients' HLA class II profiles [98]. The results of these studies confirmed that vatreptacog alfa peptides with the E158D mutation lack the necessary condition (i.e., high-affinity peptide–HLA class II binding) to elicit an immune response. However, the E296V and M298Q mutant analogs do bind with high affinity to several HLA class II variants. Furthermore, all ADA-positive patients expressed at least one of the HLA class II variants shown to bind a vatreptacog alfa peptide with high affinity [98]. Finally, mutant peptides carrying the E296V and M298Q mutations were detected on HLA class II proteins isolated from dendritic cells. Together, these findings indicate that the E296V and M298Q mutations can result in peptide sequences that act as strong T-cell epitopes in patients with one or more of these "high-risk" alleles [98].

A bioengineered analog of rFVIIa, vatreptacog alfa, has >99% sequence identity to native FVIIa, with only three amino acid substitutions. However, the development of ADAs in treated patients, together with the results of the immunogenicity prediction studies, points to the greatest lesson learned from the vatreptacog alfa case, namely, that even modest sequence changes can create new T-cell epitopes and significantly alter the immunogenicity profile of a therapeutic protein.

10 Immunogenicity Prediction in the Future

As the vatreptacog alfa case shows, considerable challenges may hinder the clinical development of bioengineered therapeutic coagulation proteins, especially when dealing with sequence-modified analogs [1]. One of the greatest challenges lies in developing a protein that offers improved clinical efficacy without introducing an increased immunogenicity risk. As highlighted by the vatreptacog alfa clinical trial program, a primary challenge is the need to include a sufficient number of patients and exposures in phase 2 trials to allow reliable detection of an immunogenic signal [1].

However, a growing area of interest in immunogenicity assessment lies in the development of accurate strategies for predicting the immunogenicity profiles of therapeutic proteins [17]. Animal models are useful for evaluating some of the factors that influence immunogenicity (e.g., product formulation, dosing regimen, and other non-sequence-related affects) [78]; however, as most therapeutic proteins show species differences in amino acid sequence, they are often recognized by animals as foreign [78, 95]. Consequently animal models have restricted predictive value for evaluating mutations in amino acid sequences. The animal and human MHC-II repertoires are not equivalent, as they differ at the amino acid level; therefore, the results of in vivo prediction studies conducted in animals that are not transgenic for HLA class II should be interpreted with great caution [67, 92]. The challenges inherent in using animal models to predict immunogenicity are evident in the vatreptacog alfa case: two studies using rats [56] and mice [98] failed to detect the increased immunogenicity of vatreptacog alfa versus rFVIIa. To overcome some of these challenges, specialized mouse models are being developed [78].

Improved immunogenicity prediction methods and models are clearly needed [57]. Since the early immunogenicity risk assessment for vatreptacog alfa was performed, more than a decade ago, the landscape of immunogenicity prediction has changed considerably, and the tools available today offer a greater level of reliability and prediction accuracy. There was very good concordance between the different immunogenicity prediction tools used for the post hoc assessment of vatreptacog alfa. However, no single tool used alone can accurately predict immunogenicity or address all questions relating to immunogenicity [17, 77]. If the proper sample format is used, T-cell proliferation assays are better predictors of immunogenicity risk than peptide–HLA class II-binding affinities but are more useful in estimating risk in the population as a whole than in individuals. Therefore, strategic combination of multiple predictive approaches is needed. When used in combination, in silico peptide–HLA class II binding predictions, in vitro antigen presentation assays, and ex vivo T-cell proliferation assays provide useful indications of the immunogenicity risk posed by specific neo-sequences in bioengineered therapeutic proteins [67, 77]. Today, a number of drug developers [3] are incorporating this combined approach into preclinical development programs [17, 77]. Such methods are also recommended by both the European Medicines

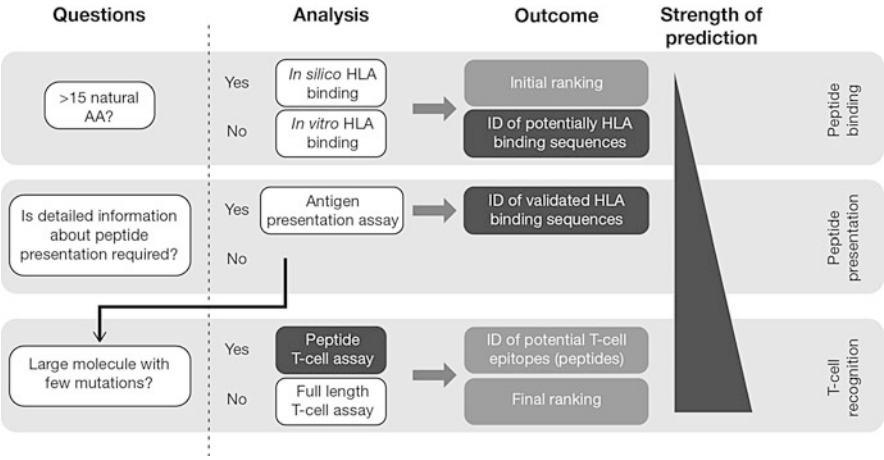


Fig. 10 Suggested pathway for the use of immunogenicity prediction tools. This multistep approach evaluates the immunogenic potential of a therapeutic protein. The successive steps evaluate, in turn, peptide binding, peptide presentation, and T-cell recognition; each step provides a greater predictive strength and thus validates results from the previous step. The first step is in silico prediction of peptide–HLA class II binding, followed by a confirmatory in vitro peptide binding assay, to identify putative HLA-binding peptide sequences ranked in order of potential immunogenicity. The second step is an antigen presentation assay, which evaluates antigen processing and peptide presentation by HLA class II molecules on the surface of antigen-presenting (e.g., dendritic) cells. This step validates the putative HLA class II-binding sequences identified in the first step. The third step is a T-cell proliferation assay, which identifies potential T-cell epitopes, assesses T-cell response, and establishes a final ranking of peptides according to immunogenic potential

Agency (EMA) and the United States (US) Food and Drug Administration (FDA) in their current guidelines for preclinical immunogenicity risk assessment [95, 96]. Indeed, the EMA guidelines now indicate a move away from preclinical animal studies toward combining these non in vivo tools in strategies to perform early immunogenicity risk assessment [95].

A suggested early immunogenicity risk assessment strategy – based on lessons learned from the vatreptacog alfa case – is presented in Fig. 10. If the sequence of interest consists of 15 or more natural amino acids, it is possible to perform an in silico analysis to identify HLA-binding peptides. If the sequence is less than 15 amino acids long or includes unnatural amino acids, the in silico analysis cannot be used, so the in vitro peptide binding assay should be performed instead. The next suggested step is to identify which peptides are presented on the dendritic cell surface by using the DC presentation assay. There are then two sample format options to investigate whether or not there is a T-cell repertoire toward the peptide–HLA complex. If the molecule of interest is large with few mutations (i.e., low degree of foreignness), peptides may be used as samples, designed using information from the in silico or in vitro DC presentation assays. Alternatively, small molecules with more mutations (i.e., high degree of foreignness) may require a format that assays the full length of the molecule.

It is hoped that guidelines for standardizing immunogenicity prediction testing will emerge from the increasing and continued implementation of these novel tools in the early stages of drug discovery and development [17]. Unquestionably, the overriding short-term aim of all these endeavors is to reduce immunogenicity risk in the clinic. A longer-term goal is to improve the prediction of clinical immunogenicity. This will require validation of the prediction tools but will in turn enable more efficient development of bioengineered protein drugs that carry great efficacy potential. This validation will be based on clinical data and the use of the tools to test molecules that are known to be immunogenic. An even longer-term goal is to address the question of what will be required to establish immunogenicity prediction tools as mandatory aspects of the development process for biopharmaceuticals. If these goals can all be met with satisfaction, then the development of "personalized" pharmaceuticals to target patients with low immunogenicity risk becomes a feasible – though still distant – possibility.

References

1. Lentz SR, Ehrenforth S, Karim FA, Matsushita T, Weldingh KN, Windyga J, Mahlangu JN, and adeptTM2 Investigators (2014) Recombinant factor VIIa analog in the management of hemophilia with inhibitors: results from a multicenter, randomized, controlled trial of vatreptacog alfa. J Thromb Haemost 12(8):1244–1253
2. de Paula EV, Kavakli K, Mahlangu J, Ayob Y, Lentz SR, Morfini M, Nemes L, Šalek SZ, Shima M, Windyga J, Ehrenforth S, Chuansumrit A, and 1804 (adept™1) Investigators (2012) Recombinant factor VIIa analog (vatreptacog alfa [activated]) for treatment of joint bleeds in hemophilia patients with inhibitors: a randomized controlled trial. J Thromb Haemost 10 (1):81–89
3. Persson E, Olsen OH, Bjørn SE, Ezban M (2012) Vatreptacog alfa from conception to clinical proof of concept. Semin Thromb Hemost 38(3):274–281
4. Mahlangu JN, Coetzee MJ, Laffan M, Windyga J, Yee TT, Schroeder J, Haaning J, Siegel JE, Lemm G (2012) Phase I, randomized, double-blind, placebo-controlled, single-dose escalation study of the recombinant factor VIIa variant BAY 86–6150 in hemophilia. J Thromb Haemost 10(5):773–780
5. Pipe S (2009) Antihemophilic factor (recombinant) plasma/albumin-free method for the management and prevention of bleeding episodes in patients with hemophilia A. Biologics 3:117–125
6. Shibeko AM, Woodle SA, Mahmood I, Jain N, Ovanesov MV (2014) Predicting dosing advantages of factor VIIa variants with altered tissue factor-dependent and lipid-dependent activities. J Thromb Haemost 12(8):1302–1312
7. Escobar MA (2013) Advances in the treatment of inherited coagulation disorders. Haemophilia 19(5):648–659
8. Kaufman RJ, Powell JS (2013) Molecular approaches for improved clotting factors for hemophilia. Blood 122(22):3568–3574
9. Mahlangu JN, Koh PL, Ng HJ, Lissitchkov T, Hardtke M, Schroeder J (2013) The TRUST trial: anti-drug antibody formation in a patient with hemophilia with inhibitors after receiving the activated factor VII product Bay 86–6150. Presented at 55th annual meeting of the american society of hematology. Abstract 573, December 7, 13
10. Mannucci PM, Tuddenham EG (2001) The hemophilias-from royal genes to gene therapy. N Engl J Med 344(23):1773–1779

11. Bolton-Maggs PH, Pasi KJ (2003) Haemophilias A and B. Lancet 361(9371):1801–1809
12. Coppola A, Di CM, Di Minno MN, Di Palo M, Marrone E, Ieranñ P, Arturo C, Tufano A, Cerbone AM (2010) Treatment of hemophilia: a review of current advances and ongoing issues. J Blood Med 1:183–195
13. Srivastava A, Brewer AK, Mauser-Bunschoten EP, Key NS, Kitchen S, Llinas A, Ludlam CA, Mahlangu JN, Mulder K, Poon MC, Street A, and Treatment Guidelines Working Group on Behalf of the World Federation of Hemophilia (2013) Guidelines for the management of hemophilia. Haemophilia 19(1):e1–e47
14. Franchini M, Mannucci PM (2014) The history of hemophilia. Semin Thromb Hemost 40 (5):571–576
15. Collins PW, Hirsch S, Baglin TP, Dolan G, Hanley J, Makris M, Keeling DM, Liesner R, Brown SA, Hay CR (2007) Acquired hemophilia A in the United Kingdom: a 2-year national surveillance study by the United Kingdom Haemophilia Centre Doctors' Organisation. Blood 109(5):1870–1877
16. Franchini M, Gandini G, Di PT, Mariani G (2005) Acquired hemophilia A: a concise review. Am J Hematol 80(1):55–63
17. Jawa V, Cousens LP, Awwad M, Wakshull E, Kropshofer H, De Groot AS (2013) T-cell dependent immunogenicity of protein therapeutics: preclinical assessment and mitigation. Clin Immunol 149(3):534–555
18. Scott DW (2014) Inhibitors – cellular aspects and novel approaches for tolerance. Haemophilia 20(Suppl 4):80–86
19. Pratt KP, Thompson AR (2009) B-cell and T-cell epitopes in anti-factor VIII immune responses. Clin Rev Allergy Immunol 37(2):80–95
20. Weber CA, Mehta PJ, Ardito M, Moise L, Martin B, De Groot AS (2009) T cell epitope: friend or foe? Immunogenicity of biologics in context. Adv Drug Deliv Rev 61(11):965–976
21. Minno GD, Santagostino E, Pratt K, Königs C (2014) New predictive approaches for ITI treatment. Haemophilia 20(Suppl 6):27–43
22. De Groot AS, Terry F, Cousens L, Martin W (2013) Beyond humanization and de-immunization: tolerization as a method for reducing the immunogenicity of biologics. Expert Rev Clin Pharmacol 6(6):651–662
23. Bardi E, Astermark J (2015) Genetic risk factors for inhibitors in haemophilia A. Eur J Haematol 94(Suppl 77):7–10
24. Saini S, Hamasaki-Katagiri N, Pandey GS, Yanover C, Guelcher C, Simhadri VL, Dandekar S, Guerrera MF, Kimchi-Sarfaty C, Sauna ZE (2015) Genetic determinants of immunogenicity to factor IX during the treatment of haemophilia B. Haemophilia 21(2):210–218
25. Kelley M, Ahene AB, Gorovits B, Kamerud J, King LE, McIntosh T, Yang J (2013) Theoretical considerations and practical approaches to address the effect of anti-drug antibody (ADA) on quantification of biotherapeutics in circulation. AAPS J 15(3):646–658
26. Creeke PI, Farrell RA (2013) Clinical testing for neutralizing antibodies to interferon-beta in multiple sclerosis. Ther Adv Neurol Disord 6(1):3–17
27. Whelan SF, Hofbauer CJ, Horling FM, Allacher P, Wolfsegger MJ, Oldenburg J, Male C, Windyga J, Tiede A, Schwarz HP, Scheiflinger F, Reipert BM (2013) Distinct characteristics of antibody responses against factor VIII in healthy individuals and in different cohorts of hemophilia A patients. Blood 121(6):1039–1048
28. Pandey GS, Yanover C, Howard TE, Sauna ZE (2013) Polymorphisms in the F8 gene and MHC-II variants as risk factors for the development of inhibitory anti-factor VIII antibodies during the treatment of hemophilia a: a computational assessment. PLoS Comput Biol 9(5), e1003066
29. Morfini M, Haya S, Tagariello G, Pollmann H, Quintana M, Siegmund B, Stieltjes N, Dolan G, Tusell J (2007) European study on orthopaedic status of haemophilia patients with inhibitors. Haemophilia 13(5):606–612
30. Wight J, Paisley S (2003) The epidemiology of inhibitors in haemophilia A: a systematic review. Haemophilia 9(4):418–435

31. Hedner U, Lee CA (2011) First 20 years with recombinant FVIIa (NovoSeven). Haemophilia 17(1):e172–e182
32. Lacroix-Desmazes S, Navarrete AM, André S, Bayry J, Kaveri SV, Dasgupta S (2008) Dynamics of factor VIII interactions determine its immunologic fate in hemophilia A. Blood 112(2):240–249
33. van den Hoorn T, Paul P, Jongsma ML, Neefjes J (2011) Routes to manipulate MHC class II antigen presentation. Curr Opin Immunol 23(1):88–95
34. Barbosa MD, Vielmetter J, Chu S, Smith DD, Jacinto J (2006) Clinical link between MHC class II haplotype and interferon-beta (IFN-beta) immunogenicity. Clin Immunol 118 (1):42–50
35. Kempton CL, Meeks SL (2014) Toward optimal therapy for inhibitors in hemophilia. Blood 124(23):3365–3372
36. Hoffman M, Dargaud Y (2012) Mechanisms and monitoring of bypassing agent therapy. J Thromb Haemost 10(8):1478–1485
37. Abshire T, Kenet G (2008) Safety update on the use of recombinant factor VIIa and the treatment of congenital and acquired deficiency of factor VIII or IX with inhibitors. Haemophilia 14(5):898–902
38. Key NS, Aledort LM, Beardsley D, Cooper HA, Davignon G, Ewenstein BM, Gilchrist GS, Gill JC, Glader B, Hoots WK, Kisker CT, Lusher JM, Rosenfield CG, Shapiro AD, Smith H, Taft E (1998) Home treatment of mild to moderate bleeding episodes using recombinant factor VIIa (NovoSeven) in haemophiliacs with inhibitors. Thromb Haemost 80(6):912–918
39. Kavakli K, Makris M, Zulfikar B, Erhardtsen E, Abrams ZS, Kenet G, and NovoSeven trial (F7HAEM-1510) investigators (2006) Home treatment of haemarthroses using a single dose regimen of recombinant activated factor VII in patients with haemophilia and inhibitors. A multi-centre, randomised, double-blind, cross-over trial. Thromb Haemost 95(4):600–605
40. Santagostino E, Mancuso ME, Rocino A, Mancuso G, Scaraggi F, Mannucci PM (2006) A prospective randomized trial of high and standard dosages of recombinant factor VIIa for treatment of hemarthroses in hemophiliacs with inhibitors. J Thromb Haemost 4(2):367–371
41. Young G, Shafer FE, Rojas P, Seremetis S (2008) Single 270 microg kg(−1)-dose rFVIIa vs. standard 90 microg kg(−1)-dose rFVIIa and APCC for home treatment of joint bleeds in haemophilia patients with inhibitors: a randomized comparison. Haemophilia 14(2):287–294
42. Parameswaran R, Shapiro AD, Gill JC, Kessler CM, and HTRS Registry Investigators (2005) Dose effect and efficacy of rFVIIa in the treatment of haemophilia patients with inhibitors: analysis from the Hemophilia and Thrombosis Research Society Registry. Haemophilia 11 (2):100–106
43. Neufeld EJ, Négrier C, Arkhammar P, Benchikh el Fegoun S, Simonsen MD, Rosholm A, Seremetis S (2015) Safety update on the use of recombinant activated factor VII in approved indications. Blood Rev 29(Suppl 1):S34–S41
44. Persson E (2004) Variants of recombinant factor VIIa with increased intrinsic activity. Semin Hematol 41(1 Suppl 1):89–92
45. Møss J, Scharling B, Ezban M, Møller ST (2009) Evaluation of the safety and pharmacokinetics of a fast-acting recombinant FVIIa analogue, NN1731, in healthy male subjects. J Thromb Haemost 7(2):299–305
46. Allen GA, Persson E, Campbell RA, Ezban M, Hedner U, Wolberg AS (2007) A variant of recombinant factor VIIa with enhanced procoagulant and antifibrinolytic activities in an in vitro model of hemophilia. Arterioscler Thromb Vasc Biol 27(3):683–689
47. Aljamali MN, Kjalke M, Hedner U, Ezban M, Tranholm M (2009) Thrombin generation and platelet activation induced by rFVIIa (NovoSeven) and NN1731 in a reconstituted cell-based model mimicking haemophilia conditions. Haemophilia 15(6):1318–1326
48. Ghosh S, Ezban M, Persson E, Pendurthi U, Hedner U, Rao LV (2007) Activity and regulation of factor VIIa analogs with increased potency at the endothelial cell surface. J Thromb Haemost 5(2):336–346

49. Brophy DF, Martin EJ, Nolte ME, Kuhn JG, Carr ME Jr (2007) Effect of recombinant factor VIIa variant (NN1731) on platelet function, clot structure and force onset time in whole blood from healthy volunteers and haemophilia patients. Haemophilia 13(5):533–541
50. Brophy DF, Martin EJ, Nolte ME, Kuhn JG, Barrett JC, Ezban M (2010) Factor VIIa analog has marked effects on platelet function and clot kinetics in blood from patients with hemophilia A. Blood Coagul Fibrinolysis 21(6):539–546
51. Sørensen B, Persson E, Ingerslev J (2007) Factor VIIa analogue (V158D/E296V/M298Q-FVIIa) normalises clot formation in whole blood from patients with severe haemophilia A. Br J Haematol 137(2):158–165
52. Gray LD, Hussey MA, Larson BM, Machlus KR, Campbell RA, Koch G, Ezban M, Hedner U, Wolberg AS (2011) Recombinant factor VIIa analog NN1731 (V158D/E296V/M298Q-FVIIa) enhances fibrin formation, structure and stability in lipidated hemophilic plasma. Thromb Res 128(6):570–576
53. Tranholm M, Kristensen K, Kristensen AT, Pyke C, Røjkjaer R, Persson E (2003) Improved hemostasis with superactive analogs of factor VIIa in a mouse model of hemophilia A. Blood 102(10):3615–3620
54. Holmberg HL, Lauritzen B, Tranholm M, Ezban M (2009) Faster onset of effect and greater efficacy of NN1731 compared with rFVIIa, aPCC and FVIII in tail bleeding in hemophilic mice. J Thromb Haemost 7(9):1517–1522
55. Persson E, Kjalke M, Olsen OH (2001) Rational design of coagulation factor VIIa variants with substantially increased intrinsic activity. Proc Natl Acad Sci U S A 98(24):13583–13588
56. Sommer C, Norbert JP, Salanti Z, Clausen JT, Jensen LB (2007) Immunogenicity of novel recombinant human activated factor VII analogues on factor VII neonatally-tolerized rats. Thromb Haemost 98(4):721–725
57. Mahlangu JN, Weldingh KN, Lentz SR, Kaicker S, Karim FA, Matsushita T, Recht M, Tomczak W, Windyga J, Ehrenforth S, Knobe K, adept™2 Investigators (2015) Changes in the amino acid sequence of the rFVIIa analog, vatreptacog alfa, are associated with clinical immunogenicity. J Thromb Haemost 13(11):1989–1998
58. Kobayashi H, Wood M, Song Y, Appella E, Celis E (2000) Defining promiscuous MHC class II helper T-cell epitopes for the HER2/neu tumor antigen. Cancer Res 60(18):5228–5236
59. Nielsen M, Lund O, Buus S, Lundegaard C (2010) MHC class II epitope predictive algorithms. Immunology 130(3):319–328
60. Lazarski CA, Chaves FA, Jenks SA, Wu S, Richards KA, Weaver JM, Sant AJ (2005) The kinetic stability of MHC class II:peptide complexes is a key parameter that dictates immunodominance. Immunity 23(1):29–40
61. Chen X, Hickling TP, Vicini P (2014) A mechanistic, multiscale mathematical model of immunogenicity for therapeutic proteins: part 2-model applications. CPT Pharmacometrics Syst Pharmacol 3, e134
62. Ekins S, Mestres J, Testa B (2007) In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. Br J Pharmacol 152(1):9–20
63. Lundegaard C, Lund O, Nielsen M (2011) Prediction of epitopes using neural network based methods. J Immunol Methods 374(1–2):26–34
64. Bond KB, Sriwanthana B, Hodge TW, De Groot AS, Mastro TD, Young NL, Promadej N, Altman JD, Limpakarnjanarat K, McNicholl JM (2001) An HLA-directed molecular and bioinformatics approach identifies new HLA-A11 HIV-1 subtype E cytotoxic T lymphocyte epitopes in HIV-1-infected Thais. AIDS Res Hum Retroviruses 17(8):703–717
65. Brusic V, Bajic VB, Petrovsky N (2004) Computational methods for prediction of T-cell epitopes–a framework for modelling, testing, and applications. Methods 34(4):436–443
66. Cohen T, Moise L, Ardito M, Martin W, De Groot AS (2010) A method for individualizing the prediction of immunogenicity of protein vaccines and biologic therapeutics: individualized T cell epitope measure (iTEM). J Biomed Biotechnol. pii: 961752. doi:10.1155/2010/961752
67. De Groot AS, Moise L (2007) Prediction of immunogenicity for therapeutic proteins: state of the art. Curr Opin Drug Discov Devel 10(3):332–340

68. De Groot AS, Bosma A, Chinai N, Frost J, Jesdale BM, Gonzalez MA, Martin W, Saint-Aubin C (2001) From genome to vaccine: in silico predictions, ex vivo verification. Vaccine 19 (31):4385–4395
69. De Groot AS, Knopp PM, Martin W (2005) De-immunization of therapeutic proteins by T-cell epitope modification. Dev Biol (Basel) 122:171–194
70. Koren E, Zuckerman LA, Mire-Sluis AR (2002) Immune responses to therapeutic proteins in humans–clinical significance, assessment and prediction. Curr Pharm Biotechnol 3 (4):349–360
71. Koren E, De Groot AS, Jawa V, Beck KD, Boone T, Rivera D, Li L, Mytych D, Koscec M, Weeraratne D, Swanson S, Martin W (2007) Clinical validation of the "in silico" prediction of immunogenicity of a human recombinant therapeutic protein. Clin Immunol 124(1):26–32
72. Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, and Nielsen M (2008) NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8–11. Nucleic Acids Res 36(Web Server issue):W509–W512
73. McMurry J, Sbai H, Gennaro ML, Carter EJ, Martin W, De Groot AS (2005) Analyzing Mycobacterium tuberculosis proteomes for candidate vaccine epitopes. Tuberculosis (Edinb) 85(1–2):95–105
74. Tatarewicz SM, Wei X, Gupta S, Masterman D, Swanson SJ, Moxness MS (2007) Development of a maturing T-cell-mediated immune response in patients with idiopathic Parkinson's disease receiving r-metHuGDNF via continuous intraputaminal infusion. J Clin Immunol 27 (6):620–627
75. Inaba H, Martin W, De Groot AS, Qin S, De Groot LJ (2006) Thyrotropin receptor epitopes and their relation to histocompatibility leukocyte antigen-DR molecules in Graves' disease. J Clin Endocrinol Metab 91(6):2286–2294
76. Brennan FR, Morton LD, Spindeldreher S, Kiessling A, Allenspach R, Hey A, Muller PY, Frings W, Sims J (2010) Safety and immunotoxicity assessment of immunomodulatory monoclonal antibodies. MAbs 2(3):233–255
77. Holgate RG, Baker MP (2009) Circumventing immunogenicity in the development of therapeutic antibodies. IDrugs 12(4):233–237
78. Perry LC, Jones TD, Baker MP (2008) New approaches to prediction of immune responses to therapeutic proteins during preclinical development. Drugs R D 9(6):385–396
79. Nielsen M, Justesen S, Lund O, Lundegaard C, and Buus S (2010) NetMHCIIpan-2.0 – Improved pan-specific HLA-DR predictions using a novel concurrent alignment and weight optimization training procedure. Immunome Res 6:9
80. Nielsen M and Lund O (2009) NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. BMC Bioinformatics 10:296
81. Justesen S, Harndahl M, Lamberth K, Nielsen LL, Buus S (2009) Functional recombinant MHC class II molecules and high-throughput peptide-binding assays. Immunome Res 5:2
82. Ullman EF, Kirakossian H, Switchenko AC, Ishkanian J, Ericson M, Wartchow CA, Pirio M, Pease J, Irvin BR, Singh S, Singh R, Patel R, Dafforn A, Davalian D, Skold C, Kurn N, Wagner DB (1996) Luminescent oxygen channeling assay (LOCI): sensitive, broadly applicable homogeneous immunoassay method. Clin Chem 42(9):1518–1526
83. Eglen RM, Reisine T, Roby P, Rouleau N, Illy C, Bossé R, Bielefeld M (2008) The use of AlphaScreen technology in HTS: current status. Curr Chem Genomics 1:2–10
84. Rammensee HG (1995) Chemistry of peptides associated with MHC class I and class II molecules. Curr Opin Immunol 7(1):85–96
85. Chicz RM, Urban RG, Lane WS, Gorga JC, Stern LJ, Vignali DA, Strominger JL (1992) Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. Nature 358(6389):764–768
86. Sung SS (2008) Monocyte-derived dendritic cells as antigen-presenting cells in T-cell proliferation and cytokine production. Methods Mol Med 138:97–106
87. Chung CY, Ysebaert D, Berneman ZN, Cools N (2013) Dendritic cells: cellular mediators for immunological tolerance. Clin Dev Immunol 2013:972865

88. Dudek AM, Martin S, Garg AD, Agostinis P (2013) Immature, semi-mature, and fully mature dendritic cells: toward a dc-cancer cells interface that augments anticancer immunity. Front Immunol 4:438
89. Roche PA, Furuta K (2015) The ins and outs of MHC class II-mediated antigen processing and presentation. Nat Rev Immunol 15(4):203–216
90. Steinman RM, Hemmi H (2006) Dendritic cells: translating innate to adaptive immunity. Curr Top Microbiol Immunol 311:17–58
91. Gordon JR, Ma Y, Churchman L, Gordon SA, Dawicki W (2014) Regulatory dendritic cells for immunotherapy in immunologic diseases. Front Immunol 5:7
92. De Groot AS, McMurry J, Moise L (2008) Prediction of immunogenicity: in silico paradigms, ex vivo and in vivo correlates. Curr Opin Pharmacol 8(5):620–626
93. Jaber A, Baker M (2007) Assessment of the immunogenicity of different interferon beta-1a formulations using ex vivo T-cell assays. J Pharm Biomed Anal 43(4):1256–1261
94. Smith RC, O'Bryan LM, Mitchell PJ, Leung D, Ghanem M, Wilson JM, Hanson JC, Sossick S, Cooper J, Huang L, Merchant KM, Lu J, O'Neill MJ (2015) Increased brain bio-distribution and chemical stability and decreased immunogenicity of an engineered variant of GDNF. Exp Neurol 267:165–176
95. European Medicines Agency and Committee for Medicinal Products for Human Use (CHMP) (2015) Guideline on immunogenicity assessment of biotechnology-derived therapeutic products. http://www.ema.europa.eu/ema/doc_index.jsp?curl=pages/includes/document/document_detail.jsp?webContentId=WC500194507&murl=menus/document_library/document_library.jsp&mid=0b01ac058009a3dc
96. U.S. Department of Health and Human Services: Food and Drug Administration (2009) Guidance for industry: assay development for immunogenicity testing of therapeutic proteins. http://www.fda.gov/downloads/Drugs/. . ./Guidances/UCM192750.pdf
97. Hoffman M, Volovyk Z, Persson E, Gabriel DA, Ezban M, Monroe DM (2011) Platelet binding and activity of a factor VIIa variant with enhanced tissue factor independent activity. J Thromb Haemost 9(4):759–766
98. Lamberth K, Reedtz-Runge SL, Simon J, Klementyeva K, Pandey GS, Padkjær SB, Pascal V, León IR, Gudme CN, Buus S, Sauna ZE (2017) Post hoc assessment of the immunogenicity of bioengineered factor VIIa demonstrates the use of preclinical tools. Sci Transl Med 9(372). pii: eaag1286. doi:10.1126/scitranslmed.aag1286

Top Med Chem (2017) 21: 161–178
DOI: 10.1007/7355_2016_2

Published online: 15 March 2016

The Art of Gene Redesign and Recombinant Protein Production: Approaches and Perspectives

Anton A. Komar

Abstract In recent years, the demand for recombinant proteins for use in research laboratories or in medical settings has increased dramatically. Although a wide variety of recombinant protein expression systems and gene redesign approaches are available, obtaining active, correctly folded recombinant proteins in sufficient amounts remains a challenge in many cases. One of the main approaches to gene redesign with the potential to increase protein production involves introduction of synonymous codon substitutions in mRNAs aimed at increasing the rate/efficiency of translation. However, a number of recent studies have shown that synonymous codon substitutions can also negatively impact mRNA biogenesis, mRNA decoding, as well as protein folding and function. Maximizing the speed and output of translation may put conflicting demands on the protein synthesis machinery resulting in reduced accuracy of the decoding process and/or improper protein folding. An improved understanding of the impact of synonymous codon substitutions on mRNA/protein biogenesis and function is critically important for the development of safer and more effective recombinant protein therapeutics. This review discusses the most common approaches to gene redesign that involve synonymous codon substitutions and provides recommendations for their optimal use in light of recent developments in the field regarding the impact of synonymous codon usage on various aspects of protein production and function.

Keywords Codon usage, Gene redesign, Mistranslation, mRNA turnover, Protein folding, Protein synthesis, Rare synonymous codons, Recombinant protein therapeutics, Synonymous codons

A.A. Komar (✉)
Center for Gene Regulation in Health and Disease and Department of Biological, Geological and Environmental Sciences, Cleveland State University, Cleveland, OH 44115, USA

DAPCEL, Inc., Cleveland, OH 44106, USA
e-mail: a.komar@csuohio.edu

Contents

1 Introduction

Production of soluble and functionally active proteins in heterologous and homologous host organisms is the cornerstone of many modern biotechnology applications. In recent years, the demand for recombinant proteins used in research laboratories or in medical settings (e.g., for therapeutic applications) has increased dramatically. Specifically, the protein therapeutic market was valued in excess of $85 billion in 2010 and is predicted to double by the end of 2018, reaching up to $165 billion, as new products (especially therapeutic monoclonal antibodies) become available (http://www.researchandmarkets.com/reports/2729030/global_protein_therapeutics_market_outlook_2018). Despite the strong existing and potential significance of efficient recombinant protein production for both research applications and development of novel therapeutics, obtaining soluble, active recombinant proteins in sufficient amounts remains challenging in many cases.

A wide variety of recombinant protein expression systems are well established. These include, but are not limited to, various cellular systems, such as bacterial, yeast, insect and mammalian systems [1–7], and cell-free in vitro systems [8, 9]. The urgent need for robust and highly scalable protein manufacturing systems has further led to the development of in vivo plant- and animal-based systems [10–13]. All of these systems have their own advantages and disadvantages [14]. The choice of system to use for a particular application depends on the specific requirements for the final recombinant protein product (e.g., requirements for proper protein processing and/or co- and posttranslational protein folding and modifications) [14]. In most cases, use of a recombinant protein expression system that closely resembles the protein's natural in vivo expression system/environment is highly desirable, but this is obviously not always achievable [14]. For example, toxicity of the final product may not allow enhanced expression of a protein in a homologous, or even heterologous, cellular system(s) [15, 16]. In such cases, cell-free protein synthesis systems on a larger scale, particularly with continuous action, may offer an alternative solution [8, 9, 15, 17, 18]. In addition, expression of

unmodified natural genes in a homologous environment frequently does not support levels of protein expression sufficient for large-scale protein production. The key to solving this problem lies in development of gene redesign approaches that result in robust expression of functionally active proteins both inside and outside their natural (homologous) cellular environments.

One of the main approaches to gene redesign facilitating protein production in heterologous and homologous organisms [19–21] takes advantage of the degeneracy of the genetic code (meaning a given amino acid may be encoded by more than one "synonymous" codon). Synonymous codons are present at different frequencies in different organisms and are decoded at different rates [22–24]. Therefore, substitution of synonymous codons in a gene can dramatically affect the rate/efficiency of synthesis of the encoded protein without altering its amino acid sequence [19–21]. In a given organism, frequently used codons are typically translated more rapidly than infrequently used ones due to the fact that tRNAs corresponding to the frequently used codons are relatively more abundant [25–31]. Many synonymous codons that are frequently used in eukaryotes (especially mammals) are utilized with low frequency in prokaryotes [22–24] such as the bacteria *Escherichia coli*, one of the most common hosts for heterologous protein production [14]. The impact of these differences on recombinant protein production is now well appreciated, and it has been clearly demonstrated that the level of protein expression in heterologous and homologous organisms can be increased through suitable selection of synonymous (frequent) codons along target mRNAs [19–21].

In addition to the effects of differential codon usage, the secondary structure of messenger RNAs (mRNA) has been recognized as a factor that can have a negative impact on translation and reduce protein yields by slowing or blocking translation initiation and/or the movement of ribosomes along the mRNA [32–39].

Several other considerations important for recombinant protein production (e.g., choice of appropriate vector/promoter system(s), means of gene delivery, etc.) are outside the scope of this short review.

Approaches involving substitution of the majority of infrequently used codons with synonymous frequently used ones, often combined with elimination of extreme GC content that could contribute to formation of stable mRNA secondary structures, have been widely used by many biotechnology companies and research groups for optimization of heterologous gene/protein expression, but with mixed results ([19, 40] and references therein). Use of gene sequences optimized through the abovementioned approaches often yielded large amounts of recombinant proteins [19, 40]; however, in many cases, the products formed biologically inactive insoluble aggregates which had to be refolded (whenever it was possible) in order to regain similarity in structure and biological activity with native analogues [19, 28]. Moreover, even when proteins expressed in heterologous or homologous hosts remained soluble, they were not necessarily natively folded [41].

These and other experiments brought about awareness of the scientific community to the impact of synonymous codon usage on not only the efficiency of translation but also on other aspects of gene function, particularly, protein folding. The significance of synonymous codon usage on protein folding was highlighted by

findings showing that multiple and, more surprisingly, single synonymous substitutions/mutations can affect proteins' activity [42–44], interactions with drugs and inhibitors [43], phosphorylation profiles [45], sensitivity to limited proteolysis [43, 45, 46], spectroscopic properties [47], and aggregation propensity [47–49] and ultimately change protein structure [50].

Many recent studies have shown that synonymous substitutions or naturally occurring synonymous mutations are not neutral and may affect gene function by multiple mechanisms [51, 52], including but not limited to those mentioned above, as well as mechanisms exerting effects on mRNA splicing and/or mRNA stability [53, 54]. Synonymous codon choice has been also suggested to affect efficient interaction of nascent polypeptides with the signal recognition particle [55]. Changes in codon context caused by synonymous mutations may also induce mistranslation leading to protein misfolding [56].

While in many instances complete understanding of the exact effects caused by synonymous substitutions and/or mutations is still lacking, it nevertheless seems possible to use existing knowledge for the development of some common rules to gene design and redesign that should increase the chances of getting the desired levels and activity of the expressed recombinant proteins and reduce protein misfolding and aggregation.

This review discusses the most common approaches to gene redesign that involve synonymous codon substitutions and contains a set of recommendations for optimizing protein synthesis and folding through this approach. These recommendations take into account recent developments in the field highlighting the impact of synonymous codon usage on protein production and function.

2 Synonymous Gene Exploration in Protein Production and Folding

Designing an optimal gene for recombinant protein production requires choosing from an enormous number of possible DNA/RNA sequences. It is a combinatorial problem, giving approximately 3^N variants for a sequence with N codons. However, as discussed below, this number can be substantially reduced by taking into account a set of critical considerations.

In general, two global gene design/redesign approaches predominate (1) de novo gene design based on reverse translation from an amino acid sequence to DNA/RNA and (2) gene redesign based on recoding of a natural DNA/RNA sequence. Numerous online/web-based and stand-alone platforms are available for use in one or both of these approaches. These include, for example, Codon Optimization OnLine (COOL) [57], DNA Works [58], D-Tailor [59], EuGene [60], GeneDesign [61], Gene Designer 2.0 [62], Jcat [63], mRNA Optimiser [64], OPTIMIZER [65], Synthetic Gene Designer [66], TmPrime [67], Visual Gene Developer [68], and others (for a review see [69]). The majority of available

tools, however, start with a natural DNA/RNA sequence and employ either codon or RNA structure optimization algorithms (or both) to maximize gene expression; only TmPrime [67] is a "pure" de novo back-translation tool. GeneDesign [61] and OPTIMIZER [65] offer both possibilities – de novo back-translation from protein to DNA/RNA sequence and recoding of the natural DNA/RNA sequence.

Most of the abovementioned platforms customize codon usage by setting codon frequency percentage [70] and/or Codon Adaptation Index (CAI) [71] thresholds and then substituting rare synonymous codons with frequent ones along the entire open reading frame (ORF) of a gene to achieve the desired threshold level(s). Substitutions are selected based on known organism-specific codon biases [22–24, 68]. The COOL [57], D-Tailor [59], EuGene [60], OPTIMIZER [65], and Visual Gene Developer [68] tools also take into account the RNA structure and/or GC/AT content, aiming to reduce obstacles related to formation of stable RNA structures. mRNA Optimizer [64] and TmPrime [65] focus solely on mRNA secondary structure optimization to avoid stable secondary structures by means of maximizing the minimum free energy (MFE) of the nucleotide sequences without changing the final resulting amino acid sequence.

As mentioned above, all currently available algorithms (with the exception of TmPrime [65]) typically start from the original/natural coding sequence and then evolve the sequence through iterations of synonymous codon changes that would increase/maximize the MFE and/or codon usage frequency/CAI or both to achieve the desired outcome. However, none of the abovementioned tools typically considers the impact of synonymous codon usage on protein folding (rather than simply on translation efficiency). They also fail to take into account some other important considerations that can affect mRNA translatability and stability and, therefore, preclude efficient expression of correctly folded and functional proteins. Below, I examine some of these considerations that may facilitate gene design and redesign toward optimized expression of active, correctly folded proteins.

2.1 Codon Usage at ORF (Open Reading Frame) 5′ Termini

The occurrence of synonymous codons in protein-coding open reading frames (ORFs) of genes is not random, thus revealing the existence of evolutionary pressure on codon choice [23, 24, 28, 72–74]. Clustering of synonymous codons has been observed at specific conserved locations in mRNAs indicating that there are forces that influence the selection of these codons at specific locations within mRNA sequences [28, 33, 37, 38, 55, 75, 76]. Strategic placement of specific synonymous codons, particularly those that are rare, in gene ORFs suggests a functional role conserved in evolution rather than random chance. Therefore, the randomized and/or global substitution of rare synonymous codons with frequent ones that is offered by the majority of tools aimed at simply increasing CAI/codon usage frequency and/or MFE (see above) might not be beneficial for the production of a functional protein.

An example of nonrandom synonymous codon usage within ORFs is the observed enrichment of rare codons at the 5′ termini of genes in *E. coli* and many other prokaryotes, as well as in genes of some eukaryotes such as the yeast *Saccharomyces cerevisiae* [75, 76]. The clustering of rare codons at 5′ gene termini (typically at codon positions 1 to ~20 [33, 37, 38, 76]) clearly indicates an influence of evolutionary pressure on their selection. This particular aspect of natural codon usage may be explained by fact that rare codons in many bacteria are largely AT-rich [70]. Thus, their clustering at 5′ORF termini leads to reduced secondary structure in that region of the mRNA and, consequently, enhanced protein expression (it is known that mRNA secondary structure at 5′ ORF termini negatively affects protein expression by limiting access of the ribosomes to the ribosome binding site (RBS) on the mRNA [33, 37, 38, 55, 75]).

It should be noted, however, that the enrichment of rare codons at 5′ ORF termini has been mostly found in bacteria with genomes with overall GC content of at least 50% [77]. Recent work showed that, in general, AT-rich codons as opposed to rare codons are preferentially located at 5′ ORF termini in prokaryotes [33, 34, 37, 38, 54]. This further implicates secondary structure as the driving force for specific codon selection at 5′ ORF termini in bacteria [33, 38, 54]. Interestingly, the higher the GC content of a genome, the more mRNA stability is reduced at the region near the start codon [78].

It should be also noted that despite differences in translation apparatus and the mechanism of protein synthesis between prokaryotes and eukaryotes, many eukaryotic ORFeomes also are characterized by reduced 5′-terminal mRNA secondary structure near the start codon [78]. This indicates that reduced 5′-terminal ORF mRNA secondary structure may have been evolutionary selected in all organisms. In eukaryotes, this can be expected to facilitate start-codon recognition by the scanning ribosome [78].

Could there be additional reasons for preferential use of rare codons at the 5′ ORF termini of some natural genes, including those in *E. coli*? It has been suggested that clustering of rare codons at 5′ ORF termini may in certain cases allow slow co-translational formation of the N-terminal folding nucleus of the protein, thus facilitating overall correct protein folding in the cell [28].

Interestingly, strong enrichment of rare codons at 5′ gene termini has been preferentially observed (with very high statistical significance ($P < 0.0001$)) in genes/ORFs encoding secretory proteins [76]. It has been suggested that for genes encoding secretory proteins with N-terminal signal sequences, 5′ rare codon clusters could have a functional role related to secretion, by transiently slowing down translation prior to membrane localization of the nascent chain(s) [79]. It has been experimentally shown in yeast that local slowdown of translation caused by presence of rare codons (located ~35–40 codons downstream of signal sequences or transmembrane segments) promotes nascent-chain recognition by signal recognition particle (SRP), which assists in protein translocation across membranes [55]. Similarly, strategically located Shine-Dalgarno-like elements were identified in ORFeomes of *E. coli* secretory proteins; these elements serve to transiently slow

down translation elongation in order to allow efficient integration of the transmembrane helix of many membrane proteins [80].

Therefore, based on the considerations described above, carefully planned placement of rare/non-optimal (or AT-rich) codons in the 5′ ORF termini of mRNAs, especially for those encoding secretory and transmembrane proteins, may represent an important strategy for successful gene design and redesign enhancing proper protein production, secretion, and folding.

2.2 Conserved Rare Codon Clusters Within Gene ORFs

It is widely believed that the major influence of codon usage is on global and local translation rate. As mentioned above, frequently used codons are translated more rapidly than infrequently used ones [25–31]. However, which codons are more rare or frequent varies by organism [22–25, 70]. Surprisingly, across all organisms, rare codons appear to occur in clusters, rather than being randomly scattered across genes [28, 75]. Although there is a general tendency for rare codons to cluster at the 5′ termini of ORFs (see above), such clustering is also observed within ORFs [28, 75, 81]. These clusters are not confined to the 5′ end of ORFs or to ORFs of genes/proteins that are expressed at a low level (as might be expected if rare codons are thought of as simply correlating with reduced translation rate). Rather, they are found to occur equally in genes for all types of proteins, including abundant/highly expressed proteins [75, 81].

Analyses of ORFeomes from prokaryotic and eukaryotic organisms revealed that rare codon clustering (1) is not limited to a particular set of genes or genotype, (2) does not depend on and is not related to the overall GC content of the organism's genome, and (3) is significantly more abundant than would be expected based on random selection [75, 81]. Furthermore, for some protein families, the locations of rare codon-rich regions within mRNAs are highly conserved across homologs in different organisms; this is observed, for example, in families of cytochromes c, globins, gamma-B crystallins [28], ocular lacritins [82], and chloramphenicol acetyltransferases [28, 83].

Enrichment of rare codon clusters at specific locations in a broad range of genes and organisms suggests that evolutionary selection determines such clustering and that it must have some functional significance [28, 75, 81–83]. One hypothesis links the location of rare codon clusters to the process of protein folding in the cell [84, 85]. This proposes that sequential folding events that occur during co-translational folding of proteins might be separated by rare codon clusters, with such clusters serving to reduce the speed of translation at these positions and thereby facilitating proper folding through temporal separation of folding events on the ribosome [28, 74, 86–91]. This is consistent with the finding that there seems to be a certain hierarchy in the location of rare codon-rich regions along mRNAs. Frequently, but not always, the rarest codons seem to encode boundaries of relatively large structural units (e.g., protein domains), whereas less rare codons encode

boundaries of smaller units (e.g., protein motifs and subdomains) [28]. This might reflect the need to provide a more substantial translational delay for independent co-translation folding of larger units in comparison with smaller ones [28].

In summary, while there is a substantial body of literature underlining the overall negative effects of rare codons on levels of protein production (see [19] for a review), it is becoming increasingly clear that strategic placement of conserved rare codons clusters can have positive effects on protein biogenesis (particularly proper folding) and function. Some biotech companies, such as DAPCEL, Inc., are already using this knowledge to enhance protein production and facilitate correct co-translational protein folding.

2.3 Codon Usage at ORF (Open Reading Frame) 3′ Termini

Enrichment of rare codons at the 3′ terminus of *E. coli* ORFs (and ORFs of 11 other prokaryotes) has also been observed [76]. While significant enrichment of rare codons at the 5′ termini of genes in *E. coli* can be explained as a mechanism that facilitates interaction between ribosomes and ribosome binding sites on mRNAs (see above; [33, 37, 38, 55, 75]), the observed incidence (albeit less pronounced) of increased rare codon abundance at the 3′ termini of *E. coli* ORFs is not that easy to explain. It is possible that rare codon clusters at 3′ ORF termini could be required for more robust termination of translation and/or for reducing the rate of protein folding before release from the ribosome [76]. Queuing of ribosomes at the 3′ termini of ORFs due to presence of rare codons may also protect mRNAs from degradation. An improved understanding of the impact of codon usage at 3′ ORF termini is required before this feature can be rationally exploited in gene design and redesign strategies and/or interpretation of in vivo folding pathways.

3 Synonymous Codons and mRNA Stability

mRNA turnover plays a critical role in regulating gene expression. mRNAs with longer half-lives generally produce more protein than those with shorter half-lives simply because they are available to be translated for a longer period of time. A link between codon usage and mRNA turnover rate has been long recognized in both prokaryotes and eukaryotes [92–94], but has not been well understood until recently [53, 54]. Previously, it was generally believed that more thermodynamically stable mRNAs would also be more resistant to degradation. However, recent work showed that, at least in yeast, so-called codon optimality [53] rather than mRNA thermodynamic stability has a broad and powerful influence on in vivo mRNA degradation rates. Codon optimality is a scale that reflects the balance between the supply of specific charged tRNA molecules and the demand for their use by translating ribosomes, thus representing a measure of translation efficiency [53]. Optimal

codons (typically, these are frequent codons) are decoded faster. In the yeast study, it was found that many stable/long-lived mRNAs harbor optimal codons within their ORFs, while many unstable/short-lived mRNAs harbor non-optimal codons [53]. Moreover, it was found that substitution of optimal codons with synonymous, non-optimal codons results in dramatic destabilization of the mRNA and vice versa [53]. Interestingly, very similar results were obtained in *E. coli* [54]. These findings suggest that transcript-specific translation elongation rate is an important determinant of mRNA stability and that more rapidly translated mRNAs (at least in yeast and *E. coli*) are likely to be more stable and, thus, produce more protein. This new information presents an opportunity to upscale protein production in yeast and *E. coli* via reassignment of codon optimality in an mRNA to increase its stability and, thus, its capacity to produce protein. Whether the same paradigm exists in higher eukaryotic organisms remains to be determined. However, this approach should be applied with caution since assignment of codons that are optimal for translation rate and mRNA stability could lead to incorrect protein folding.

4 Synonymous Codons and Mistranslation/Frameshifting

Another aspect of mRNA biology that can be impacted by synonymous codon usage is the accuracy with which they are translated. Clearly, mRNAs must be translated accurately in order for fully functional proteins to be produced. Estimates of missense error rates (referred to as miscoding or mistranslation) during protein synthesis from natural mRNAs vary from 10^{-3} to 10^{-4} per codon ([95–98] and references therein). Mistranslation is the incorporation of an amino acid that is different from the one encoded by a specific codon in the mRNA. Recent research has enhanced our understanding of mistranslation mechanisms and how it is controlled [95–98]. While it is generally believed that synonymous codon changes should be silent (not changing the amino acid that is incorporated), that is not always the case [95–98]. Moreover, certain codons are mistranslated more frequently than others [95, 98]. This is apparently due to the fact that translation speed and mistranslation rate are carefully balanced during protein synthesis and situations maximizing translation speed place demands on the translational machinery that reduces accuracy [95–98]. In general, translation has multiple layers of proofreading; however, most errors occur during decoding, which takes place on the ribosome [96, 98]. The frequency of miscoding of different codons varies over a nearly 20-fold range ([95] and references therein). Mispairing at the wobble position and scarce availability of cognate competitor tRNAs appear to play major roles in mistranslation [95–98]. For example, the frequency of miscoding of the AAU (Asn) codon in *E. coli* leading to incorporation of Lys (encoded by AAG and AAA) instead of Asn is about fourfold higher than that for the AAC (Asn) codon [95]. It should be noted, however, that the AAU codon is used more frequently than the AAC codon (codon usage frequency per 1,000 codons is 29.32 for AAU vs. 20.26 for AAC [70]); thus, substitution of AAC with AAU

with the intention of maximizing codon frequency/CAI could result in increased levels of miscoding, which in turn could lead to loss of protein activity due to misfolding [56] or absence of a functionally important amino acid.

While, as described above, there is considerable evidence linking codon usage and missense errors, little is known about the relationship between codon usage and frameshifting errors. Programmed ribosomal frameshifting is utilized by many viruses and bacteria to increase the information content of their genomes; through frameshifting, multiple proteins can be produced from a single span of sequence [99, 100]. Signals in mRNAs have been identified that cause frameshifting by one base in the 5′ (−1) or 3′ (+1) direction [99, 100]. While beneficial in some cases for bacteria and viruses as mentioned above, unintended frameshifting during translation is clearly not desirable. Frameshifting errors can lead to premature termination of translation or generate abnormal proteins with toxic effects on the cell [56]. Attempts have been made to develop computational tools to assess whether codon usage can be optimized to minimize the frequency of frameshifting errors [101]. The results of this work indicate that natural synonymous codon usage is biased toward specific patterns correlated with avoidance of mistranslation and frameshifting-induced protein misfolding [101]. Overall, an understanding of the impact of codon usage on mistranslation and frameshifting errors may be helpful in minimizing the risk of producing subpopulation of proteins with different amino acid sequences when undertaking recombinant protein production from a redesigned gene.

5 The Impact of Single Synonymous Codon Substitutions

Gene redesign usually involves numerous substitutions of synonymous codons. However, recent studies have shown that some specific single synonymous mutations are deleterious for proper protein expression and, moreover, organism health ([51, 52] for a review). The majority of identified deleterious single synonymous mutations exert effects on mRNA splicing (in eukaryotes), but there are also quite a few that may alter protein folding and, as a consequence, protein activity and/or resistance to degradation [51, 52]. These single synonymous mutations can produce disease in the expressing organism, and their inadvertent introduction into genes of therapeutic proteins may produce undesirable effects. It should be noted that the exact mechanisms underlying the effects of many synonymous mutations linked to disease are not yet well understood [51, 52]. One of the major challenges in the field is to understand why some disease-causing synonymous mutations are more deleterious than others and to predict the likely effects of a single mutation.

Evaluation of mRNA stability of fragments of genes of several proteins carrying neutral vs. disease-associated mutations and synonymous vs. non-synonymous mutations revealed that deleterious synonymous mutations tend to occur in mRNA regions with higher MFE levels and often lead to a reduction in MFE [102–105]. It is not yet clear how broadly applicable this situation originally

identified for "disease-associated" mutations in the *F8* and *F9* genes encoding blood-coagulation factors VIII and IX, respectively, might be [102, 105]. Mutations in the *F8* and *F9* genes lead to blood clotting disorders known as hemophilia A and B [102, 105]. While further investigation into the deleterious effects of specific synonymous mutations is required, it is clear that known disease-associated mutations should be avoided in gene redesign efforts.

6 Concluding Remarks and Future Perspectives

Gene design and redesign approaches target protein-coding genes and aim to introduce predefined features of interest into the final protein product. These approaches frequently involve changes in synonymous codon usage intended to improve protein production in homologous and/or heterologous hosts without compromising the integrity of the encoded protein. Optimization of gene design and protein production is of strong significance due to the high, and continually increasing, demand for recombinant proteins for use in research and in therapeutic applications. Advances in DNA synthesis have enabled construction of numerous gene variants and facilitated our understanding of the impact of codon usage on gene function. Additional knowledge came from genome-wide studies aimed at uncovering the impact of synonymous mutations on gene function and phenotype and understanding their association with various diseases.

It has become clear that synonymous codon usage and synonymous mutations do not only alter the speed of protein synthesis but affect many critical aspects of mRNA and protein biogenesis (ranging from mRNA stability to protein mistranslation and folding), thus ultimately changing the phenotype associated with the protein. Importantly, it was revealed that even a single synonymous mutation may be deleterious to protein function. While complete understanding of the effects caused by multiple and single synonymous mutations remains lacking, it is possible, as done in this review, to use existing knowledge to develop some common rules to gene design and redesign that should increase the probability of achieving the desired quantity and activity of an expressed recombinant protein.

A combination of evolutionary, computational, and synthetic biology should ultimately enable (1) full genome-based understanding of the impact of individual synonymous mutations on gene function, mRNA biogenesis, protein production, and protein folding; (2) efficient manufacturing of safer, more effective, and even potentially individualized protein therapeutics; and (3) improved understanding of evolutionary processes.

7 Notes

1. Carefully planned placement of rare/non-optimal (or AT-rich) codons in the 5′ termini of mRNA ORFs, especially those encoding secretory and transmembrane proteins, may represent an important strategy for successful gene design and redesign enhancing proper protein production, secretion, and folding.
2. Enrichment of rare codon clusters at specific locations in a broad range of genes implies that they have functional significance. Therefore, strategic placement of evolutionarily conserved rare codon clusters within ORFs may facilitate correct protein folding.
3. Use of optimal synonymous codons during gene design and redesign may lead to substantial stabilization of the mRNA and enhancement of protein production (at least in yeast and *E. coli*).
4. Mistranslation as a result of synonymous codon changes may lead to incorrect protein folding; this should be taken into consideration when planning production of recombinant proteins.
5. Although a variety of methods are available for gene redesign, approaches that take into account the effect(s) of synonymous codon substitutions on translation efficiency, protein folding, and protein activity will allow the most productive manufacturing of safer and more effective protein therapeutics.

Acknowledgments I apologize to those whose work or original publications could not be cited in this article because of space limitations. I thank Patricia Stanhope Baker for help with manuscript preparation. This work was supported in part by grants to A.A.K. from the Human Frontier Science Program (grant # RGP0024/2010), AHA (grant # 13GRNT17070025), and NIH (grant # 1R15HL121779).

References

1. Schmidt FR (2004) Recombinant expression systems in the pharmaceutical industry. Appl Microbiol Biotechnol 65:363–372
2. Berlec A, Strukelj B (2013) Current state and recent advances in biopharmaceutical production in *Escherichia coli*, yeasts and mammalian cells. J Ind Microbiol Biotechnol 40:257–274
3. Khan KH (2013) Gene expression in Mammalian cells and its applications. Adv Pharm Bull 3:257–263
4. Assenberg R, Wan PT, Geisse S, Mayr LM (2013) Advances in recombinant protein expression for use in pharmaceutical research. Curr Opin Struct Biol 23:393–402
5. Young CL, Robinson AS (2014) Protein folding and secretion: mechanistic insights advancing recombinant protein production in *S. cerevisiae*. Curr Opin Biotechnol 30:168–177
6. Sugiki T, Fujiwara T, Kojima C (2014) Latest approaches for efficient protein production in drug discovery. Expert Opin Drug Discov 9:1189–1204
7. van Oers MM, Pijlman GP, Vlak JM (2015) Thirty years of baculovirus-insect cell protein expression: from dark horse to mainstream technology. J Gen Virol 96:6–23
8. Carlson ED, Gan R, Hodgman CE, Jewett MC (2012) Cell-free protein synthesis: applications come of age. Biotechnol Adv 30:1185–1194

9. Whittaker JW (2013) Cell-free protein synthesis: the state of the art. Biotechnol Lett 35:143–152
10. Yusibov V, Streatfield SJ, Kushnir N (2011) Clinical development of plant-produced recombinant pharmaceuticals: vaccines, antibodies and beyond. Hum Vaccin 7:313–321
11. Abiri R, Valdiani A, Maziah M, Shaharuddin NA, Sahebi M, Yusof ZY, Atabaki N, Talei D (2015) A critical review of the concept of transgenic plants: insights into pharmaceutical biotechnology and molecular farming. Curr Issues Mol Biol 18:21–42
12. Houdebine LM (2000) Transgenic animal bioreactors. Transgenic Res 9:305–320
13. Bösze Z, Baranyi M, Whitelaw CB (2008) Producing recombinant human milk proteins in the milk of livestock species. Adv Exp Med Biol 606:357–393
14. Demain AL, Vaishnav P (2009) Production of recombinant proteins by microbes and higher organisms. Biotechnol Adv 27:297–306
15. Klammt C, Schwarz D, Löhr F, Schneider B, Dötsch V, Bernhard F (2006) Cell-free expression as an emerging technique for the large scale production of integral membrane protein. FEBS J 273:4141–4153
16. Saïda F (2007) Overview on the expression of toxic gene products in *Escherichia coli*. Curr Protoc Protein Sci 50:1–5
17. Ryabova LA, Morozov IY, Spirin AS (1998) Continuous-flow cell-free translation, transcription-translation, and replication-translation systems. Methods Mol Biol 77:179–193
18. Murray CJ, Baliga R (2013) Cell-free translation of peptides and proteins: from high throughput screening to clinical production. Curr Opin Chem Biol 17:420–426
19. Gustafsson C, Govindarajan S, Minshull J (2004) Codon bias and heterologous protein expression. Trends Biotechnol 22:346–353
20. Elena C, Ravasi P, Castelli ME, Peirú S, Menzella HG (2014) Expression of codon optimized genes in microbial systems: current industrial applications and perspectives. Front Microbiol 5:21
21. Quax TE, Claassens NJ, Söll D, van der Oost J (2015) Codon bias as a means to fine-tune gene expression. Mol Cell 59:149–161
22. Sharp PM, Cowe E, Higgins DG, Shields DC, Wolfe KH, Wright F (1998) Codon usage patterns in *Escherichia coli, Bacillus subtilis, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Drosophila melanogaster* and *Homo sapiens*; a review of the considerable within-species diversity. Nucleic Acids Res 16:8207–18211
23. Hershberg R, Petrov DA (2008) Selection on codon bias. Annu Rev Genet 42:287–299
24. Sharp PM, Emery LR, Zeng K (2010) Forces that influence the evolution of codon bias. Philos Trans R Soc Lond B Biol Sci 365:1203–1212
25. Ikemura T (1985) Codon usage and tRNA content in unicellular and multicellular organisms. Mol Biol Evol 2:13–34
26. Buchan JR, Stansfield I (2007) Halting a cellular production line: responses to ribosomal pausing during translation. Biol Cell 99:475–487
27. Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science 324:218–223
28. Komar AA (2009) A pause for thought along the co-translational folding pathway. Trends Biochem Sci 34:16–24
29. Ingolia NT (2014) Ribosome profiling: new views of translation, from single codons to genome scale. Nat Rev Genet 15:205–213
30. Dana A, Tuller T (2014) The effect of tRNA levels on decoding times of mRNA codons. Nucleic Acids Res 42:9171–9181
31. Gardin J, Yeasmin R, Yurovsky A, Cai Y, Skiena S, Futcher B (2014) Measurement of average decoding rates of the 61 sense codons in vivo. Elife 3, eLife.03735
32. Hatfield GW, Roth DA (2007) Optimizing scaleup yield for protein production: computationally optimized DNA assembly (CODA) and translation engineering. Biotechnol Annu Rev 13:27–42

33. Kudla G, Murray AW, Tollervey D, Plotkin JB (2009) Coding-sequence determinants of gene expression in *Escherichia coli*. Science 324:255–258
34. Tuller T, Waldman YY, Kupiec M, Ruppin E (2010) Translation efficiency is determined by both codon bias and folding energy. Proc Natl Acad Sci U S A 107:3645–3650
35. Kim HJ, Lee SJ, Kim HJ (2010) Optimizing the secondary structure of human papillomavirus type 16 L1 mRNA enhances L1 protein expression in *Saccharomyces cerevisiae*. J Biotechnol 150:31–36
36. Castillo-Méndez MA, Jacinto-Loeza E, Olivares-Trejo JJ, Guarneros-Pena G, Hernandez-Sanchez J (2012) Adenine-containing codons enhance protein synthesis by promoting mRNA binding to ribosomal 30S subunits provided that specific tRNAs are not exhausted. Biochimie 94:662–672
37. Goodman DB, Church GM, Kosuri S (2013) Causes and effects of N-terminal codon bias in bacterial genes. Science 342:475–479
38. Bentele K, Saffert P, Rauscher R, Ignatova Z, Bluthgen N (2013) Efficient translation initiation dictates codon usage at gene start. Mol Syst Biol 9:675
39. Li GW (2015) How do bacteria tune translation efficiency? Curr Opin Microbiol 24:66–71
40. Wu G, Zheng Y, Qureshi I, Zin HT, Beck T, Bulka B, Freeland SJ (2007) SGDB: a database of synthetic genes re-designed for optimizing protein over-expression. Nucleic Acids Res 35: D76–D79
41. de Marco A, Vigh L, Diamant S, Goloubinoff P (2005) Native folding of aggregation-prone recombinant proteins in *Escherichia coli* by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones. Cell Stress Chaperones 10:329–339
42. Komar AA, Lesnik T, Reiss C (1999) Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation. FEBS Lett 462:387–391
43. Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV, Gottesman MM (2007) A "silent" polymorphism in the MDR1 gene changes substrate specificity. Science 315:525–528
44. Yu CH, Dang Y, Zhou Z, Wu C, Zhao F, Sachs MS, Liu Y (2015) Codon usage influences the local rate of translation elongation to regulate co-translational protein folding. Mol Cell 59:744–754
45. Zhou M, Guo J, Cha J, Chae M, Chen S, Barral JM, Sachs MS, Liu Y (2013) Non-optimal codon usage affects expression, structure and function of clock protein FRQ. Nature 495:111–115
46. Zhang G, Hubalewska M, Ignatova Z (2009) Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. Nat Struct Mol Biol 16:274–280
47. Sander IM, Chaney JL, Clark PL (2014) Expanding Anfinsen's principle: contributions of synonymous codon selection to rational protein design. J Am Chem 136:858–861
48. Hu S, Wang M, Cai G, He M (2013) Genetic code-guided protein synthesis and folding in *Escherichia coli*. J Biol Chem 288:30855–30861
49. Kim SJ, Yoon JS, Shishido H, Yang Z, Rooney LA, Barral JM, Skach WR (2015) Protein folding. Translational tuning optimizes nascent protein folding in cells. Science 348:444–448
50. Buhr F, Jha S, Thommen M, Mittelstaet J, Kutz F, Schwalbe H, Rodnina MV, Komar AA (2016) Synonymous codons direct cotranslational folding toward different protein conformations. Mol Cell 61:341–351. http://www.sciencedirect.com/science/article/pii/S1097276516000095
51. Sauna ZE, Kimchi-Sarfaty C (2011) Understanding the contribution of synonymous mutations to human disease. Nat Rev Genet 12:683–691
52. Hunt RC, Simhadri VL, Iandoli M, Sauna ZE, Kimchi-Sarfaty C (2014) Exposing synonymous mutations. Trends Genet 30:308–321
53. Presnyak V, Alhusaini N, Chen YH, Martin S, Morris N, Kline N, Olson S, Weinberg D, Baker KE, Graveley BR, Coller J (2015) Codon optimality is a major determinant of mRNA stability. Cell 160:1111–1124

54. Boël G, Letso R, Neely H, Price WN, Wong KH, Su M, Luff JD, Valecha M, Everett JK, Acton TB, Xiao R, Montelione GT, Aalberts DP, Hunt JF (2016) Codon influence on protein expression in *E. coli* correlates with mRNA levels. Nature 529:358–363
55. Pechmann S, Chartron JW, Frydman J (2014) Local slowdown of translation by nonoptimal codons promotes nascent-chain recognition by SRP in vivo. Nat Struct Mol Biol 21:1100–1105
56. Drummond DA, Wilke CO (2008) Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. Cell 134:341–352
57. Chin JX, Chung BK-S, Lee D-Y (2014) Codon optimization on-line (COOL): a web-based multi-objective optimization platform for synthetic gene design. Bioinformatics 30:2210–2212
58. Hoover DM, Lubkowski J (2002) DNA Works: an automated method for designing oligonucleotides for PCR-based gene synthesis. Nucleic Acids Res 30, e43
59. Guimaraes JC, Rocha M, Arkin AP, Cambray G (2014) D-Tailor: automated analysis and design of DNA sequences. Bioinformatics 30:1087–1094
60. Gaspar P, Oliveira JL, Frommlet J, Santos MAS, Moura G (2012) EuGene: maximizing synthetic gene design for heterologous expression. Bioinformatics 28:2683–2684
61. Richardson SM, Wheelan SJ, Yarrington RM, Boeke JD (2006) GeneDesign: rapid, automated design of multikilobase synthetic genes. Genome Res 16:550–556
62. Villalobos A, Ness JE, Gustafsson C, Minshull J, Govindarajan S (2006) Gene designer: a synthetic biology tool for constructing artificial DNA segments. BMC Bioinformat 7:285
63. Grote A, Hiller K, Scheer M, Münch R, Nörtemann B, Hempel DC, Jahn D (2005) JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res 33:W526–W531
64. Gaspar P, Moura G, Santos MAS, Oliveira JL (2013) mRNA secondary structure optimization using a correlated stem-loop prediction. Nucleic Acids Res 41, e73
65. Puigbò P, Guzmán E, Romeu A, Garcia-Vallvé S (2007) Optimizer: a web server for optimizing the codon usage of DNA sequences. Nucleic Acids Res 35:W126–W131
66. Wu G, Bashir-Bello N, Freeland S (2005) The synthetic gene designer: a flexible web platform to explore sequence space of synthetic genes for heterolo-gous expression. In: 2005 I.E. computational systems bioinformatics conference, workshops and poster abstracts, 2005 Aug 8–11. Stanford University, California, pp 258–259
67. Li MH, Bode M, Huang MC, Cheong WC, Lim LS (2012) *De novo* gene synthesis design using TmPrime software. Methods Mol Biol 852:225–234
68. Jung S-K, McDonald K (2011) Visual gene developer: a fully programmable bioinformatics software for synthetic gene optimization. BMC Bioinformat 12:340
69. Gould N, Hendy O, Papamichail D (2014) Computational tools and algorithms for designing customized synthetic genes. Front Bioeng Biotechnol 2:41
70. Nakamura Y, Gojobori T, Ikemura T (2000) Codon usage tabulated from the international DNA sequence databases: status for the year 2000. Nucleic Acids Res 28:292
71. Sharp PM, Li WH (1987) The codon Adaptation Index – a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res 15:1281–1295
72. Plotkin JB, Kudla G (2011) Synonymous but not the same: the causes and consequences of codon bias. Nat Rev Genet 12:32–42
73. Pechmann S, Frydman J (2011) Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. Nat Struct Mol Biol 20:237–243
74. Chaney JL, Clark PL (2015) Roles for synonymous codon usage in protein biogenesis. Annu Rev Biophys 44:143–166
75. Clarke TF 4th, Clark PL (2008) Rare codons cluster. PLoS One 3, e3412
76. Clarke TF 4th, Clark PL (2010) Increased incidence of rare codon clusters at 5′ and 3′ gene termini: implications for function. BMC Genomics 11:118
77. Allert M, Cox JC, Hellinga HW (2010) Multifactorial determinants of protein expression in prokaryotic open reading frames. J Mol Biol 402:905–918

78. Gu W, Zhou T, Wilke CO (2010) A universal trend of reduced mRNA stability near the translation-initiation site in prokaryotes and eukaryotes. PLoS Comput Biol 6, e1000664
79. Zalucki YM, Beacham IR, Jennings MP (2009) Biased codon usage in signal peptides: a role in protein export. Trends Microbiol 17:146–150
80. Fluman N, Navon S, Bibi E, Pilpel Y (2014) mRNA-programmed translation pauses in the targeting of *E. coli* membrane proteins. Elife 3:eLife.03440
81. Chartier M, Gaudreault F, Najmanovich R (2012) Large-scale analysis of conserved rare codon clusters suggests an involvement in co-translational molecular recognition events. Bioinformatics 28:1438–1445
82. McKown RL, Raab RW, Kachelries P, Caldwell S, Laurie GW (2013) Conserved regional 3′ grouping of rare codons in the coding sequence of ocular prosecretory mitogen lacritin. Invest Ophthalmol Vis Sci 54:1979–1987
83. Widmann M, Clairo M, Dippon J, Pleiss J (2008) Analysis of the distribution of functionally relevant rare codons. BMC Genomics 9:207
84. Purvis IJ, Bettany AJ, Santiago TC, Coggins JR, Duncan K, Eason R, Brown AJ (1987) The efficiency of folding of some proteins is increased by controlled rates of translation in vivo. A hypothesis. J Mol Biol 193:413–417
85. Krasheninnikov IA, Komar AA, Adzhubeĭ IA (1988) Role of the rare codon clusters in defining the boundaries of polypeptide chain regions with identical secondary structures in the process of co-translational folding of proteins. Dokl Akad Nauk SSSR 303:995–999
86. Tsai CJ, Sauna ZE, Kimchi-Sarfaty C, Ambudkar SV, Gottesman MM, Nussinov R (2008) Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima. J Mol Biol 383:281–291
87. Kramer G, Boehringer D, Ban N, Bukau B (2009) The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. Nat Struct Mol Biol 16:589–597
88. Zhang G, Ignatova Z (2011) Folding at the birth of the nascent chain: coordinating translation with co-translational folding. Curr Opin Struct Biol 21:25–31
89. Waudby CA, Launay H, Cabrita LD, Christodoulou J (2013) Protein folding on the ribosome studied using NMR spectroscopy. Prog Nucl Magn Reson Spectrosc 74:57–75
90. O'Brien EP, Ciryam P, Vendruscolo M, Dobson CM (2014) Understanding the influence of codon translation rates on cotranslational protein folding. Acc Chem Res 47:1536–1544
91. Gloge F, Becker AH, Kramer G, Bukau B (2014) Co-translational mechanisms of protein maturation. Curr Opin Struct Biol 24:24–33
92. Hoekema A, Kastelein RA, Vasser M, de Boer HA (1987) Codon replacement in the PGK1 gene of *Saccharomyces cerevisiae*: experimental approach to study the role of biased codon usage in gene expression. Mol Cell Biol 7:2914–2924
93. Caponigro G, Muhlrad D, Parker R (1993) A small segment of the MAT alpha 1 transcript promotes mRNA decay in *Saccharomyces cerevisiae*: a stimulatory role for rare codons. Mol Cell Biol 13:5141–5148
94. Deana A, Ehrlich R, Reiss C (1996) Synonymous codon selection controls in vivo turnover and amount of mRNA in *Escherichia coli* bla and ompA genes. J Bacteriol 178:2718–2720
95. Kramer EB, Farabaugh PJ (2007) The frequency of translational misreading errors in *E. coli* is largely determined by tRNA competition. RNA 13:87–96
96. Zaher HS, Green R (2009) Fidelity at the molecular level: lessons from protein synthesis. Cell 136:746–762
97. Kramer EB, Vallabhaneni H, Mayer LM, Farabaugh PJ (2010) A comprehensive analysis of translational missense errors in the yeast *Saccharomyces cerevisiae*. RNA 16:1797–1808
98. Ribas de Pouplana L, Santos MA, Zhu JH, Farabaugh PJ, Javid B (2014) Protein mistranslation: friend or foe? Trends Biochem Sci 39:355–362
99. Dinman JD (2012) Mechanisms and implications of programmed translational frameshifting. Wiley Interdiscip Rev RNA 3:661–673

100. Caliskan N, Peske F, Rodnina MV (2015) Changed in translation: mRNA recoding by -1 programmed ribosomal frameshifting. Trends Biochem Sci 40:265–274
101. Huang Y, Koonin EV, Lipman DJ, Przytycka TM (2009) Selection for minimization of translational frameshifting errors as a factor in the evolution of codon usage. Nucleic Acids Res 37:6799–6810
102. Hamasaki-Katagiri N, Salari R, Simhadri VL, Tseng SC, Needlman E, Edwards NC, Sauna ZE, Grigoryan V, Komar AA, Przytycka TM, Kimchi-Sarfaty C (2012) Analysis of *F9* point mutations and their correlation to severity of haemophilia B disease. Haemophilia 18:933–940
103. Edwards NC, Hing ZA, Perry A, Blaisdell A, Kopelman DB, Fathke R, Plum W, Newell J, Allen CE, Shapiro SGA, Okunji C, Kosti I, Shomron N, Grigoryan V, Przytycka TM, Sauna ZE, Salari R, Mandel-Gutfreund Y, Komar AA, Kimchi-Sarfaty C (2012) Characterization of coding synonymous and non-synonymous variants in ADAMTS13 using ex vivo and in silico approaches. PLoS One **7**:e38864
104. Salari R, Kimchi-Sarfaty C, Gottesman MM, Przytycka TM (2013) Sensitive measurement of single-nucleotide polymorphism-induced changes of RNA conformation: application to disease studies. Nucleic Acids Res 41:44–53
105. Hamasaki-Katagiri N, Salari R, Wu A, Qi Y, Schiller T, Filiberto AC, Schisterman EF, Komar AA, Przytycka TM, Kimchi-Sarfaty C (2013) A gene-specific method for predicting hemophilia-causing point mutations. J Mol Biol 425:4023–4033

Index