

Chapter 11

Molecular Modeling and Drug Design

Techniques in Microbial Drug Discovery



Chandrabose Selvaraj

Contents

11.1	Introduction.....	187
11.2	Global Battle Against Infectious Diseases.....	188
11.3	Methods in Drug Design.....	189
11.3.1	Structure-Based Drug Design.....	189
11.3.2	Ligand-Based Drug Design (LBDD).....	190
11.3.2.1	Pharmacophore Modeling.....	191
11.3.2.2	QSAR Modeling.....	194
11.3.2.3	CoMFA.....	195
11.3.2.4	Comparative Molecular Similarity Indices Analysis (CoMSIA).....	196
11.4	Virtual Screening (VS) for Lead Discovery.....	196
11.4.1	Protein Modeling.....	197
11.4.1.1	Homology Modeling.....	198
11.4.1.2	Threading.....	200
11.4.1.3	Ab Initio Method.....	201
11.4.1.4	Protein Validation Server.....	202
11.4.2	Protein and Ligand Preparation.....	203
11.4.3	Active Site Prediction.....	205
11.4.4	Molecular Docking.....	205
11.4.5	Scoring Methods.....	208
11.4.6	Molecular Dynamics (MD) Simulations.....	210
11.4.7	QM/MM Simulations.....	211
11.5	Drug Delivery Approach Using Computational Methods.....	213
11.6	Polymer Used as Carrier.....	214
11.6.1	Drug-Polymer Interaction.....	215
11.7	Computational Methods Used in Toxicity Studies.....	216
11.8	Outcome of Drug Research in Bacterial Inhibitors.....	218
11.9	Future Aspects of Computational Methods in Targeting Bacterial Infections.....	219
	References.....	219

C. Selvaraj (✉)

School of Basic Sciences, Indian Institute of Technology, Mandi,
Kamand, Himachal Pradesh, India

© Springer Nature Switzerland AG 2019

N. A. Shaik et al. (eds.), *Essentials of Bioinformatics, Volume II*,
https://doi.org/10.1007/978-3-030-18375-2_11

185

Abbreviations

AMR	Antimicrobial resistance
BLAST	Basic Local Alignment Search Tool
CADD	Computer aided drug designing
CG	Coarse-grained
CG-MD	Coarse-grained molecular dynamics
CHARMM	Chemistry at Harvard Macromolecular Mechanics
CoMFA	Comparative molecular field analysis
CoMSIA	Comparative molecular similarity indices analysis
DADA	D-alanyl-D-alanine
DFT	Density functional theory
DPD	Dissipative particle dynamics
ESBLs	Extended spectrum β -lactamases
FDA	Food and Drug Administration
FEP	Free-energy perturbation method
GA	Genetic algorithms
GISA	Glycopeptides-intermediately-resistant <i>S. aureus</i>
GPU	Graphical processor unit
HTS	High throughput screening
IUPAC	International Union of Pure and Applied Chemistry
LB	Ligand-based
LBDD	Ligand-based drug design
LBVS	Ligand-based virtual screening
MD	Molecular dynamics
MDR	Multi-drug resistance
MRSA	<i>Staphylococcus aureus</i> resistant to methicillin
NCBI	National Center for Biotechnology Information
NMR	Nuclear magnetic resonance
PCA	Principle component analysis
PCR	Polymerase chain reaction
PK	Pharmacokinetic
PLS	Partial least squares
QM/MM	Quantum mechanics/molecular mechanics
QSAR	Quantitative structure-activity relationship
SB	Structure-based
SBDD	Structure-based drug design
SBVS	Structure-based virtual screening
SI	Sequence identity
TEIC	Teicoplanin
VANC	Vancomycin

11.1 Introduction

A potential drug molecule is one that effectively binds and modulates a molecular target in such a manner that is less toxic, safe and effective in the disease context for which it is doled out. The drug discovery development is a complex process, which can take 12–15 years and entail costs of more than \$1 billion. In the modern era of drug discovery, development involves the cooperation of many disciplines such as chemistry, biology, mathematics and computer science (Herrling 2005). A chemical moiety with significant therapeutic value is extensively analyzed for its safety and efficacy before it is marketed. The multistep process, termed ‘drug discovery,’ includes identification and validation of the drug target and of the lead molecule. The drug development process is categorized, basically, into the two major phases of drug discovery and drug development. The drug discovery process involves two important approaches; identification and validation of a potential disease-oriented target molecule and another approach is phenotypic screening to identify and refine the potential small molecules that can interact with target (Ernst and Obrecht 2008). This molecular interaction can be to block, promote or modify the activity of the target. In recent years, the drug discovery process has undergone radical changes due to the entry of various novel techniques in genomics; proteomics have been developed in drug target identification and validation has become more specific (Umashankar and Gurunathan 2015). In the past decade, emergence of microbial resistance (Amini and Tavazoie 2011) and complicated new diseases and unexpected adverse side effects have accelerated the identification of potent lead molecules (Ashrafuzzaman 2014). Infectious diseases, particularly Gram-positive bacterial infection, are among the major serious threats to public health worldwide: they are difficult to treat and are associated with high morbidity and mortality rates. Gram-negative bacteria are highly adaptive pathogens that produce resistance to antibiotics through several mechanisms. The production of β -lactamases and hydrolyzation of the β -lactam ring represents the most common resistant mechanism in Gram-negative bacteria against β -lactam antibiotics. Most bacteria can develop and adapt themselves according to their surroundings and subsequently develop several protective mechanisms to reduce their susceptibility to antibiotics. In some cases, bacteria allow horizontal gene transfer within and between species to become more resistant to antibiotics (Palumbi 2001; Thomas and Nielsen 2005). This horizontal gene transfer provides the most important mechanism to accelerate the spectrum of β -lactamases (ESBLs), causing severe problems in drug resistant in the health care world (Giske et al. 2008; Hawkey and Jones 2009). Bacterial strains capable of producing ESBLs are resistant to several antibiotics, including penicillins and cephalosporins, and they are resistant to other antibacterials such as quinolones and aminoglycosides. This antibiotic resistance shows a strong correlation between the segment of the population that uses antibiotics and the prevalence of antibiotic-resistant bacteria in the same population; the correlation has been found on both national and regional levels (Bronzwaer et al. 2002; Albrich et al. 2004).

11.2 Global Battle Against Infectious Diseases

In the middle of the seventeenth century, smallpox infection was the most fatal and feared of diseases. The discovery of penicillin developed a new generation of antibiotics that cured a wide range of infectious diseases. Several researches focused on understanding what mechanisms the microbes used to survive antibiotics, and several pharmaceutical and biotech companies nearly stampeded to identify a significant bacterial target and to create novel methodologies against the bacteria. Recent evidence suggests that mutation with humans is not the only way bacteria develop antibiotic resistance; they can also transfer genetic instructions for avoiding an antibiotic to other bacterial species. In the late 1800s, pathogen-specific medical diagnosis lent a hand to the identification of microbes that caused specific diseases. Molecular genetics technique, polymerase chain reaction (PCR) and, more recently, sophisticated, high throughput rapid sequencing of the genome of the pathogen are all used to observe the individual genetic variants, facilitating identification of the familial base of drug immunity. Other factor-based, diagnostic tools including microchip and serological techniques and enzyme-linked immunosorbent assay can be more sensitive than traditional techniques in finding and measuring antibodies to pathogens (Pallen et al. 2010). Current data suggest that Gram-positive bacteria cause 45–70% of infectious diseases and are behind the increase in rates of drug resistance in many infections. The pace of drug resistance among bacterial pathogens is increasing; virtually no new antibiotics are being developed (Spellberg et al. 2004). Gram-positive organisms such as the bacteria of the genera *Staphylococcus*, *Streptococcus* and *Enterococcus* are the predominant bacterial spp causing clinical infection, hence, recent attention has focused on the multi-drug resistance (MDR) and antimicrobial resistance (AMR) (Menichetti 2005; Doernberg et al. 2017).

Sulfonamide synthetic antimetabolites were first used clinically in 1932 for a wide range of both Gram-negative and Gram-positive bacteria. These synthetic metabolites inhibit dihydropteroate synthetase leads to repressed DNA replication. Until 1938, β -lactam was another widely used antibiotic. The 28 members that include antibiotics/ β -lactamase inhibitor combinations are broadly classified into three subclasses: penicillins, cephalosporins and carbapenems, which are critically used in very broad-spectrum activity against most aerobic and anaerobic Gram-positive and Gram-negative bacteria (Walsh 2003; Collignon et al. 2009; Lewis 2013). Recently, glycopeptides like vancomycin (VANC) and teicoplanin (TEIC) have been widely used against Gram-positive bacteria; these share a mechanism of natural process similar to that of β -lactams, except their interruption on cell wall synthesis via an interaction with the D-alanyl-D-alanine (DADA) moiety of peptidoglycan precursors inhibits the cross-linking stabilization step in bacterial cell wall formation (Malabarba and Goldstein 2005). The cyclic lipopeptide daptomycin has an extensive range of activity on Gram-positive bacterial infection and also on MRSA. Structurally, daptomycin comprises a 13-member hydrophobic polypeptide with a lipophilic side chain having a unique mechanism of natural process, which is leads insertion of the lipophilic region into the bacterial cell wall, oligomerizing

into pore-like constructions, through which a significant efflux of potassium ions results in rapid bacterial cell death (Silverman et al. 2003; Steenbergen et al. 2005).

11.3 Methods in Drug Design

Drug development commences with the identification of a molecular target and lead molecules followed by lead optimization and preclinical *in vitro* and *in vivo* studies to recognize potent compounds that fulfill the primary criteria for the drug development (Bleicher et al. 2003). But, the development of lead molecules through *in vitro* and *in vivo* methods takes a long time and is very expensive (DiMasi et al. 2003); hence, in recent years *in silico* drug designing has been widely used to predict active lead molecules. Here, we look at discovery. Traditional drug discovery (*in vitro* and *in vivo*) requires about 12–14 years and costs up to \$1.2–\$1.4 billion dollars to get a drug from discovery to market (Hileman 2006). About 90% of the drugs entering clinical trials fail to obtain FDA approval and reach the consumer market (Tollman 2001). Lately, high throughput screening (HTS) experiments are used to sort thousands of molecules with robotic automation; however, HTS is still expensive and requires a great amount of resources. Therefore, computer-aided drug designing (CADD) can cut cost- and time-associated drawbacks and ensure the best possible lead compounds are used in animal studies. CADD tools have not merely been applied to distinguish potential lead molecules; they can also predict effectiveness and possible side effects and aid in improving bioavailability of the possible drug molecules (Yang et al. 2016). CADD plays a crucial role in the identification of many pharmaceutically available drugs, ones that have obtained FDA approval and reached the consumer market (Kitchen et al. 2004; Clark 2006; Talele et al. 2010). CADD methods are broadly classified into two categories: structure-based (SB) drug discovery and ligand-based (LB) drug discovery.

11.3.1 Structure-Based Drug Design

Structure-based drug design (SBDD) methods are prominent tools in modern medicinal chemistry that utilize three-dimensional structural information from biological targets (Salum et al. 2008). Understanding the mechanism of small molecule reorganization and interaction with biological macromolecules is of great importance in pharmaceutical research and development. In recent years, due to wide range of application such as molecular docking, molecular dynamic simulation, and structure-based virtual screening (SBVS), SBDD has played a crucial role in the identification of potential drug molecules against various drug target (Kalyanamoorthy and Chen 2011). In SBDD, binding site topology (including clefts, cavities and sub-pockets) and the electrostatic properties of the target molecule were carefully examined (Wilson and Lill 2011).

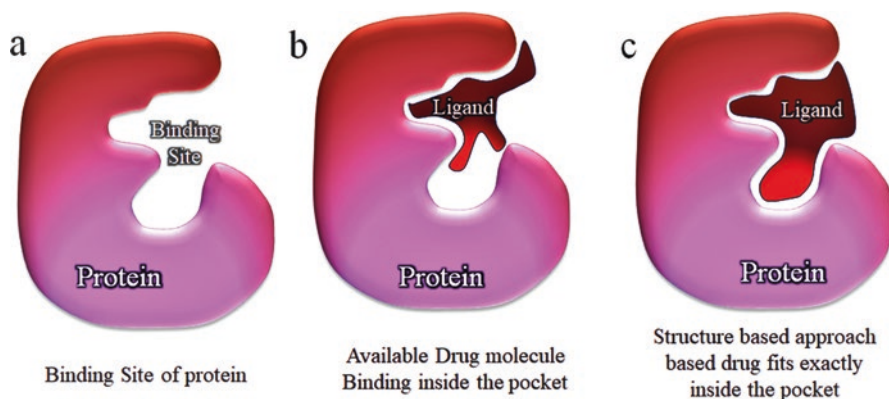


Fig. 11.1 Mechanism of SBDD showing the design of a new molecule as per the binding site feature of a protein

SBDD is an iterative method involving multiple steps for finding a lead. The first step of SBDD includes the cloning, purification and structure elucidation of the target proteins or nucleic acid by NMR, X-ray crystallography or homology modeling, identification of potential ligand molecules and evaluation of biological properties, such as potency, affinity and efficacy, as carried out through various experimental analyses (Fang 2012). It also provides the structural descriptions of the target-ligand complex for understanding the binding mode and conformations, characterization of key molecular interaction, characterization of unknown binding sites, mechanistic studies and elucidation of ligand-induced conformational changes (Kahsai et al. 2011). Methods used in SBDD such as molecular dynamics give insight into not only how ligands bind with target proteins but also consider the target flexibility and interaction of pathway. SBDD has contributed to several compounds reaching the clinical trial stage and getting FDA approval to go into the market (Burger and Abraham 2006; Wang et al. 2010; Hanson et al. 2015). Thus, SBDD is a cyclic process consisting of several steps, starting from a known target structure, then going on to several *in silico* studies, which are conducted to identify potential ligands. The mechanism of structure-based drug design is explained in Fig. 11.1, which shows the binding site feature of the protein (Fig. 11.1a); the available drug molecules displaying the binding phenomenon with the binding site, with a few empty spaces that may be filled with water molecules (Fig. 11.1b); and finally the new drug, designed as per the binding site feature that perfectly fits with the binding site (Fig. 11.1c).

11.3.2 Ligand-Based Drug Design (LBDD)

LBDD is an one the often used method in computer aided drug design effectively used in the absence of the 3D structure of the target and the binding site is not accurately known, then a ligand-based drug design (LBDD) approach is a popular

technique in the case of experimentally active compounds that bind to the biological target of interest. The common assumption in drug identification is that similar compounds with similar chemical properties may exhibit similar biological activity. Ligand-based virtual screening (LBVS) is based on the exploration of molecular descriptors gathered from known active compounds. In general, similar characteristics of a compound series are identified and subsequently applied as molecular filters. These filtering methods are used to discover potential lead molecules for experimental evaluation and reduce the chemical space to be explored in further screening steps (Geppert et al. 2010; Sliwoski et al. 2013). This is the main principle and motivation of LBDD, where a compound with interesting biological properties can act as template for finding potential lead molecules. Basically, three approaches—2D fingerprints, 3D methods and pharmacophores—are widely used for defining and quantifying chemical similarity in LBDD.

11.3.2.1 Pharmacophore Modeling

Pharmacophore model prediction is an essential way to describe those steric and electronic features needed for optimal interaction of lead with receptor molecules. According to the International Union of Pure and Applied Chemistry (IUPAC), pharmacophore is “the ensemble of steric and electronic features ... necessary to ensure the optimal supramolecular interactions with a specific biological target structure to trigger or to block its biological activity.” (Kaserer et al. 2015). In drug discovery approaches with small molecules, it is important to analyze the assignment of proper protonation and tautomeric states of the lead molecules. Pharmacophore describes a set of interactions required to bind in the cavity of target molecules and a set of spatially arranged spheres of a certain type and diameter. These spheres are commonly known as pharmacophoric features (Fig. 11.2). They include hydrophobic centroids, hydrogen-bond acceptor, hydrogen-bond donor, positively ionizable groups and negatively ionizable groups— all common features which target their

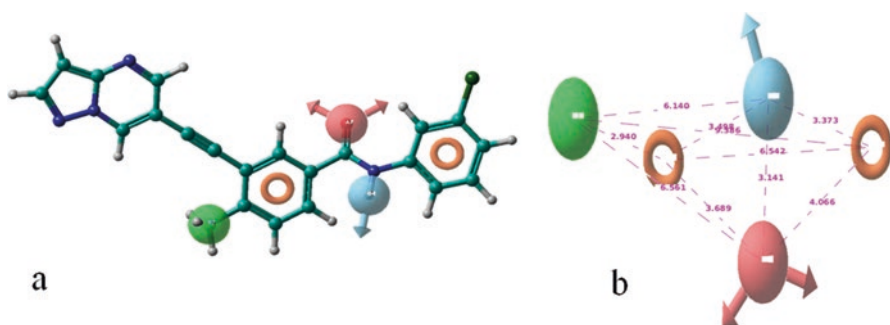


Fig. 11.2 Basic pharmacophore features (a) and (b) show the superimposed lead molecule with the pharmacophore model

corresponding sites. For example, a hydrophobic feature corresponds to hydrophobic protein side chains in the cavity; and a hydrogen-bond acceptor feature has a hydrogen bond-donating counterpart in the protein (Langer and Hoffmann 2006; Wolber and Langer 2005). A pharmacophore model was built from a collection of known partial agonists, and it was validated with a newly discovered partial agonist. Pharmacophore models are frequently employed in virtual screening processes to find a potential lead molecule. For example, Mustata et al. developed a potential lead molecule against Myc-Max via a pharmacophore model generated using known disruptors. In another study, Petersen et al. identified a novel PPAR γ partial agonist using a pharmacophore model (Mustata et al. 2009; Petersen et al. 2011). Pharmacophore-based screening processes match all the atoms or functional groups and the geometric relations between them to the pharmacophore in the query. Basically, two steps are involved in a pharmacophore-based search: in the first step, software checks all the lead molecules as to whether it has the atom type or functional groups required by the pharmacophore; then it checks whether the spatial arrangement of this element matches the query.

2D pharmacophore searching

Searching of a 2D database to find potential lead molecules is one of the crucial steps in drug discovery. Pharmacophore-based virtual screening has been used for the identification of potential hit molecules in drug development process. This approach can be used to screen virtually millions of compounds for hit identification. However, problems can arise from substructure when the number of compounds identified reaches into the thousands. This problem can be rectified by collecting these compounds based on similarity between compound in the database and in the query (Vyas et al. 2008). The structure activity relationship of these compounds can be generated in these processes even before synthetic paths are made for lead optimization based on the biochemical data (Enyedy et al. 2003). Beyond structure similarity, activity similarity has also been the subject of several studies.

3D pharmacophore searching

3D pharmacophore modeling acts as an efficient filter for virtual screening of large compound libraries due to its simplicity and abstract nature. The computational complexity of the hit identification process in virtual screening is greatly reduced by the sparse pharmacophoric representation of ligand-protein interaction. The generation of a query pharmacophore model that specifies the type and geometric constraints of the chemical feature is the first step in a typical pharmacophore-based virtual screening experiment. Both ligand-based and structure-based models can be created and used separately or in combination via parallel virtual screening. Ligand-based screening is generally used when crystallographic solution structure or modeled structure is lacking. Both ligand-based and structure-based pharmacophores

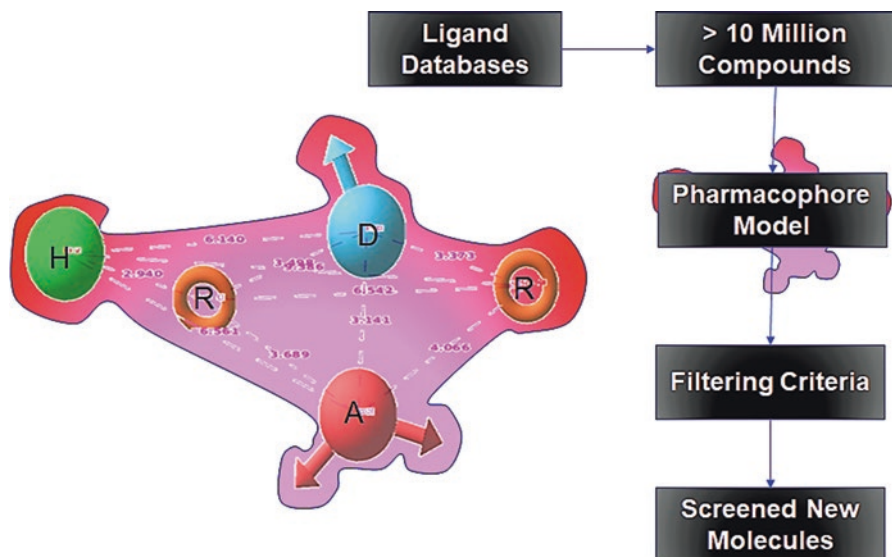


Fig. 11.3 Working method of 3D pharmacophore searching against small molecule databases

significantly screen the potential novel compounds with similar features and activity that can bind the same site of the proteins based on the features of the known compounds as mentioned in the Fig. 11.3. Several software products such as Catalyst, Sybyl/Unity, MOE and Phase are widely used methods for ligand-based pharmacophore building. Structure-based methods in pharmacophore modeling have gained significant interest in recent years, and several new approaches have been described, including the application of pharmacophore fingerprints for lead identification (Karnachi and Kulkarni 2006; Langer and Hoffmann 2006).

Fingerprinting

Pharmacophore fingerprints are defined as the binary encoded information about the presence or absence of pharmacophore features such as the centers and the three inter-center distances between them. By default, the seven center types that are probably the most important for the ligand-receptor interactions defined are: hydrogen-bond acceptor (A) and donor (D), groups with formal negative (N) and positive (P) charges, hydrophobic (H) and aromatic ring (R), and distance in a single molecule or a compound collection. Generally, fingerprinting focuses two or four-point pharmacophore fingerprints, but a larger number can be used, and utilization of up to nine pharmacophores has been described (Martin and Hoeffel 2000; Cato 2000). Traditionally, pharmacophore triplets are a widely used method and are most effective in terms of information content versus complexity; they are usually

generated for a set of compounds instead of an individual one. For each compound, the flow energy conformer is calculated by every possible combination of three or four features and used to set the corresponding bit in the fingerprint. The obtained fingerprint is termed the 'union key' (Cato 2000). The generation of pharmacophore fingerprints for proteins with known binding site can be calculated from complementary site-points in the binding site. Methods such as ChemProtein module of Chem-X or the GRID program are often used for generation of site-points using a variety of probe atoms (Mason and Cheney 2000; Mason and Beno 2000). Chem-X is one of the most popular software packages. The fingerprinting in this module is defined according to all the potential pharmacophores that can be present in some low-energy conformer of the molecules. Another method, the Oriented Substituent Pharmacophore PPropErtY space (OSPPREYS) approach, introduced by Martin and Hoeffel, is aimed towards better representation of diversity and similarity in combinatorial libraries in the 3D pharmacophore space (Martin and Hoeffel 2000). Pharmacophore fingerprint methods have a wide range of applications; they can be used to measure molecular similarity (Willett 2006), to design libraries, to assess their diversity and to search them for novel active compounds (Beno and Mason 2001).

11.3.2.2 QSAR Modeling

Quantitative structure-activity relationship (QSAR) is a highly popular approach for ligand-based drug designing. This method significantly quantifies the correlation between the chemical structures of a series of compounds and a chemical or biological process. The basic mechanism underlying the QSAR method is that structurally similar molecules or those compounds having similar physiochemical properties yield similar activity (Akamatsu 2002; Verma and Hansch 2009). The first step of developing a QSAR model is identification of a group of chemical entities or potential lead molecules which show the desired biological activity. The developed QSAR model is then used to optimize the active compounds to maximize the relevant activity, and then it is tested experimentally for the desired activity. Mainly, four steps are involved in QSAR model prediction (Fig. 11.4). In the first step, potential lead molecules are identified with experimentally measured values of the desired biological activity. In second step, molecular descriptors associated with various structural and physiochemical properties of the molecules are identified, and in the third step, the correlation between molecular description and biological activity is discovered to explain the variation in activity in the dataset. Finally, the statistical stability and predictive power of the QSAR model is tested.

In the classical or the 2D QSAR method, various electronic, hydrophobic and steric features are correlated with biological activity for a congeneric series of compounds (Acharya et al. 2011). In the classical method the molecular descriptors used for correlation with activity are mostly representative of fragments of the parent molecule. The major advantage of the classical method is that it is more effective for a congeneric series of molecules; however, the fragment-based descriptors are

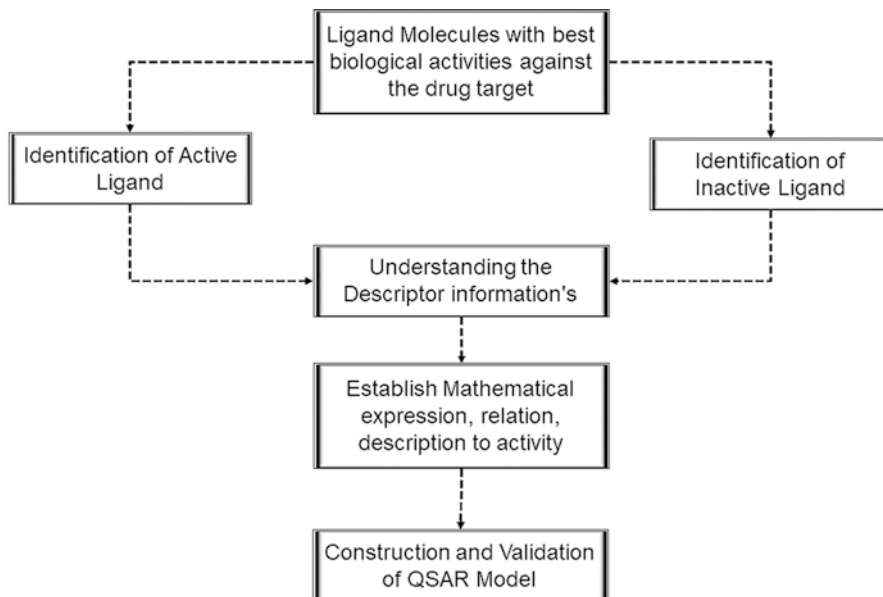


Fig. 11.4 Working method of QSAR modeling and predictions

usually inadequate to capture 3D conformational features of the crucial step for its activity (Winkler 2002; Bernard et al. 2005; González et al. 2009). To describe the 3D features of molecules the new 3D QSAR method was developed in which various geometric, physical characteristics and quantum chemical descriptors are used to describe the 3D features of a molecule; those descriptors are then combined to create a pharmacophore that can explain the biological activity of ligands (Chang and Swaan 2006). Then, a developed pharmacophore model is subjected to stability and statistical analysis to obtain the final 3D QSAR model. Several techniques including CoMFA, CoMSIA and catalyst are currently used for this drug designing approach.

11.3.2.3 CoMFA

Comparative molecular field analysis (CoMFA) is one of the 3D QSAR techniques mainly used to describe structure activity relationships in a quantitative manner. In this method a set of molecules is identified and aligned based on their 3D structures on a 3D grid and the values of steric and electrostatic potential energies are calculated at each grid point. The identified lead molecules should have a similar binding mode (identical binding) to the same kind of receptor. In the next step, a certain group of molecules is selected as a training set to derive the CoMFA model. The residual molecules are considered a test set, which independently proves the

validity of the derived models. A pharmacophore hypothesis of this method is generated to orient the superposition of all molecules and to afford a rational and consistent alignment. It calculates the values in each grid point, i.e., the energy of molecules via a carbon atom, a positively or negatively charged atom, a hydrogen-bond donor or acceptor, or a lipophilic probe, correlating these values with the biological activity. Principle component analysis (PCA) and partial least squares (PLS) are the most widely used methods for development of pharmacophore in CoMFA. The developed model is then tested for statistical significance and robustness (Gohda et al. 2000; Akamatsu 2002; Yasuo et al. 2009). The result of this approach can be represented as counter maps that indicate points of the lattice where variations in field values are related to variations in biological activity. These maps can be used to estimate the regions of molecules where some types of interactions have a favorable or unfavorable influence on the biological activity. Recently, several modifications have been described which significantly are used as alternatives to CoMFA (Sen et al. 2012).

11.3.2.4 Comparative Molecular Similarity Indices Analysis (CoMSIA)

CoMSIA is another 3D QSAR method, introduced by Klebe and his coworkers (1994) based on the calculation of similarity indices between the alignment's molecules and a common probe atom placed at the interaction grid. Most of the features of CoMSIA are similar to CoMFA; however, there are differences: The molecular field expression includes five different properties such as hydrophobic, hydrogen-bond donor and acceptor terms in addition to steric and coulombic contributions, and it calculates similarity indices instead of interaction energies by comparing each ligand molecule with a common probe. The statical evaluation of these field properties are correlated with the biological property by PLS analysis, but the counter maps are more contiguous and easier to interpret in CoMSIA because they are no cut-off values (Flower 2002; Klebe et al. 1994). To calculate the similarity indices, a Gaussian-type functional form is used to describe steric, electrostatic and hydrophobic compounds of the energy function, and it avoids using the arbitrary cut-off value for the energy calculation (Acharya et al. 2011). The Gaussian function also provides a smoother description of potential energy in regions near the van der Waals radius atom (Klebe et al. 1994).

11.4 Virtual Screening (VS) for Lead Discovery

The discovery of novel leads with potential interaction with targets is one of the important steps in drug discovery. This approach is conventionally achieved by wet-lab high throughput screening (HTS) in many pharmaceutical industries, but due to the high cost and low hit rate, the alternative method is developed with broad application of the cheaper and faster screening of *in silico* approaches (Clark 2008;

Ripphausen et al. 2010). Alternative virtual screening (VS) uses computational power to test a large set of small molecules in a limited time at low cost. VS is a stepwise process with a cascade of sequential filters able to narrow down and choose a set of lead-like hits with potential biological activity against intended drug targets. It can be broadly classified into two categories, ligand-based virtual screening (LBVS) and structure-based virtual screening (SBVS). A broad range of computational techniques that can be applied in this process includes drug likeness screening, counting scheme, functional group filters, topological drug classification, pharmacophore points filter and pharmacophore-based virtual screening. Molecular docking is a computationally intensive method that has been applied to very large databases of chemical structures.

Protein-ligand docking has become one of the widely used tools in modern drug discovery approaches to predict the most likely binding mode of small molecules at a particular receptor to explore specific interactions that may be formed and to estimate ligand-binding affinity. A number of protein-ligand methods are available to date, from academic groups to commercial software vendors. The binding free energy between protein and ligand molecules employs rather heuristic terms and these functions are referred as scoring function. Scoring functions is a very important step, which includes protein preparation, ligand database preparation, docking calculation and post processing. Basically, the scoring process composed of three different aspects relevant to docking and design. The first aspect is the ranking of the conformations generated by the docking research for one ligand interacting with a given protein; this aspect is crucial for detecting the binding mode that best approximates the experimentally observed situation. The second aspect is ranking the different ligands with respect to binding to one protein; that is, prioritizing ligands according to their affinity, which is essential in virtual screening and the third aspect is ranking one or different ligands with respect to their binding affinity to different compounds which is essential for the consideration of selectivity and specificity of ligands (Leach and Hann 2000; Lewis et al. 2000). The amount and quality of available information on the target protein is one of the key factors in designing a virtual screening project (Klebe 2006). The information on the coordinates of the features of the 3D structure of the known targets is valuable data and can be used to improve the quality of the results. The predictions of 3D structure of biomolecules are obtained by the three exemplary methods of NMR spectroscopy, X-ray crystallography and homology modeling. Currently, PDB contains more than 70,000 experimentally solved 3D structures of proteins that can be used as targets in VS and in homology modeling.

11.4.1 Protein Modeling

Proteins are the fundamental structural elements in living organisms; they act as catalytic agents, signal transmitters, transporters and molecular machines in cells (Nelson et al. 2008). Mostly, most the proteins are not functions individually; they must

interact with other molecules to carry out their cellular roles, if any alteration in the protein interface leads to a pathological condition. Hence, the protein interface may be used as potential targets for rational drug designing approaches (Rask-Andersen et al. 2011; Jubb et al. 2015). Many experimental methods including NMR and X-ray crystallography have been used to identify and characterize the protein-protein interface at the level of individual atoms and residues, and various mass spectrometry-based approaches such as chemical cross-linking and hydrogen/deuterium exchange have been used, which typically report the location of interface at lower resolution (Hoofnagle et al. 2003; Kaveti and Engen 2006; Gobl et al. 2014; Shi 2014). Though these experiments provide valuable knowledge of the protein recognition mechanism, technical challenges such as expressing and purifying aggregation-prone protein samples, obtaining high quality crystals and protein size constraints are both labor-intensive and time-consuming. Hence, in the absence of an experimentally determined structure, an alternative computational approach such as comparative or homology modeling is used to predict the 3D model of proteins related to at least one known protein structure. The model gives the 3D structure based on its alignment to one or more known protein structures (Pieper et al. 2002).

11.4.1.1 Homology Modeling

Comparative or homology modeling is one the easiest methods among the three-structure prediction approach. In homology modeling, the structure process consists of fold assignment, target-template alignment, model building and model evaluation. There are several computer programs and web servers that automate the comparative modeling of proteins. Generally, the 3D structure of proteins can be achieved by several different approaches and is strongly dependent on the sequence identity (SI) or the percentage of identical amino acid residues present among the target sequence and their templates (Santos Filho and Alencastro 2003). Ab initio is the another method used for prediction of 3D structure of protein and mostly suitable, when there is no suitable template with significant sequential identity to the target sequence. If the sequence identity between target and template protein is above 30%, comparative or homology modeling is a suitable approach (Baker and Sali 2001; D'Alfonso et al. 2001). In practice, homology modeling consists of the seven important steps, which are template recognition and initial alignment, alignment correction, backbone generation, loop modeling, side chain modeling, model optimization and model validation (Peitsch et al. 2000; Westbrook et al. 2002; Orengo et al. 2002; Lo Conte et al. 2002).

Template selection is the initial step in safe homology modeling. The percentage of sequence identity between the sequence of interest (query) and a possible template can be detected by different software. The template model can be found using the query sequence from a database such as the protein data bank (Westbrook et al. 2002), SCOP (Lo Conte et al. 2002) and CATH (Orengo et al. 2002). Three main classes of protein comparison methods are involved in fold identification. Initially, the target sequence is subjected to pairwise sequence alignment with each database

sequence independently to find its homologous sequence (Fiser 2010). Computational programs such as BLAST (Schäffer et al. 2001), FASTA (Srivastava et al. 2009) and CDART are frequently used methods for searching the related protein sequence and structure of the template. The second class of method employed is a multiple sequence alignment profile to compare the sequence using profile analysis profile-profile comparisons, Hidden Markov models and intermediate sequence search (Rychlewski et al. 2000 Yona and Levitt 2002; Zhou and Zhou 2005; Fiser 2010). SAM and PSI-BLAST (Karplus et al. 2003) are the most often used programs for this approach. The third class of method is also a pairwise alignment method, where the target sequence adopts any one of the many known 3D -folds predicted by an optimization of the alignment with respect to a structure-dependent scoring function independently for each sequence-structure pair; i.e., the target sequence is threaded through a library of 3D-folds (Kelley et al. 2000).

The next important step is a sequence alignment between the target and template structure. Mostly, fold assignment methods are widely used in this process and it is agreed that profile-based alignment produce better quality models than sequence-based alignments. In addition, HMM-based alignments produce higher quality model than PSSM-based method alignments produced by PSI-BLAST (Yan et al. 2013). A pairwise comparison of protein sequence and protein structure is matched against a library of 3D profiles, this method is also known as fold assignment. Once a list of potential templates is obtained using different searching methods, it is necessary to select a potential template more appropriate for the modeling problem. The selection of highest sequence similarity is the simplest template selection rule for modeling the protein (Retief 2000). After the selection of a potential template, a suitable method is used to construct the 3D model from template and alignments. Generally rigid-body assembly, segment matching, spatial restraint and artificial evolution are used for model building. This rigid-body assembly model relies on the natural dissection of the protein into conserved core regions, variable loops that connect them and side chains that decorate the backbone. The segment matching based on the construction of a model by using a subset of atomic positions from template structure and by identifying and assembling short. All atom segments in the model that fit the guiding positions can evaluated by scanning all the known protein structures (Xiang 2006). Several programs are available for modeling the query sequence. Andrej et al. developed MODELLER, which remains one of the most widely used comparative modeling methods. The spatial restraints approach is implemented in MODELLER. It starts by aligning the target sequence with the related known 3D structure, and the output obtained by this method contains a molecular structure that includes main chain and side chain non-hydrogen atoms similar to the known structure. In addition to MODELLER, other tools including Swiss Model, RAMP, PrISM, COMPOSER, CONGEN+2 and DISGEO/Co-sensus are often used in homology modeling (Schwede et al. 2003; Vyas et al. 2012). This homology modeling approach is described in several available programs, both in the commercial and public arena.

Model evaluation and validation is necessary to construct a model with good stereochemistry; the most important factor in the assessment of constructed mod-

els is the scoring function, and programs evaluate the location of each residue in a model with respect to the expected environment as found in the high-resolution X-ray structure. The stereochemistry of the modeled protein can be verified by the analysis of parameters like bond lengths and angles, torsional angles and chirality of residues using PROCHECK (Laskowski et al. 1993), WHATCHECK (Hooft et al. 1996), PROSA (Sippl 1993) and Molprobity (Davis et al. 2007; Chen et al. 2010). The reliability of a predicted model is also subject to a check of other parameters such as planarity of the peptide bond, chirality of the C α , bond length and angles in the main chain, the planarity of aromatic system, the inner backing of globular proteins and the elements of the secondary structure, hydrophobic and hydrophilic residues of the predicted protein structure (Schwartz et al. 2001).

11.4.1.2 Threading

In comparative modeling it has been observed that the careful alignment of the corresponding amino acid residues of the unknown proteins with a similar sequence, often closely related homologues, tend to have similar 3D structure with similar conformations. When no sequences are clearly related to the modeling target, the alternative method of threading is employed to predict structure via fold recognition. Protein threading, i.e., sequence-structure alignment, is a promising template based on fold recognition, which identifies a suitable fold from a structure library for the query sequence and provides an alignment between the query protein and the fold (Shan et al. 2001). The word 'threading' was first coined by Jones et al. (1992); the original term was 'optimal sequence threading,' later it shortened 'threading.' In this method, the query sequence is threaded onto the backbones of the template structures. Threading requires four basic components: (1) a template library representing the 3D protein structure to be used as the template; (2) an energy function to describe the fitness of any template; (3) a threading algorithm to search for the lowest energy among the possible alignments for a given sequence-template pair; (4) a criterion to estimate the confidence level of the predicted structure. The threading method is further classified into two broad categories, singleton threading, in which the threading considers only the preference of amino acids in the query sequence at single sites of the templates; and a category that uses the preference on pairs of amino acids in the query sequence within contact distance when they are aligned to a given structure. Singleton threading constructs a 1D structure profile for each amino acid residue position in a template using the 3D structural information, such as secondary structure type, degree of environmental polarity and fraction of residue surface accessible to solvent. Typically in threading, it is assumed that the backbones of the structures are rigid and only the amino acid side chains of the query and the template are different. Threading exploits the fact that proteins with different functions can possess a similar structure even though they may have little to no sequence similarity. Loopp and therader are software (learning, observing and outputting protein patterns (Tobi and Elber 2000; Meller and Elber 2001; Teodorescu et al. 2004) can be used for structure prediction via fold recognition. Both loop and

threader rely on similar strategies, yet they use different energy and scoring functions to generate possible alignments with feasible templates. THREADER uses solved protein structure as a scaffold on which to place the target protein sequence and analyze secondary structure information about the target sequence used to force alignment between predicted secondary structures of the target. It uses a set of basic knowledge-based potentials such as statistical data compiled from known protein structure and pairwise pseudo-energy to indicate misfolded proteins.

The strategy of LOOPP is similar to THREADER, but it differs in its implementation of an empirical energy function and its scoring method. The most notable aspect of LOOPP is its extensive parameterization, which is based on the structure from the protein data bank (PDB) and a database of close to five million decoy structures (Berman et al. 2000; Tobi and Elber 2000). Three novel implementations of common protocol—the pairwise contact model, gap penalties and Z-scores—differentiate LOOPP from other threading methodologies. It creates a new pairwise interaction model (empirical energy function) acting as the key to devising a truly novel threading algorithm. Basically, two main types of empirical energy functions exist in this method: (1) those that pairwise residues contacts for residues within a specified distance of one another; (2) those based on the environment of an amino acid residue at a point in the structural lattice (Meller and Elber 2001). Several threading programs including the NCBI threading package (Bryant and Lawrence 1993), PROFIT (Sippl and Weitckus 1992), PROSPECT (Xu et al. 1998), CASP-3 (CASP 1999), TOPITS (Rost and Sander 1995) and SAS (Milburn et al. 1998) are used for singleton and pairwise interactions. The NCBI threading package provides a good statistical assessment of a threading result, and recently CASP-3 was used as a top performer in threading with pairwise interactions.

11.4.1.3 Ab Initio Method

Ab initio method is one of the modeling technique often used for structure prediction when the sequence of the query proteins has either no or a low amount of similarity and in this method the query protein is folded with a random conformation. The ab initio method is based on the thermodynamic hypothesis proposed by Anfinsen, according to which the native structure corresponds to the global free energy minimum under a given set of conditions (Floudas et al. 2006). Basically, the ab initio category has two subclasses, fragment-based and biophysics-based methods. These are often called, respectively, first-principles methods that employ database information and first-principles methods without database information (Floudas 2007). All types of proposed approaches rely on minimization of the energy function over the conformation parameters. The typical method has four basic steps for finding the conformation with the lowest energy: (1) start with an unfolded/arbitrarily folded conformation; (2) generate alternative conformations using some heuristics; (3) estimate their corresponding energy; and (4) again, generate the alternative conformation until the final criterion is reached. Parameters like energy function accuracy, search algorithm efficiency and selection of the

best models play a crucial role in the structure prediction *ab initio* method. In the basic modeling, folding process, and quantum mechanics is used to model and estimate the interactions of atoms. Currently, a high performance computing facilities force field (FF) or energy function are employed to express a variety of atomic interactions such as van der Waals, torsion angles, electrostatics and bond length. Energy functions are usually associated with the search procedure to locate the conformation that has the minimum energy function value. The most popular optimization methods are molecular dynamics and Monte Carlo simulation (Adcock and McCammon 2006). The category of *ab initio* prediction with database information focuses only on predicting as accurately as possible a protein's final configuration. In this approach, the structure prediction starts with the primary amino acid sequence, which is searched for different conformations, leading to the prediction of native folds. After the folds have been recognized and predicted, the model assessment is performed to verify the quality of the structure. ROSETTA and I-TASSER are widely used fragment-based enhanced methodologies for *ab initio* structure prediction of a protein. TASSER was initially created in 2004 by Zhang and Skolnick (2004), and later the enhanced versions Chunk-TASSER (Zhou and Skolnick 2007) and I-TASSER were developed in structure prediction (Wu et al. 2007). TASSER is a hierarchical approach that encompasses three phases, thus its name: threading/assembly/refinement ("TASSER"). The first step, threading, is an iterative sequence-structure alignment algorithm that uses the program PROSPECTOR_3 (Skolnick et al. 2004). The second step, assembly, uses parallel hydrophobic Monte Carlo sampling by rearranging the template fragments (Zhang et al. 2005). The final step, refinement, is performed using a clustering program called SPICKER (Zhang and Skolnick 2004), and the full atom optimization is conducted using the CHARMM22 force field. ROSETTA prediction involves the identification of small fragments from the structural databases consistent with a local sequence preference.

11.4.1.4 Protein Validation Server

Protein structure has proved to be a crucial piece of information for biochemical research. From the millions of currently sequenced proteins only a small fraction is experimentally solved for structure, and the only feasible way to bridge the gap between sequence and structure data is computational modeling. Unlike experimental structure, the accuracy of a computationally modeled structure can be estimated by a broad range of the accuracy spectrum. Over the past two decades, several approaches have been developed to analyze the accuracy of the protein structure and model. They use stereochemistry checks, molecular mechanics energy-based functions and statistical potentials to tackle problems. Typically, features like molecular environment, hydrogen bonding, secondary structure, solvent exposure, planarity, chirality, phi/psi preference, chi angles, non-bonded contact distances, unsatisfied donors/acceptors, pairwise residue interaction and molecular packing are analyzed in these approaches. A good quality protein should resemble a native protein, with

spatial features of the residues complying with empirically characterized constraints on torsional angles captured in Ramachandran plots (Ramachandran et al. 1963). PROCHECK (Laskowski et al. 1993) and MolProbity (Chen et al. 2010) are widely used programs for determining whether a modeled protein structure has native-like features. Traditionally, several studies have examined protein structures using an all atom-based description. Ramachandran's plot with backbone dihedral angle ϕ (N-C α) and ψ (C α -C) is a representative microscopic description of the protein structure. Dihedral angle prediction has several applications in protein structure prediction; which include secondary structure prediction (Rost 2001; Wood and Hirst 2005; Kountouris and Hirst 2009), generation of multiple alignments (Huang and Zou 2006a, b; Miao et al. 2008), identification of protein fold (Karchin et al. 2003; Zhang et al. 2008) and fragment-free tertiary structure prediction (Faraggi et al. 2009). Quality assessment is an important step in the modeling process, wherein processes like template level, alignment level, selected fragment level and structural level error are analyzed. A template structure for a target sequence is identified by considering the significance of the score that indicates the fitness of the target to the template. In principle, most frequently the statistical significance of a raw score is considered as either in the form of the E-value (homology search) or the Z-score (used in threading algorithms). Z-score are calculated as measured value minus population mean, divided by the standard deviation of the population. So, a Z-score is negative if the value of X is less than the mean, and it is positive if the measured value is greater than the mean value. WHAT IF uses this criterion a lot to calculate Z-score. The Z-score provides basic information about the root mean square of a population with a Z value and it should be 1.0.

11.4.2 Protein and Ligand Preparation

The success of the various drug designing approaches depends largely on whether reasonable starting structures are used for both the protein and the ligand. The protein structure that is retrieved from PDB (X-ray structure) consists of heavy atoms and may contain water molecules, cofactors, activators, ligands and metal ions as well as several protein subunits and does not have the information on bond orders, topologies. Because of the above structural issues, several protein preparation approaches have been developed (Sastry et al. 2013; Pitt et al. 2013). The determination of protonation states of the amino acid in protein molecules is the first crucial step in protein preparation. Several freely available software packages including PROPKA (Li et al. 2005), H++ (Anandakrishnan et al. 2012) and SPORES (ten Brink and Exner 2010) are widely used for determining the first step of the protein preparation. The next important step is to assign hydrogen atoms and optimize protein hydrogen bonds according to an optimal hydrogen bond network. PDB2PRO software is a widely used tool for these tasks (Dolinsky et al. 2007). The next step is assignment of partial charges, capping of residues, treating metals, filling missing loops and missing side chains and minimizing the protein structure to relieve steric

clashes; also, a crucial decision must be made regarding whether water molecule will be left in or removed from the binding site. To tackle the above mentioned challenging problems, freely available tools such as 3D-RISM (Kovalenko 2003; Young et al. 2007; Abel et al. 2008), SZMAP (Myrianthopoulos et al. 2016), JAWS (Michel et al. 2009) and WaterMap (Young et al. 2007; WaterMap, Schrödinger 2014) are utilized in commercial software (Jorgensen and Tirado-Rives 2005; SZMAP Software Inc.). In the case of a co-crystallized protein structure with substrates and cofactors, Protein Preparation Wizard of Maestro (Maestro, Schrödinger, LLC) is used to assign proper bond orders and generate accessible tautomer and ionization states prior to virtual screening.

The selection of the type of ligand molecule chosen for docking is another important step in virtual screening. The type can be obtained from various databases like ZINC or pubchem, or it can be sketched by means of ChemsSketch or ChemDraw tools (Dias and de Azevedo 2008). A wide variety of small molecule databases are available for virtual screening-based drug designing. Many of them are free and possess desirable characteristic lead molecules. ZINC is a public access database, contains number of commercially available compound that are mostly developed in the pharmaceutical chemistry department at the University of California, San Francisco. NCI is an another open database developed by the Developmental Therapeutics program of the National Cancer Institute, NIH; it currently contains over 250,000 molecules from both organic synthesis and natural sources. ASINEX is a regularly updated commercial database currently containing 600,000 screening compounds, 27,000 macrocycles, 23, 000 fragments and 7000 building blocks. SPECS is a monthly updated database containing more than 240,000 novel drugs—drug-like small molecules obtained from an academic research institute. MAYBRIDGE is one of the widely used commercial databases containing a screening hit discovery collection more than 53,000 and offering a fragment library of 30,000. CHEMBRIDGE encompasses one million drug-like and lead-like molecules in two non-overlapping collections of respectively 460,000 and 620,000 compounds. After selection of potential lead molecules, it should be preprocessed before docking. There are several thousand small molecules in a ligand database, so one must avoid performing manual steps in data preparation. Typically, information on available ligands is stored in 2D form in databases, serving as a data repository. Currently, several thousand small molecules are available in various databases; Table 11.2 shows widely used small molecule repositories. The 2D structure retrieved from these repositories of atom and bond types must be checked and corrected; protonation states and charges have to be assigned. Then, 3D structures must be converted for calculating ligand conformation like rotational barriers or side-chain rotamers allowed. In addition, protein-ligand interactions including site-points that guarantee proper hydrogen-bonding directionality must be assigned (Claussen et al. 2001). LigPrep is the most widely used module for ligand preparation implemented in Schrödinger (LigPrep, Schrödinger 2011). In this module, ionization/tautomeric states are generated with either a pair of fast rule-based programs or with Epik, which is based on the more accurate Hammett and Taft methodologies (Shelley et al. 2007; Epik, Schrödinger 2011).

11.4.3 Active Site Prediction

Binding site prediction and characterization of small molecules is more important for drug discovery. Often, possible binding sites for potential small molecules are known for co-crystal structures of the target or a closely related protein with natural ligand molecules. Recently, Hajduk and coworkers used heteronuclear-NMR-based screening to identify and characterize the ligand binding site on a protein surface (Hajduk et al. 2005). By screening a large number of lead-like molecules against 23 target proteins, the results revealed that 90% of the ligand molecules bonded to specific locations on the protein surface, depicting that certain properties of small-molecule binding sites should be common to general molecular recognition. Mostly computational studies have been used to predict the binding site for an unknown or if a new binding site is to be identified, e.g., allosteric molecules. Computational methods like Q-SITEFINDER, POCKET (Levitt and Banaszak 1992), SURFNET (Laskowski 1995), APROPOS (Peters et al. 1996), LIGSITE (Hendlich et al. 1997), CAST, CASTp (Binkowski et al. 2003) and PASS (Brady and Stouten 2000) are often used for binding site prediction. Computational methods for the identification of a binding site can be categorized into three major classes: (1) geometric algorithms to find the shape concave invagination in the protein molecules; (2) energies-based method; and (3) method considering dynamic of protein structures. Geometric algorithms find a putative binding site through detection of cavities on a protein surface. In this algorithm, grids are used to describe the molecular surface of the protein, and the boundary of the binding site is determined by rolling a spherical probe over the grid surface. This kind of algorithm is used in SURFNET, LIGSITE and POCKET, where spheres are placed between all pairs of target atoms and then the radius of sphere is reduced until each sphere contains only a pair of atoms. An et al. (2005) developed the Pocket Finder algorithm and expanded the geometric method by counteracting a smoothed van der Waals potential for the target protein to identify candidate ligand binding sites. The new technique of Sitemap, developed by Schrödinger, Inc., identifies the known binding site in >96% of cases by linking together site-points that contribute to tight protein ligand binding. Sitemap provides quantitative and geographical information that helps guide efforts to modify ligand structure to enhance properties (Halgren 2007; Halgren 2009) (Table 11.1).

11.4.4 Molecular Docking

In a modern drug discovery approach, protein-ligand and protein-protein interaction mechanisms play a significant role in predicting orientation of the ligand when it is bound to a protein receptor or enzyme using shape and electrostatic interaction to quantify it. Molecular docking is an attractive scaffold for understanding protein-ligand interaction in a rational drug design and drug discovery; in the mechanistic study a molecule is placed into the binding site of the receptor molecules mainly in

Table 11.1 Widely used small molecule repositories with basic information about the class of the compounds and their size

Database	Type	Size	Citations
PubChem	Biologic activities of small molecules	40,000,000	Wheeler et al. (2006)
Accelrys Available Chemicals Directory (ACD)	Consolidated catalog from major chemical suppliers	7,000,000	Accelrys (2012)
PDBChem	Ligands and small molecules referred in PDB	14,572	Dimitropoulos et al. (2006)
Zinc	Annotated commercially available compounds	21,000,000	Irwin and Shoichet (2005)
LIGAND	Chemical compounds with target and reactions data	16,838	Goto et al. (2002)
DrugBank	Detailed drug data with comprehensive drug target information	6711	Wishart et al. (2006)
ChemDB	Annotated commercially available molecules	5,000,000	Chen et al. (2005, 2007)
WOMBAT Data base	Bioactivity data for compounds reported in medicinal chemistry journals	331,872	Ekins et al. (2007); Hristozov et al. (2007)
MDDR (MDL Drug Data Report)	Drugs under development or released; descriptions of therapeutic	180,000	Hristozov et al. (2007)
3D MIND	molecules with target interaction and tumor cell line screen data	100,000	Mandal et al. (2009)

a non-covalent fashion to form a stable complex of potential efficacy and more specificity (Rohs et al. 2005; Guedes et al. 2014). The information obtained from a docking study can be used to study the binding energy, free energy and stability of drug-biomolecular complexes with optimized conformation and with the intention of possessing less binding free energy. The basic two steps involved in molecular docking, usually related to sampling methods and scoring schemes, are (1) prediction of ligand conformation and position and orientation within these sites (usually referred as pose) and (2) assessment of binding affinity (Fig. 11.5).

Most of the docking tools employed the searching algorithms including genetic algorithms (GA), Monte Carlo algorithms, molecular dynamics algorithms and conformational search algorithms in the molecular docking method. Conformational search algorithms perform in the docking approach by applying systematic and stochastic search methods (Agrafiotis et al. 2007; Yuriev et al. 2011). The basic methodology of molecular docking falls into three categories: induced fit docking, where both ligand and receptor molecules are flexible; rigid body docking, where ligand and receptor molecules are rigid; and flexible docking method, in which it is also the case that both interacting molecules are flexible (Meng et al. 2011). The molecular docking process involves the following major steps: (1) Preparation of protein—

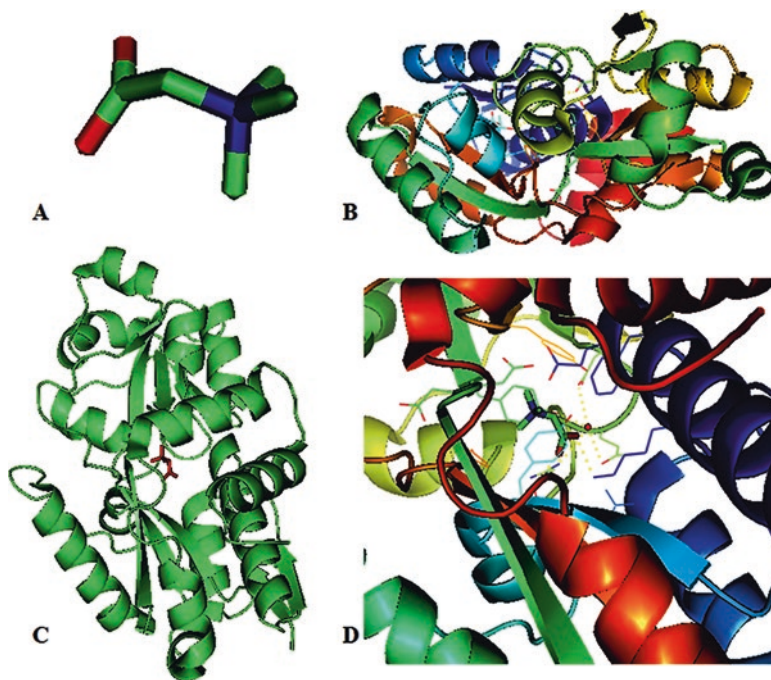


Fig. 11.5 Basic steps involved in molecular docking approach. (a) Three-dimensional structure of lead molecules; (b) three-dimensional structure of the protein; (c) ligand is docked into the binding site of the protein; (d) binding affinity and interactions of ligand molecules with protein

before docking, the 3D structure of the receptor molecule (retrieved from either PDB or molecular modeling) should be pre-processed by stabilizing the charges, filling the missing residues, and generating and removing free water molecules from the cavity. (2) Active site prediction—the binding site of the receptor molecules should be predicted in this step; the water molecules and hetero atoms are removed. (3) Ligand preparation—the small molecules can be retrieved from small molecule databases while choosing the ligand molecules; the LIPINSKY'S RULE OF 5 should be utilized. (4) Docking—the final step, where the ligand is docked against the protein and the interactions are analyzed; the scoring function finds the docking scores based on best pose of docked ligands complex. Over the last two decades, approximately 60 different docking tools and programs have been developed for both academic and commercial use, including DOCK (Venkatachalam et al. 2003), Auto Dock (Österberg et al. 2002), FlexX (Rarey et al. 1996), Surflex (Jain 2003), GOLD (Jones et al. 1997), ICM (Schapira et al. 2003), Glide (Friesner et al. 2004), Cdocker, LigandFit (Venkatachalam et al. 2003), MCDock, FRED (McGann et al. 2003), MOE-Dock (Corbeil et al. 2012), LeDock (Zhao and Caffisch 2013), AutoDock Vina (Trott and Olson 2010), Dock (Ruiz-Carmona et al. 2014) and UCSF Dock (Allen et al. 2015). Table 11.2 shows the basic information on the currently used docking tools and scoring functions.

Table 11.2 Basic characteristics of widely used docking tools

S. No	Docking programs	Docking approach	Scoring function
1	DOCK	Shape-fitting (sphere sets)	Chem Score, GB/SA solvation scoring, other
2	Auto Dock	Genetic algorithm Lamarckian genetic algorithm simulated annealing	Auto Dock (force-field methods)
3	Flex X	Incremental construction	FlexX Score, PLP, Screen Score, Drug Score
4	FRED	Shape-fitting (Gaussian)	Screen Score, PLP, Gaussian shape score, user-defined
5	Glide	Monte Carlo sampling	Glide Score, Glide Comp
6	GOLD	Genetic algorithm	Gold Score, Chem Score user defined
7	Ligand Fit	Monte Carlo sampling	Lig Score, PLP, PMF
8	Surflex	Surflex-Dock search algorithm	Bohm's scoring function
9	ICM (Internal Coordinate Modelling)	Monte Carlo minimization	Virtual library screening scoring function
10	MVD (Molegro Virtual Docker)	Evolutionary algorithm	MolDock score
11	FITTED (Flexibility Induced Through Targeted Evolutionary Description)	Genetic algorithm potential of mean force	(PMF), Drug Score
12	GlamDock	Monte Carlo method	ChillScore
13	vLifeDock	Genetic algorithm	PLP score, XCscore
14	iGEMDOCK	Genetic algorithm	Empirical scoring function

11.4.5 Scoring Methods

Molecular docking approaches use scoring functions to calculate the binding energies of the predicted ligand-receptor complexes. Scoring function is a key element of a protein-ligand docking algorithm, determining the accuracy of the algorithms (Gohlke and Klebe 2001; Schulz-Gasch and Stahl 2004; Jain 2006; Rajamani and Good 2007; Gilson and Zhou 2007). Speed and accuracy are the important aspects basic to a scoring function. Several scoring functions have been used mainly to delineate correct poses from incorrect poses, or binders from inactive compounds within a reasonable computation time. Overall, scoring functions can be divided in the three categories of as force field-based, empirical-based and knowledge-based scoring functions (Kitchen et al. 2004). A classical force-field scoring function estimates the binding energy of a complex by calculating the sum of bonded terms

such as bond stretching, angle bending and dihedral variation, and non-bonded terms including electrostatic and van der Waals interactions. Electrostatics terms use a set of derived force-field parameters such as AMBER or CHARMM (Miller et al. 2017) and are calculated by a coulombic formulation. In addition to the above electrostatic terms, the force field-based scoring function also considers hydrogen bond, solvation and entropy contributions. The software such as DOCK (Kuntz et al. 1982), GLOD (Shoichet et al. 1993) and Auto Dock (Morris et al. 1998) offer users such functions. Force fields are mathematical expressions describing the dependence of energy of a system on the coordinates of its particles. The force-field scoring function shows some differences in the treatment of hydrogen bonds in terms of the energy function used, and it is further refined with other techniques such as linear interaction energy (Michel et al. 2006) and free-energy perturbation method (FEP) (Kollman 1993; Briggs et al. 1996) to improve accuracy in predicting binding energies. To reduce computational expense, alternative approaches such as Poisson-Boltzmann/surface area (PB/SA) and the generalized-Born/surface area (GB/SA) models were used to measure accuracy by treating water as a continuum dielectric medium (Rocchia et al. 2002; Liu and Zou 2006; Lyne et al. 2006; Thompson et al. 2008; Guimaraes and Cardozo 2008).

Empirical scoring function is another method to evaluate the types of physical events involved in the formation of the ligand-receptor complex. The binding energy of a complex is calculated by summing up a set of empirical energy terms including van der Waals energy, electrostatic energy, hydrogen bonding energy and desolvation terms. Each empirical energy term component is multiplied, and corresponding coefficients are determined by reproducing the binding affinity data of a training set of protein-ligand complexes with known three-dimensional structure using least squares fitting (Ballester and Mitchell 2010). Due to the simple energy terms and the nature of their fitting to known binding affinities of the training set, empirical scoring functions are computationally more efficient and faster than force-field-based methods. Molecular docking tools such as Surflex and FlexX and Glidescore (Friesner et al. 2004; Halgren et al. 2004), PLP (Gehlhaar et al. 1995; Gehlhaar et al. 1999), SYBYL/F-Score (Rarey et al. 1996), LigScore (Kramer et al. 1998) and Chemscore are some examples of programs that use empirical scoring functions (Jain 2003). Table 11.3 provides the widely used scoring functions implemented in the most frequently used molecular docking programs.

Table 11.3 Provides widely used empirical scoring functions in frequently used molecular docking tools

Force-field-based	Empirical	Knowledge-based
DOCK	Auto Dock	SMoG
Auto Dock	Gold Score	Drug Score
Glide Score	Chem Score	PMF_Score
ICM	X_Score	Motif Score
LigandFit	F_Score	RF_Score
Molegro Virtual Docker	Fresno	PESD_SVM
SYBYL_G-Score	SCORE	Pose Score

A third approach includes knowledge-based scoring functions that use statistical analysis, which are directly derived from the structural information in an experimentally determined protein-ligand complex to obtain interatomic contact frequencies and distance between the ligand and protein. Further, this approach uses pairwise energy potentials derived from a known ligand-receptor complex to obtain a general function (Huang et al. 2006). These potentials are constructed by considering the frequency distribution and the score is calculated by summing up of the individual interactions. Compared to force field and empirical scoring functions, knowledge-based scoring functions offer a good balance between accuracy and speed and are relatively robust and also enable the scoring process to be as fast as the empirical scoring function (Muegge 2006; Huang and Zou 2006a, b). Recently, a consensus scoring method has been developed which combines several scores to assess the docking conformation.

11.4.6 Molecular Dynamics (MD) Simulations

Molecular dynamics (MD) simulations of recent years play a critical role in computational drug discovery. Simulation studies can provide detail concerning individual particle motion as a function of time and use physics-based energy functions and explicit representations of atomic systems to model protein dynamics. MD simulation studies provide basic information to evaluate the stability and functions of the protein and to monitor the specific behaviors over the course of many simulations and provide information about target structure or properties unobtainable from static native structure. MD simulation was first developed in the late 70s when Alder and Wainwright performed it using a hard-sphere model. The first molecular simulation of BPTI was done in 1975 with a crude molecular mechanics potential for only 9.2 ps (Adcock and McCammon 2006). Molecular dynamics simulation mimics the physical motion of each atom in the macromolecule present in the actual environment. Each atom of a protein molecule can interact for a certain period of time, which helps in the computation of their trajectory in and around the protein molecules. A variety of properties such as free energy, kinetics measures and other macroscopic quantities of macromolecules can be calculated by using the trajectories. Several studies revealed the role of classical MD simulations to obtain different conformations of proteins and nucleic acids, including early attempts to stimulate spontaneously complex phenomena such as protein folding (Frenkel and Smit 2001). In recent research, MD simulation has been widely used to overcome the major limitation of static structure-based drug design and also to characterize routinely applied ligand docking calculations which do not sample the major protein conformational rearrangements during ligand binding (Carlson 2002; Fanelli et al. 2008). MD simulation is a multistep process that starts with the knowledge of the potential energy of the system with respect to its position coordinates, and these position coordinates help to compute the force acting on the individual atoms of the system. The next important step is simulation environment, which gives the actual environment including optimum pressure and temperature. In general, protein

simulation is done in a canonical ensemble (NVT), particularly the initial equilibrium steps, or it is done in an isothermal-isobaric (NPT) ensemble. For simulation, the protein molecule should be kept in the unit cell and solvated with a suitable explicit solvent. Several explicit water models include TIP3P, TIP4P (Jorgensen et al. 1983), TIP5P (Mahoney and Jorgensen 2001), SPC and SPC/E (Berendsen et al. 1987) are the most popular models used to imitate the specific nature and complexity of molecule hydration, including orientation of solvent dipoles and effective electrostatic shielding, subtle hydrogen bond network rearrangements, saturation of hydrophobic surface and accompanying changes in entropy.

There are two main families of MD simulation methods, classical and quantum simulation, which are distinguished based on the model chosen to represent a physical system. A basic ball-and-stick model of molecules was used in classical molecular simulation, where the atoms correspond to soft balls and elastic sticks correspond to bonds. Several force fields are widely used in the molecular simulation approach. AMBER (Case et al. 2005), NAMD (Phillips et al. 2005), CHARMM (Brooks et al. 1983) and GROMOS (Pronk et al. 2013) are widely used force fields which differ principally in the way they are parameterized, but they generally give similar results. Quantum simulation or first principle MD simulation began in 1980s with the seminal work of Car and Parinello, explicitly taking into account the quantum nature of the chemical bond. Due to the invention of high configurational computer and the advent of graphical processor unit (GPU) architectures, MD simulation software can efficiently run on innovative hardware infrastructures, surpassing alternate conventional methods. Even these methods, running on specialized hardware fails to describe the slow unbinding events. In fast-paced drug discovery programs, this is the major issue limiting the use of MD-based simulation for kinetic prediction (Borhani and Shaw 2012). However, sampling issues have led the development of several innovative algorithms that form the basis of the enhanced sampling method, speeding up the description of slow processes and accelerating the rare events characterized by high-in-free-energy states (Abrams and Bussi 2014). Sampling methods including free energy perturbation (Jorgensen and Thomas 2008), umbrella sampling, replica exchange, meta-dynamics (Laio and Parrinello 2002), steered MD (Israelewitz et al. 2001), accelerated MD (Hamelberg et al. 2004) milestoning (Faradjian and Elber 2004), transition-path sampling (Bolhuis et al. 2002), Monte Carlo sampling of conformational space, quantum mechanics/molecular mechanics (QM/MM) and molecular docking simulation are recently used methods for studying protein-ligand binding and estimating the associated energy and kinetics (Durrant and McCammon 2011; Harvey and Fabritius 2012).

11.4.7 QM/MM Simulations

Most of biological systems such as enzymes are heavy atoms, too large to be described at any level of ab initio theory, and classical molecular mechanics force field is not sufficiently flexible to model processes in which chemical bonds are

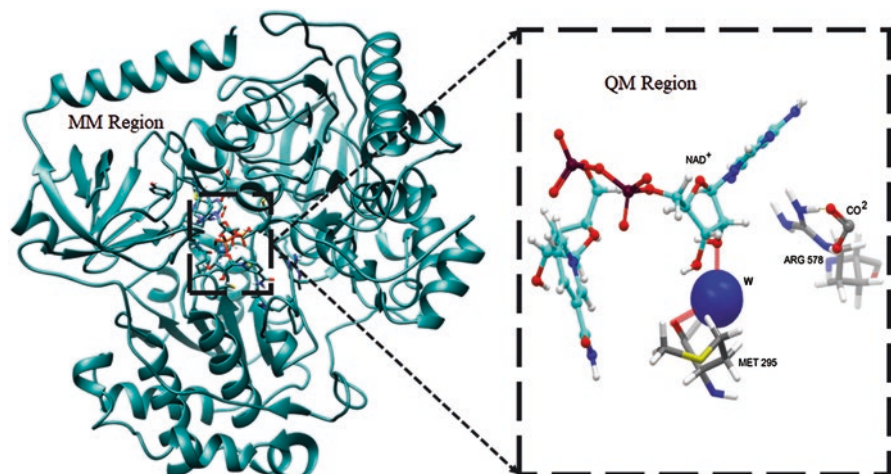


Fig. 11.6 Showing the focused QM region inside the MM region of the whole protein

broken or formed and make a proper model of the complex environment of the reaction, which involves efficient thermal averaging of the energy landscape. To overcome these issues, an alternative approach has been developed that treats a small part of the system at the level of quantum chemistry (QM) while retaining the computationally cheaper force field (MM) for the large part (Fig. 11.6).

This hybrid strategy QM/MM simulation was introduced by Warshel and Levitt and become a power full tool for the analysis of the enzyme reaction mechanism, playing a significant role in exciting applications like drug design (Gao and Truhlar 2002; Shaik et al. 2010; van der Kamp and Mulholland 2013; Lonsdale and Mulholland 2014). Basically, three classes of interaction are involved in QM/MM potential energy: interaction between atoms in the QM region, interaction between atoms in the MM region and interactions between QM and MM atoms. Quantum mechanics calculations are also an essential complement or alternative in the interpretation of outcomes of experiments by theoretical prediction of a molecular characteristic such as electrical and magnetic ones and properties related to geometrical derivatives (Cohen et al. 2012). QM treats molecules as a collection of nuclei and electrons, without any reference to chemical bonds, which is important in understanding the behavior of system at the atomic level. This method applies the laws of QM to approximate the wave function of Schrödinger equation in terms of the motions of electrons (Atkins and de Paula 2006; Tannor 2008). QM methods are a more accurate but they entail an expensive and time-consuming calculation. Calculations are employed in semi-empirical methods such as AM1 and PM3 only for valence electrons in the system. The combined QM-MM methods provide the accuracy of QM description with the low cost of MM (Lin and Truhlar 2007; Menikarachchi and Gascon 2010; Honarparvar et al. 2014). Quantum mechanics-based methods such as ab initio and the density functional theory (DFT) method fall

Table 11.4 Accuracy of different quantum mechanics methods

S. No	Types of quantum mechanics	Accuracy	Maximum atoms
1	Semi-empirical	Low	2000
2	Hartree–Fock and density functional	Medium	500
3	Perturbation and variation methods	High	50
4	Coupled cluster	Very high	20

within the approximate range of a few picometers to nanometers. These electronic structures allow accurate theoretical studies to be certain to extend to both macromolecules (synthetic polymers and proteins) and condensed matter (liquid and solids). DFT provides all the information on the system and avoids the wave function calculation. DFT is rooted in the Hohenberg–Kohn theorems, according to which the exact energy of a molecular system depends on its electron density; the latter being a function of the electronic coordinates. The total energy of a system can be calculated by the sum of several functionals such as kinetic energy, nucleus–electron potential energy, electron–electron repulsion energy and exchange–correlation functional. The choice of QM method, choice of MM force field, segregation of the system into QM and MM regions, simulation types and the advanced conformational sampling are the five important aspects of QM-MM calculation of an enzyme. The choice of QM method is crucial: there are different QM methods ranging from fast, semi-empirical methods to more accurate and more computationally expensive methods; however, not all the methods are applicable to all systems for reasons of accuracy, practicality or due to lack of parameters. The Table 11.4 shows the accuracy of different quantum methods.

11.5 Drug Delivery Approach Using Computational Methods

In drug delivery approach, potential drug molecule must have the capability to sustain its effectiveness, posing key challenges to effective drug delivery; an administered drug must penetrate obstacles such as endo or epithelial membranes and also survive the host's defenses to be effective. Hence, to overcome these challenges requires some form of drug encapsulation such as the unique molecular encapsulation architecture known as a drug delivery system (Allen et al. 2004; Blanco et al. 2015). This new approach of controlling the pharmacokinetics, thermodynamics, non-specific toxicity, immunogenicity, biorecognition and efficacy of drugs was generated to minimize drug degradation and loss and to prevent harmful side effects and increase drug bioavailability and the fraction of the drug that accumulates in the required zone (Reddy and Swarnalatha 2010). Several mechanisms are involved in a drug delivery system such as drug formulation, medical device or dosage technology to carry the drug inside the body and a mechanism for the release. Most of the commercial applications of nanoparticles in medicine are directed to drug delivery,

for which several solutions have been proposed, including liposomal and lipid-based colloidal nanoDDS, nanoparticulate polymeric micelles (as drug carrier and polymer-based nanoparticulate DDS. Molecular modeling and computational chemistry provide several tools such as quantum mechanical *ab initio* methods, molecular dynamics, free energy perturbation and docking to quantify drug-carrier, carrier-medium and drug-medium interactions (Neumann et al. 2004).

11.6 Polymer Used as Carrier

Polymers are naturally occurring substances with high molar masses and a large number of repeating units; they play a significant role in the development of drug delivery systems by releasing both hydrophilic and hydrophobic drug molecules. Covalent bond formation of polymers with drug molecules carries the drug molecules to their respective site. Hence, there are several advantages of polymers acting as inert carriers to which a drug can be conjugated; for example, polymers improve pharmacokinetic and pharmacodynamic properties of drug molecules. Polymers is an important constituents of pharmaceutical forms such as solid dosage as in tablets and capsules; they can be dispersed in a system like a suspension, emulsion, cream or ointment; and they can be made into a particulate system, microcapsules, microparticles and nanoparticles; and they are accepted that formulation in clinical performance of pharmaceutical dosage forms (Duncan 2003; Raizada et al. 2010). The main function of a polymeric carrier is to carry and transport drug molecules to the site of action. This polymeric drug delivery system significantly protects the drug molecule from interaction with other macromolecules including proteins and nucleic acids, which could alter the chemical structure of the drug molecules. Both non-biodegradable and biodegradable polymers have been used in drug delivery systems. Based on their desirable physical properties, polymers are selected and used in both non-biological and biological settings. Polymers such as polymethyl methacrylate, polyvinyl alcohol, polyurethane and polyethylene are a few examples of polymer use in non-biological processes. In recent years, polymers have been used as carrier molecules due to their unique features such as chemical inertness, freedom from impurities, appropriate physical structure and ability to be processed readily. Polyethylene-co-vinyl acetate, polymethyl methacrylate, polyvinyl alcohol, poly-N-vinyl pyrrolidine, polyacrylic acid and polyacrylamide are often used in controlled drug delivery system (Poddar et al. 2010; Harekrishna Roy et al. 2013). Smart polymers are those having the capability to change their properties in response to the changes in biological conditions (Yang and Pierstorff 2012). Several stimuli including temperature, pressure, pH electric field, magnetic field, light, change in concentration, ionic strength and potential may influence the changes in nature of polymer properties (Schmaljohann 2006). For example, a temperature-responsive polymer brings about changes in hydrophilicity/hydrophobicity of polymers, enhancing their membrane permeation. This alteration in polymer properties can be used to allow adhesion to a cell surface, to

break down a cellular membrane and to release biologically active compounds. Recently, polymers have been used for developing controlled drug release systems and sustained release formulations, which help regulate drug administration by preventing under- or overdosing. These advanced drug-releasing systems play a significant role in improving bioavailability, minimizing side effects and other types of inconveniences (Liechty et al. 2010).

11.6.1 Drug-Polymer Interaction

Most computational studies for drug delivery use molecular dynamics simulation, which mimics the natural pathway of molecular motion to sample successive configuration. Newton's law and Maxwell–Boltzmann distribution assign initial velocity of molecules at a given temperature. The interactions between molecules at each time are computed and then equations of motion are solved numerically with an appropriate time step to update the velocities and position for the next successive steps (Frenkel and Smit 2002).

In classical molecular dynamics simulations, the interaction of molecules can be described by a force field with certain functional forms and several parameters. A force field such as AMBER (Cornell et al. 1995), OPLS (Jorgensen et al. 1996) and CHARMM (Mackerell et al. 1998) is widely used to study polymer and peptide drug interactions. Interactions such as hydrogen bonding (Zhang et al. 2012; Miyazaki et al. 2011), dipole-dipole interaction (Marsac et al. 2009; Khougaz and Clas 2000), ionic interaction (Yoo et al. 2009; Kindermann et al. 2011) and van der Waals interaction (Marsac et al. 2009) generally occur between drug and polymer. Dissipative particle dynamics (DPD) is a widely used mesoscale simulation for identifying and defining chemically distinct components and defining interaction parameters between various chemical species. In this model, a fluid system is simulated using a set of interacting particles. Each particle represents a cluster of small molecules instead of a single molecule. Drug, polymer, surfactant and solvent are represented as distinct bead types. Polymer bead number length is determined by

$$N_{\text{DPD}} = \frac{Mp}{Mm C_{\infty}},$$

where Mp is polymer molecular weight, Mm monomer molecular weight and C_{∞} polymer characteristic ratio. However, a detailed mechanism on drug-polymer interactions is lacking, such as how chemically substituted cellulosic polymers interact with drug molecules at a molecular level and how different structural variables such as molecular weight and substitution pattern affect the drug-polymer interaction. In addition to the classical MD and DPD, another two levels of molecular models such as coarse-grained molecular dynamics (CG-MD) simulations,

which are used to model excipients such as modified cellulosic polymers at a monomer level resolution and drugs at a similar level. The full spectrum of the CG-MD approach contains contributions from several different fields and continuum transport modeling, in which diffusion equations for transport of polymer, drug and solvent through a capsule are determined by solving the relevant differential equation. Several software packages can integrate these equations, including the popular GROMACS (Van der Spoel et al. 2005), NAMD (Phillips et al. 2005), CHARMM (Klauda et al. 2010) and AMBER (Wang et al. 2004) packages. Many of the coarse-grained methods utilize one of these integrators to perform simulations.

11.7 Computational Methods Used in Toxicity Studies

Toxicity is a measurement of the adverse effect of chemicals, and specific types of these adverse effects are known as toxicity endpoints, for example, carcinogenicity or genotoxicity. These adverse effects can be quantitatively or qualitatively measured to identify harmful effects caused by substances on humans and animals (Rowe et al. 2010). A number of factors determine the toxicity of chemicals, including route of exposure, dose, duration of exposure, ADME properties (absorption, distribution, metabolism and excretion), biological properties and chemical properties (Raies and Bajic 2016). A number of in vitro models have been used to determine toxicity such as high throughput screening (AltTox) and in vivo animal models. Recently, computational toxicity methods have been widely used to potentially minimize the need for animal testing and reduce the cost and time of the toxicity test to improve toxicity prediction and safety assessment. The major advantage of computational toxicity methods is their ability to estimate chemicals for toxicity even before they are synthesized (Madan et al. 2013). In silico toxicology analysis encompasses a wide range of computational tools including database storage of chemical data, their toxicity and chemical properties, and software for generating molecular descriptors, simulation tools for systems biology and molecular dynamics and modeling methods for toxicity. Rule-based and structural alerts are often-used computational methods for determining toxicity based on chemical properties and how drugs should be altered to reduce their toxicity. Another method, read-across, is used to predicting the unknown toxicity of a chemical through the use of similar chemicals (analogs) with known toxicity from the same chemical category (Dimitrov and Mekenyan 2010; Modi et al. 2012; Benigni et al. 2013; Venkatapathy and Wang 2013;). There are two approaches—an analog, or one-to-one approach, and a category, or many-to-one approach—for developing a read-across method. Both approaches are quite sensitive, identifying similar chemicals by calculating their properties and the similarities between them. The main advantage of read-across is its transparency (Cronin 2011): it is easy to interpret and implement (Enoch 2009), and it can model quantitative and qualitative toxicity endpoints and allow for

a wide range of types of descriptors and similarity measures to be used to express similarity between chemicals (Dimitrov and Mekenyan 2010).

Quantitative structure-activity relationship (QSAR) is another widely used method that employs molecular descriptors to predict a chemical's toxicity. Generally, the QSAR method predicts toxicity (T) of a lead molecule using a vector feature of chemical properties (θ_p) and a function f that calculates T given θ_p is

$$T = f(\theta_p).$$

There are two QSAR models: local QSAR, which is generated from congeneric chemicals, and global QSAR, which is made from diverse chemicals. Local QSAR is used to predict toxicity based on the mode of action of specific chemicals, hence, local QSAR are more accurate as they are customized for specific chemicals (Valerio 2009). Mainly two basic steps are involved in the development of a QSAR model: the generation of molecular descriptors and then of models to fit the data. The number of molecular descriptors, as based on simulated annealing, generic algorithm or principal component analysis, can be used to determine the chemicals (Deeb and Goodarzi 2012; Devillers 2013). If there are a small number of descriptors, using two-dimensional scatterplots of each descriptor versus its biological activity can help identify significant descriptors (Devillers 2013). There are many tools available that provide pre-built QSAR model such as OECD QSAR Toolbox (OECD 2015), TopKat (Accelrys 2015) and METEOR (Lhasa Limited, Meteor Nexus 2014). The major advantage of QSAR is that it's easy to interpret and it can model categorical and continuous toxicity endpoints and molecular descriptors and toxic and non-toxic chemicals. However, it may not be always employable, as a large number of chemicals are needed in the model development for QSAR to achieve statistical significance (Valerio 2009; Deeb and Goodarzi 2012).

Pharmacokinetic (PK) models relate to the concentration of drug molecules in tissues to time, estimating the amount of chemicals in different parts of the body and quantifying ADME (absorption, distribution, metabolism and excretion) processes (Jack et al. 2013; Sung et al. 2014). Mainly, the PK models are used to relate chemical concentration in a part of the body to time of toxic responses. A PK model can be categorized as two models: compartment and non-compartmental (Sung et al. 2014). A compartment model consists of one more compartments, with each compartment represented by differential equations (Sung et al. 2014). One compartment model represents the whole body as a single compartment, assuming rapid equilibrium of chemical concentration within the body but not considering the time to distribute of the chemical. Two-compartment models consist of two compartments, the central and peripheral with both compartments represented by differential equations. These models provide mechanistic insight based on pharmacokinetic models including concentration and time, physiological descriptors of tissues and ADME processes such as volumes, blood flows, chemical binding/partitioning, metabolism and excretion (Jack et al. 2013; El-Masri 2013).

11.8 Outcome of Drug Research in Bacterial Inhibitors

Bacterial infection is one of the major threats to human health because it frequently causes severe diseases not only in the form of primary agents but also after pathologies caused by other agents. Compared to Gram-negative bacteria, Gram-positive bacteria have a much thicker peptidoglycan layer, which is responsible for the increasing occurrence of bacterial resistance to antibiotics in medicinal practice (Springer et al. 2010; Nikaido 2003). Since the discovery of several antibiotics in the mid-twentieth century, resistance has been a concern (Peters et al. 2008). Although the emergence of antibacterial resistance is not new, it continues to be a major health concern. The report from the Centers for Disease Control and Prevention on antimicrobial resistance revealed that more than 21% of hospital-acquired infections were caused by an antimicrobial resistant pathogen. Hence, there is a need for new alternatives in the treatment of infections by multi-resistant bacteria. Among the several pathogens, *Staphylococcus aureus* resistant to methicillin (MRSA), *Streptococcus pneumoniae*, resistant to penicillin, glycopeptide-intermediately-resistant *S. aureus* (GISA), methicillin-resistant *S. epidermidis*, glycopeptide-resistant *enterococcus spp* and vancomycin-resistant *Enterococci* (VRE) are the more important etiological agents of hospital and community infections and are responsible for high rates of morbidity and mortality in hospitalized patients (Woodfor and Livermore 2009; Livermore 2009; Arias and Murray 2009). Several fluoroquinolones, ramoplanin, beta-lactams and the quinupristin/dalfopristin are currently used in the market. Moellering et al. (1999) studied the clinical efficacy and safety of quinupristin-dalfopristin in the treatment of a patient with a vancomycin-resistant infection. From the studies it was noted that the overall clinical and bacteriologic success rate was 66%. In another study, Nichols et al. (1999) compared quinupristin-dalfopristin with cefazolin, oxacillin and vancomycin in two randomized, open-label clinical trials.

Oxazolidinones, an antimicrobial class of agents, are a unique family of drug molecule possessing activity against *Staphylococcus aureus* and glycopeptide-intermediately-resistant *S. aureus* (Rybak et al. 2000; Wootton et al. 2000) and they are also more effective against a wide range of Gram-positive bacteria and *Mycobacterium tuberculosis*. Linezolid was the first approved derivative with acceptable tolerability in humans for the treatment of pneumonia, skin and soft tissues infections caused by VRE (Cammarata et al. 2000). Daptomycin is another antibacterial agent used to treat a wide range of Gram-positive bacteria. Recent studies from the US and Europe revealed that daptomycin was active against all *Staphylococcus aureus* and Gram-negative bacteria such as *Leuconostoc*, which are characteristically resistant to glycopeptides (Barry et al. 2001; King and Phillips 2001). The effectiveness of daptomycin has been proved in various animal models of Gram-positive infection. Several global randomized, double blind phase II trials have investigated the efficacy of daptomycin in the treatment of community-acquired pneumonia (Pertel et al. 2008).

11.9 Future Aspects of Computational Methods in Targeting Bacterial Infections

The drug-resistant capability of Gram-positive bacteria is a serious issue in clinical practice, and several antibacterial agents have already been approved by the US Federal Drug Administration for several infections, while other agents are still undergoing clinical trials. However, a lack of effective antibiotics in development implies that future treatment strategies for the resistant bacteria may have to show enhanced therapeutic efficacy. The battle against antibiotic resistance can be carried out on two fronts: either in advancing research efforts toward the discovery of novel and potential agents or by enhancing the effectiveness of the currently available ones. With the increasing prevalence of bacterial resistance, there is need to identify potential lead molecules to combat them. Conventional drug development research requires huge investment and at least 12–15 years experimentation, and even so, it often does not reach the market; hence, alternative approaches and strategies are required to develop safe and effective novel antimicrobial therapies. The current scenario of antibiotic research and development is not very effective, so a computational approach such as structure-based drug design, ligand-based drug design, pharmacophore modeling and molecular docking are useful for understanding the mechanism of bacterial resistance to antibiotics. In addition to the experimental approach, computational biology combination therapy has great potential in the future discovery of antimicrobial drugs.

References

- Abel R, Young T, Farid R, Berne BJ, Friesner RA (2008) Role of the active-site solvent in the thermodynamics of factor Xa ligand binding. *J Am Chem Soc* 130:2817–2831
- Abrams C, Bussi G (2014) Enhanced sampling in molecular dynamics using metadynamics, replica-exchange, and temperature-acceleration. *Entropy* 16:163–199
- Accelrys Available Chemicals Directory (ACD) (2012) Accelrys, Inc., San Diego, CA
- Accelrys. TOPKAT (TOxicity Prediction by Komputer Assisted Technology) (2015). Available at: <http://accelrys.com/products/discovery-studio/admet.html>
- Acharya C, Coop A, Polli JE, MacKerell AD (2011) Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. *Curr Comput Aided Drug Des* 7:10–22
- Adcock SA, McCammon JA (2006) Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem Rev* 106:1589–1615
- Agrafiotis DK, Gibbs AC, Zhu F, Izrailev S, Martin E (2007) Conformational sampling of bioactive molecules: a comparative study. *J Chem Inf Model* 47:1067–1086
- Akamatsu M (2002) Current state and perspectives of 3D-QSAR. *Curr Top Med Chem* 2:1381–1394
- Albrich WC, Monnet DL, Harnarth S (2004) Antibiotic selection pressure and resistance in *Streptococcus pneumoniae* and *Streptococcus pyogenes*. *Emerg Infect Dis* 10:514–517
- Allen TD, Eby LT, Poteet ML, Lentz E, Lima L (2004) Career benefits associated with mentoring for protégé: a meta-analysis. *J Appl Psychol* 89:127–136

- Allen GD, Breshears DD, McDowell NG (2015) On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene. *Ecosphere* 6:1–55
- AltTox. Toxicity testing overview. Available at: <http://alttox.org/mapp/toxicity-testing-overview/>
- An J, Totrov M, Abagyan R (2005) Pocketome via comprehensive identification and classification of ligand binding envelopes. *Mol Cell Proteomics* 4:752–761
- Anandakrishnan R, Aguilar B, Onufriev AV (2012) H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulation. *Nucleic Acids Res* 40:W537–41–W541
- Arias CA, Murray BE (2009) Antibiotic-resistant bugs in the 21 st century: a super-clinical challenge. *N Engl J Med* 360:439–443
- Ashrafuzzaman M (2014) Aptamers as both drugs and drug-carriers. *Biomed Res Int* 2014:697923
- Atkins P, de Paula J. (2006) *Atkins' Physical Chemistry*. 8th ed. New York: Macmillan.
- Baker D, Sali A (2001) Protein structure prediction and structural genomics. *Science* 294(5540):93–96
- Ballester PJ, Mitchell JBO (2010) A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. *Bioinformatics (Oxford, England)* 26:1169–1175
- Barry AL, Fuchs PC, Brown SD (2001) In vitro activities of daptomycin against 2,789 clinical isolates from 11 North American medical centers. *Antimicrob Agents Chemother* 45:1919–1922
- Benigni R, Battistelli CL, Bossa C, Colafranceschi M, Tcheremenskaia O (2013) Mutagenicity, carcinogenicity, and other end points. In: Reisfeld B, Mayeno AN (eds) *Computational toxicology*, vol 930. Humana Press, New York, pp 67–98
- Beno B, Mason J (2001) The design of combinatorial libraries using properties and 3D-pharmacophore fingerprints. *Drug Discov Today* 6:251–258
- Berendsen HJC, Grigera JR, Straatsma TP (1987) The missing term in effective pair potentials. *J Phys Chem* 91(24):6269–6271
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. *Nucleic Acids Res* 28:235–242
- Bernard D, Coop A, MacKerell AD (2005) Conformationally sampled pharmacophore for peptidic delta opioid ligands. *J Med Chem* 48:7773–7780
- Binkowski T, Naghibzadeh S, Liang J (2003) CASTp: computed atlas of surface topography of proteins. *Nucleic Acids Res* 31:3352–3355
- Blanco E, Shen H, Ferrari M (2015) Principles of nanoparticle design for overcoming biological barriers to drug delivery. *Nat Biotechnol* 33:941–951
- Bleicher KH, Bohm HJ, Muller K, Alanine AI (2003) Hit and lead generation: beyond high-throughput screening. *Nat Rev Drug Discov* 2:369–378
- Bolhuis PG, Chandler D, Dellago C, Geissler PL (2002) Transition path sampling: throwing ropes over rough mountain passes, in the dark. *Annu Rev Phys Chem* 53:291–318
- Borhani DW, Shaw DE (2012) The future of molecular dynamics simulations in drug discovery. *J Comput Aided Mol Des* 26:15–26
- Brady GP, Stouten PFW (2000) Fast prediction and visualization of protein binding pockets with PASS. *J Comput Aided Mol Des* 14:383–401
- Briggs JM, Marrone TJ, McCammon JA (1996) Computational science new horizons and relevance to pharmaceutical design. *Trends Cardiovasc Med* 6:198–206
- Bronzwaer SL, Bronzwaer SL, Cars O, Buchholz U, Mølsted S, Goettsch W, Veldhuijzen IK, Kool JL, Sprenger MJ, Degener JE (2002) A European study on the relationship between antimicrobial use and antimicrobial resistance. *Emerg Infect Dis* 8:278–282
- Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM – a program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 4:187–217
- Bryant SH, Lawrence CE (1993) An empirical energy function for threading protein sequence through the folding motif. *Proteins* 16(1):92–112
- Burger A, Abraham DJ (2006) *Burger's medicinal chemistry and drug discovery, drug discovery and drug development*; Wiley, 2003. In: Langer T, Hoffmann RD (eds) *Pharmacophores and Pharmacophore Searches*. Wiley-VCH Verlag GmbH & Co, Weinheim

- Cammarata SK, Timm JA, Hemsall KA, Todd WM, Oliphant TH, Hafkin B (2000) Efficacy of linezolid in pneumonia due to penicillin intermediate and resistant *Streptococcus pneumoniae*. [abstract SO.O10]. In: Programs and abstracts of the 9th International Congress on Infectious Diseases; Buenos Aires, Argentina, April 10–13. International Society of Infectious Diseases, Boston
- Carlson HA (2002) Protein flexibility and drug design: how to hit a moving target. *Curr Opin Chem Biol* 6:447–452
- Case DA, Cheatham TE 3rd, Darden T, Gohlke H, Luo R (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26:1668–1688
- CASP (1999). Proteins, Structure, Function, and Genetics. CASP3 Proceeding 37:1–237
- Cato S (2000) Pharmacophore perception, development, and use in drug design. International University Line, La Jolla, CA, Chapter Exploring Pharmacophores with Chem-X,, pp 107–125
- Chang C, Swaan PW (2006) Computational approaches to modeling drug transporters. *Eur J Pharm Sci* 27:411–424
- Chen J, Swamidass SJ, Dou Y, Bruand J, Baldi P (2005) ChemDB: a public database of small molecules and related cheminformatics resources. *Bioinformatics* 21(22):4133–4139
- Chen JH, Linstead E, Swamidass SJ, Wang D, Baldi P (2007) ChemDB update-full-text search and virtual chemical space. *Bioinformatics* 23:2348–2351
- Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson SJ, Richardson DC (2010) MolProbity: allatom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* D66:12–21
- Clark DE (2006) What has computer-aided molecular design ever done for drug discovery? *Expert Opin Drug Discovery* 1:103–110
- Clark DE (2008) What has virtual screening ever done for drug discovery? *Expert Opin Drug Discovery* 3:841–851
- Claussen H, Buning C, Rarey M, Lengauer T (2001) FlexE: efficient molecular docking considering protein structure variations. *J Mol Biol* 308:377
- Cohen AJ, Mori-Sánchez P, Yang W (2012) Challenges for density functional theory. *Chem Rev* 112:289–320
- Collignon P, Powers JH, Chiller TM, Aidara-Kane A, Aarestrup FM (2009) World Health Organization ranking of antimicrobials according to their importance in human medicine: a critical step for developing risk management strategies for the use of antimicrobials in food production animals. *Clin Infect Dis* 49:132–141
- Corbeil CR, Williams CI, Labute P (2012) Variability in docking success rates due to dataset preparation. *J Comput Aided Mol Des* 26:775–786
- Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A 2nd generation force-field for the simulation of proteins, nucleic acids and organic molecules. *J Am Chem Soc* 117:5179–5197
- Cronin MTD (2011) In silico tools for toxicity prediction. In: Wilson AGE (ed) *New horizons in predictive toxicology: current status and application*. Royal Society of Chemistry, Cambridge
- D'Alfonso G, Tramontano A, Lahm A (2001) Structural conservation in single-domain proteins: implications for homology modeling. *J Struct Biol* 134:246–256
- Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, Snoeyink J, Richardson JS, Richardson DC (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35:W375–W383
- Deeb O, Goodarzi M (2012) In silico quantitative structure toxicity relationship of chemical compounds: some case studies. *Curr Drug Saf* 7:289–297
- Devillers J (2013) Methods for building QSARs. In: Reisfeld B, Mayeno AN (eds) *Computational toxicology*, vol 930. Humana Press, New York, pp 3–27
- Dias R, de Azevedo WF Jr (2008) Molecular docking algorithms. *Curr Drug Targets* 9(12):1040–1047
- DiMasi JA, Hansen RW, Grabowski HG (2003) The price of innovation: new estimates of drug development costs. *J Health Econ* 22:151–185

- Dimitropoulos D, Ionides J, and Henrick K (2006) Using PDBeChem to Search the PDB Ligand Dictionary, in *Current Protocols in Bioinformatics*; John Wiley & Sons, 14.13.11–14.13.13.
- Dimitrov S, Mekenyan O (2010) An introduction to readacross for the prediction of the effects of chemicals. In: Cronin MTD, Madden JC (eds) *In silico toxicology: principles and applications*. The Royal Society of Chemistry, Cambridge, pp 372–384
- Doernberg SB, Lodise TP, Thaden JT, Munita JM, Cosgrove SE, Arias CA, Boucher HW, Corey GR, Lowy FD, Murray B, Miller LG, Holland TL (2017) Gram-positive bacterial infections: research priorities, accomplishments, and future directions of the antibacterial resistance leadership group. *Clin Infect Dis* 64:S24
- Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, Baker NA (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res* 35:W522–W525
- Duncan R (2003) The dawning era of polymer therapeutics. *Nat Rev Drug Discov* 2:347–360
- Durrant J, McCammon JA (2011) Molecular dynamics simulations and drug discovery. *BMC Biol* 9:71
- Ekins S, Mestres J, Testa B (2007) In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. *Br J Pharmacol* 152:9–20
- El-Masri H (2013) Modeling for regulatory purposes (risk and safety assessment). In: Reisfeld B, Mayeno AN (eds) *Computational toxicology*, vol 930. Humana Press, New York, pp 297–303
- Enoch SJ (2009) Chemical category formation and readacross for the prediction of toxicity. In: Puzyn T, Leszczynski J, Cronin MT (eds) *Recent advances in QSAR studies*, vol 8. Springer, Dordrecht, pp 209–219
- Enyedy IJ, Sakamuri S, Zaman WA, Johnson KM, Wang S (2003) Pharmacophore-based discovery of substituted pyridines as novel dopamine transporter inhibitors. *Bioorg Med Chem Lett* 13:513–517
- Epik v2.2 (2011) Portland, OR: Schrödinger, Inc
- Ernst A, Obrecht D (2008) Case study of parallel synthesis in hit identification, hit exploration, hit-to-lead, and lead-optimization programs in high-throughput lead optimization in drug discovery. In: Kshirsagar T (ed) *CRC Press, Boca Raton, USA*, pp 99–116
- Fanelli F, Gohlke H, Kuhn LA, Morris GM, Orozco M, Pertinhez TA, Rizzi M, Sotriffer C (2008) Target flexibility: an emerging consideration in drug discovery and design. *J Med Chem* 51:6237–6255
- Fang Y (2012) Ligand-receptor interaction platforms and their applications for drug discovery. *Expert Opin Drug Discovery* 7:969–988
- Faradjian AK, Elber R (2004) Computing time scales from reaction coordinates by milestoning. *J Chem Phys* 120:10880–10889
- Faraggi E, Xue B, Zhou Y (2009) Improving the prediction accuracy of residue solvent accessibility and real-value backbone torsion angles of proteins by guided-learning through a two-layer neural network. *Proteins* 74:847–856
- Fiser A (2010) Template-based protein structure modeling. *Methods Mol Biol* (Clifton, NJ) 673:73–94
- Floudas CA (2007) Computational methods in protein structure prediction. *Biotechnol Bioeng* 97:207–213
- Floudas C, Fung H, McAllister S, Moennigmann M, Rajgaria R (2006) Advances in protein structure prediction and de novo protein design: a review. *Chem Eng Sci* 61:966
- Flower DR (2002) *Drug design: cutting edge approaches*. Royal Society of Chemistry, Cambridge
- Frenkel D, Smit B (2001) *Understanding molecular simulation*. Academic Press, Inc, San Diego, p 638
- Frenkel D, Smit B (2002) *Understanding molecular simulation: from algorithms to applications*. Academic Press, San Diego
- Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* 47:1739–1749
- Gao JL, Truhlar DG (2002) Quantum mechanical methods for enzyme kinetics. *Annu Rev Phys Chem* 53:467–505

- Gehlhaar DK, Verkhivker GM, Rejto PA, Sherman CJ, Fogel DB, Freer ST (1995) Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. *Chem Biol* 2:317–324
- Gehlhaar DK, Bouzida D, Rejto PA (1999) In: Parrill L, Reddy MR (eds) *In rational drug design: novel methodology and practical applications*, vol 719. American Chemical Society, Washington, DC, pp 292–311
- Geppert H, Vogt M, Bajorath J (2010) Current trends in ligand-based virtual screening: molecular representations, data mining methods, new application areas, and performance evaluation. *J Chem Inf Model* 50:205–216
- Gilson MK, Zhou HX (2007) Calculation of protein-ligand binding affinities. *Annu Rev Biophys Biomol Struct* 36:21–42
- Giske CG, Monnet DL, Cars O, Carmeli Y (2008) Clinical and economic impact of common multidrug-resistant Gram-negative bacilli. *Antimicrob Agents Chemother* 52:813–821
- Göbl C, Madl T, Simon B, Sattler M (2014) NMR approaches for structural analysis of multidomain proteins and complexes in solution. *Prog Nucl Magn Reson Spectrosc* 80:26–63
- Gohda K, Mori I, Ohta D, Kikuchi T (2000) A CoMFA analysis with conformational propensity: an attempt to analyze the SAR of a set of molecules with different conformational flexibility using a 3D-QSAR method. *J Comput Aided Mol Des* 14:265–275
- Gohlke H, Klebe G (2001) Statistical potentials and scoring functions applied to protein-ligand binding. *Curr Opin Struct Biol* 11:231–235
- González PM, Acharya C, MacKerell AD Jr, Polli JE (2009) Inhibition requirements of the human apical sodium-dependent bile acid transporter (hASBT) using aminopiperidine conjugates of glutamyl-bile Acids. *Pharm Res* 26:1665–1678
- Goto E, Yagi Y, Matsumoto Y, Tsubota K (2002) Impaired functional visual acuity of dry eye patients. *Am J Ophthalmol* 133:181–186
- Guedes IA, de Magalhães CS, Dardenne LE (2014) Receptor-ligand molecular docking. *Biophys Rev* 6:75–87
- Guimaraes CRW, Cardozo M (2008) MM-GB/SA rescoring of docking poses in structure-based lead optimization. *J Chem Inf Model* 48:958–970
- Hajduk PJ, Huth JR, Fesik SW (2005) Druggability indices for protein targets derived from NMR-based screening data. *J Med Chem* 48:2518–2525
- Halgren T (2007) New method for fast and accurate binding-site identification and analysis. *Chem Biol Drug Des* 69:146–148
- Halgren TA (2009) Identifying and characterizing binding sites and assessing druggability. *J Chem Inf Model* 49:377–389
- Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem* 47:1750–1759
- Hamelberg D, Mongan J, McCammon JA (2004) Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J Chem Phys* 120:11919–11929
- Hanson SM, Newstead S, Swartz KJ, Sansom MS (2015) Capsaicin interaction with TRPV1 channels in a lipid bilayer: molecular dynamics simulation. *Biophys J* 108:1425–1434
- Harvey MJ, De Fabritiis G (2012) High-throughput molecular dynamics: the powerful new tool for drug discovery. *Drug Discov Today* 17:1059–1062
- Hawkey PM, Jones AM (2009) The changing epidemiology of resistance. *J Antimicrob Chemother* 64:i3–i10
- Hendlich M, Rippmann F, Barnickel G (1997) LIGSITE: automatic and efficient detection of potential small molecule binding sites in proteins. *J Mol Graph Model* 15:359–363
- Herrling PL (2005) The drug discovery process. *Prog Drug Res* 62:1–14
- Hileman B (2006) Accounting for R&D, Many doubt the \$800 million pharmaceutical price tag. *Chemical Eng News* 84:50–51
- Honarparvar B, Kruger HG, Maguire GE, Govender T, Soliman ME (2014) Integrated approach to structure-based enzymatic drug design: molecular modeling, spectroscopy, and experimental bioactivity. *Chem Rev* 114:493–537

- Hoofnagle AN, Resing KA, Ahn NG (2003) Protein analysis by hydrogen exchange mass spectrometry. *Annu Rev Biophys Biomol Struct* 32:1–25
- Hoofst R, Vriend WG, Sander C, Abola EE (1996) Errors in protein structures. *Nature* 381:272
- Hristozov D, Oprea TI, Gasteiger J (2007) Ligand-based virtual screening by novelty detection with self-organizing maps. *J Chem Inf Model* 47(6):2044–2062
- Huang SY, Zou X (2006a) An iterative knowledge-based scoring function to predict protein-ligand interactions: I. Derivation of interaction potentials. *J Comput Chem* 27:1866–1875
- Huang SY, Zou X (2006b) An iterative knowledge-based scoring function to predict protein-ligand interactions: II. Validation of the scoring function. *J Comput Chem* 27:1876–1882
- Huang X, Lin J, Demner-Fushman D (2006) Evaluation of PICO as a knowledge representation for clinical questions. *AMIA Annual Symposium proceedings. AMIA Symposium 2006*:359–363
- Irwin JJ, Shoichet BK (2005) ZINC - a free database of commercially available compounds for virtual screening. *J Chem Inf Model* 45:177–182
- Isralewitz B, Gao M, Schulten K (2001) Steered molecular dynamics and mechanical functions of proteins. *Curr Opin Struct Biol* 11:224–230
- Jack J, Wambaugh J, Shah I (2013) Systems toxicology from genes to organs. In: Reisfeld B, Mayeno AN (eds) *Computational toxicology*, vol 930. Humana Press, New York, pp 375–397
- Jain AN (2003) Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J Med Chem* 46:499–511
- Jain AN (2006) Scoring functions for protein-ligand docking. *Curr Protein Pept Sci* 7:407–420
- Jones DT, Taylor WR, Thornton JM (1992a) The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 3:275–282
- Jones DT, Taylor WR, Thornton JM (1992b) A new approach to protein fold recognition. *Nature* 358:86–89
- Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* 267:727–748
- Jorgensen WL, Thomas LL (2008) Perspective on free-energy perturbation calculations for chemical equilibria. *J Chem Theory Comput* 4:869–876
- Jorgensen WL, Tirado-Rives J (2005) Molecular modeling of organic and biomolecular systems using BOSS and MCPRO. *J Comput Chem* 26:1689–1700
- Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935
- Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118:11225–11236
- Jubb H, Blundell TL, Ascher DB (2015) Flexibility and small pockets at protein-protein interfaces: new insights into druggability. *Prog Biophys Mol Biol* 119:2–9
- Kahsai AW, Xiao K, Rajagopal S, Ahn S, Shukla AK, Sun J, Oas TG, Lefkowitz RJ (2011) Multiple ligand-specific conformations of the β 2-adrenergic receptor. *Nat Chem Biol* 7:692–700
- Kalyaanamoorthy S, Chen YP (2011) Structure-based drug design to augment hit discovery. *Drug Discov Today* 16:831–839
- Karchin R, Cline M, Mandel-Gutfreund Y, Karplus K (2003) Hidden Markov models that use predicted local structure for fold recognition: alphabets of backbone geometry. *Proteins* 51:504–514
- Karnachi P, Kulkarni A (2006) Application of pharmacophore fingerprints to structure based design and data mining. In: Langer T, Hoffmann RD (eds) *Pharmacophores and pharmacophore searches*. Wiley-VCH, Weinheim
- Karplus K, Karchin R, Draper J, Casper J, Mandel-Gutfreund Y, Diekhans M, Hughey R (2003) Combining local-structure, fold-recognition and new fold methods for protein structure prediction. *Proteins* 53:491–496
- Kaserer T, Beck KR, Akram M, Odermatt A, Schuster D (2015) Pharmacophore models and pharmacophore-based virtual screening: concepts and applications exemplified on hydroxysteroid dehydrogenases. *Molecules* 20:22799–22832

- Kaveti S, Engen JR (2006) Protein interactions probed with mass spectrometry. *Methods Mol Biol* 316:179–197
- Kelley LA, MacCallum RM, Sternberg MJ (2000) Enhanced genome annotation using structural profiles in the program 3d-pssm. *J Mol Biol* 299:499–520
- Khogaz K, Clas SD (2000) Crystallization inhibition in solid dispersions of MK-0591 and poly(vinylpyrrolidone) polymers. *J Pharm Sci* 89:1325–1334
- Kindermann C, Matthee K, Strohmeyer J, Sievert F, Breitzkreutz J (2011) Tailor-made release triggering from hot-melt extruded complexes of basic polyelectrolyte and poorly watersoluble drugs. *Eur J Pharm Biopharm* 79:372–381
- King A, Phillips I (2001) The in vitro activity of daptomycin against 514 Gram-positive aerobic clinical isolates. *J Antimicrob Chemother* 48:219–223
- Kitchen DB, Decornez H, Furr JR, Bajorath J (2004) Docking and scoring in virtual screening for drug discovery: methods and application. *Nat Rev Drug Discov* 3:935–949
- Klauda JB, Venable RM, Freites JA, O'Connor JW, Tobias DJ, Mondragon-Ramirez C, Vorobyov I, MacKerell AD, Pastor RW (2010) Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J Phys Chem B* 114:7830–7843
- Klebe G (2006) Virtual ligand screening: strategies, perspectives and limitations. *Drug Discov Today* 11:580–594
- Klebe G, Abraham U, Mietzner T (1994) Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J Med Chem* 37:4130–4146
- Kollman PA (1993) Free energy calculations: applications to chemical and biochemical phenomena. *Chem Rev* 93:2395–2417
- Kountouris P, Hirst JD (2009) Prediction of backbone dihedral angles and protein secondary structure using support vector machines. *BMC Bioinformatics* 10:437
- Kovalenko A (2003) Three-dimensional RISM theory for molecular liquids and solid-liquid interfaces. In: Hirata F, Mezey PG (eds) 360, vol 24. Kluwer Academic Publishers, Dordrecht, pp 169–275
- Kramer MS, Cutler N, Feighner J, Shrivastava R, Carman J, Sramek JJ, Reines SA, Liu G, Snavelly D, Wyatt-Knowles E, Hale JJ, Mills SG, MacCoss M, Swain CJ, Harrison T, Hill RG, Hefti F, Scolnick EM, Cascieri MA, Chicchi GG, Sadowski S, Williams AR, Hewson L, Smith D, Carlson EJ, Hargreaves RJ, Rupniak NM (1998) Distinct mechanism for antidepressant activity by blockade of central substance P receptors. *Science* 281:1640–1645
- Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE (1982) A geometric approach to macromolecule-ligand interactions. *J Mol Biol* 161:269–288
- Laio A, Parrinello M (2002) Escaping free-energy minima. *Proc Natl Acad Sci U S A* 99:12562–12566
- Langer T, Hoffmann RD (2006) Pharmacophore modeling: application in drug discovery. *Expert Opin Drug Discovery* 1:261–267
- Laskowski RA (1995 Oct) SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J Mol Graph* 13(5):323–330, 307–308
- Laskowski RA, Moss DS, Thornton JM (1993) Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* 231:1049–1067
- Leach AR, Hann MM (2000) The in silica world of virtual libraries. *Drug Discov Today* 5:326–336
- Levitt DG, Banaszak LJ (1992 Dec) POCKET: a computer graphics method for identifying and displaying protein cavities and their surrounding amino acids. *J Mol Graph* 10(4):229–234
- Lewis K (2013) Platforms for antibiotic discovery. *Nat Rev Drug Discov* 12:371–387
- Lewis RA, Pickett SD, Clark DE (2000) Computer-aided molecular diversity analysis and combinatorial library design. *Rev Comput Chem* 16:1–51
- Lhasa Limited. Meteor Nexus (2014). Available at: <http://www.lhasalimited.org>
- Li H, Robertson AD, Jensen JH (2005) Very fast empirical prediction and rationalization of protein pKa values. *Proteins* 61:704–721
- Liechty WB, Kryscio DR, Brandon VS, Peppas NA (2010) Polymers for drug delivery. *Annu Rev Chem Biomol Eng* 1:149–173

- LigPrep v2.5 (2011) Portland, OR: Schrödinger, Inc
- Lin H, Truhlar DG (2007) QM/MM: what have we learned, where are we, and where do we go from here? *Theor Chem Acc* 117:185–199
- Liu HY, Zou X (2006) Electrostatics of ligand binding: parametrization of the generalized born model and comparison with the Poisson-Boltzmann approach. *J Phys Chem B* 110:9304–9313
- Livermore DM (2009) Has the era of untreatable infections arrived? *J Antimicrob Chemother* 64:i29–i36
- Lo Conte L, Brenner SE, Hubbard TJ, Chothia C, Murzin AG (2002) SCOP database in 2002: refinements accommodate structural genomics. *Nucleic Acids Res* 30:264–267
- Lonsdale R, Mulholland AJ (2014) QM/MM modelling of drug-metabolizing enzymes. *Curr Top Med Chem* 14:1339–1347
- Lyne PD, Lamb ML, Saeh JC (2006) Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *J Med Chem* 49:4805–4808
- MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616
- Madan AK, Bajaj S, Dureja H (2013) Classification models for safe drug molecules. In: Reisfeld B, Mayeno AN (eds) *Computational toxicology*, vol 930. Humana Press, New York, pp 99–124
- Mahoney MW, Jorgensen WL (2001) A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J Chem Phys* 112:8910–8922
- Malabarba A, Goldstein BP (2005) Origin, structure, and activity in vitro and in vivo of dalbavancin. *J Antimicrob Chemother* 55:15–20
- Mandal S, Moudgil M, Mandal SK (2009) Rational drug design. *Eur J Pharmacol* 625:90–100
- Marsac PJ, Li T, Taylor LS (2009) Estimation of drug-polymer miscibility and solubility in amorphous solid dispersions using experimentally determined interaction parameters. *Pharm Res* 26:139–151
- Martin EJ, Hoeffel TJ (2000) Oriented Substituent Pharmacophore PROPERT Space (OSPPREYS): a substituent-based calculation that describes combinatorial library products better than the corresponding product-based calculation. *J Mol Graph Model* 18:383–403
- Mason JS, Beno BR (2000) Library design using BCUT chemistry-space descriptors and multiple four-point pharmacophore fingerprints: simultaneous optimization and structure-based diversity. *J Mol Graph Model* 18:438–451
- Mason JS, Cheney DL (2000) Library design and virtual screening using multiple 4-point pharmacophore fingerprints. *Pac Symp Biocomput* 5:576–587
- McGann MR, Almond HR, Nicholls A, Grant JA, Brown FK (2003) Gaussian docking functions. *Biopolymers* 68:76–90
- Meller J, Elber R (2001) Linear programming optimization and a double statistical filter for protein threading protocols. *Proteins: Struct Funct Bioinf* 45:241
- Meng X-Y, Zhang H-X, Mezei M, Cui M (2011) Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des* 7(2):146–157
- Menichetti F (2005) Current and emerging serious Gram-positive infections. *Clin Microbiol Infect* 11(s3):22–28
- Menikarachchi LC, Gascon JA (2010) QM/MM approaches in medicinal chemistry research. *Curr Top Med Chem* 10:46–54
- Miao X, Waddell PJ, Valafar H (2008) TALI: local alignment of protein structures using backbone torsion angles. *J Bioinform Comput Biol* 6:163–181
- Michel J, Verdonk ML, Essex JW (2006) Protein-ligand binding affinity predictions by implicit solvent simulations: a tool for lead optimization? *J Med Chem* 49:7427–7439
- Michel J, Tirado-Rives J, Jorgensen WL (2009) Prediction of the water content in protein binding sites. *J Phys Chem B* 113:13337–13346
- Milburn D, Laskowski RA, Thornton JM (1998) Sequence annotated by structure: a tool facilitate the use of structural information in sequence analysis. *Protein Eng* 11:855–859

- Miller MS, Lay WK, Li S, Hacker WC, An J, Ren J, Elcock AH (2017) Reparameterization of protein force field nonbonded interactions guided by osmotic coefficient measurements from molecular dynamics simulations. *J Chem Theory Comput* 13:1812–1826
- Miyazaki T, Aso Y, Yoshioka S, Kawanishi T (2011) Differences in crystallization rate of niten-dipine enantiomers in amorphous solid dispersions with HPMC and HPMCP. *Int J Pharm* 407:111–118
- Modi S, Hughes M, Garrow A, White A (2012) The value of in silico chemistry in the safety assessment of chemicals in the consumer goods and pharmaceutical industries. *Drug Discov Today* 17:134–142
- Moellering RC, Linden PK, Reinhardt J, Blumberg EA, Bompart F (1999) The efficacy and safety of quinupristin/dalfopristin for the treatment of infection caused by vancomycin-resistant *Enterococcus faecium*. *J Antimicrob Chemother* 44:251–261
- Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 19:1639–1662
- Muegge I (2006) PMF scoring revisited. *J Med Chem* 49:5895–5902
- Mustata G, Follis AV, Hammoudeh DI, Metallo SJ, Wang H, Prochownik EV, Lazo JS, Bahar I (2009) Discovery of novel Myc-Max heterodimer disruptors with a three-dimensional pharmacophore model. *J Med Chem* 52:1247–1250
- Myriantopoulos V, Gaboriaud-Kolar N, Tallant C, Hall ML, Grigoriou S, Brownlee PM, Fedorov O, Rogers C, Heidenreich D, Wainer M, Drosos N, Mexia N, Savitsky P, Bagratuni T, Kastiris E, Terpos E, Mikros E (2016) Discovery and optimization of a selective ligand for the switch/sucrose nonfermenting-related bromodomains of polybromo protein-1 by the use of virtual screening and hydration analysis. *J Med Chem* 59(19):8787–8803
- Nelson DL, Lehninger AL, Cox MM (2008) *Lehninger principles of biochemistry*. W.H. Freeman, New York
- Neumann D, Lehr C-M, Lenhof H-P et al (2004) Computational modeling of the sugar-lectin interaction. *Adv Drug Deliv Rev* 56:437–457
- Nichols RL, Graham DR, Barriere SL, Rodgers A, Wilson SE, Zervos M, Dunn DL, Kreter B (1999) Treatment of hospitalized patients with complicated Gram-positive skin and skin structure infections: two randomized, multicentre studies of quinupristin/dalfopristin versus cefazolin, oxacillin or vancomycin. Synercid Skin and Skin Structure Infection Group. *J Antimicrob Chemother* 44:263–273
- Nikaido H (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* 67:593–656
- OECD (2015) The OECD QSAR toolbox. Available at: <http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm>
- Orengo CA, Bray JE, Buchan DW, Harrison A, Lee D (2002) The CATH protein family database: a resource for structural and functional annotation of genomes. *Proteomics* 2:11–21
- Österberg F, Morris GM, Sanner MF, Olson AJ, Goodsell DS (2002) Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins* 46:34–40
- Pallen MJ, Loman NJ, Penn CW (2010) High-throughput sequencing and clinical microbiology: progress, opportunities and challenges. *Curr Opin Microbiol* 13:625–631
- Palumbi SR (2001) Humans as the world's greatest evolutionary force. *Science* 293:1786–1790
- Peitsch MC, Schwede T, Guex N (2000) Automated protein modelling—the proteome in 3D. *Pharmacogenomics* 1:257–266
- Pertel PE, Bernardo P, Fogarty C, Matthews P, Northland R, Benvenuto M, Thorne GM, Luperchio SA, Arbeit RD, Alder J (2008) Effects of prior effective therapy on the efficacy of daptomycin and ceftriaxone for the treatment of community-acquired pneumonia. *Clin Infect Dis* 46(8):1142–1151
- Peters JM, King RW, Höög C, Kirschner MW (1996 Nov 15) Identification of BIME as a subunit of the anaphase-promoting complex. *Science* 274(5290):1199–1201
- Peters KN, Dixon DM, Holland SM, Fauci AS (2008) The research agenda of the National Institute of Allergy and Infectious Diseases for antimicrobial resistance. *J Infect Dis* 197:1087–1093

- Petersen RK, Christensen KB, Assimopoulou AN, Fretté X, Papageorgiou VP, Kristiansen K, Kouskoumvekaki I (2011) Pharmacophore-driven identification of PPAR γ agonists from natural sources. *J Comput Aided Mol Des* 25:107–116
- Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802
- Pieper U, Eswar N, Ilyin VA, Stuart A, Sali A (2002) ModBase, a database of annotated comparative protein structure models. *Nucleic Acids Res* 30:255–259
- Pitt WR, Calmiano MD, Kroeplien B, Taylor RD, Turner JP, King MA (2013) Structure-based virtual screening for novel ligands. *Methods Mol Biol* 1008:501–519
- Poddar RK, Rakha P, Singh SK, Mishra DN (2010) Bioadhesive polymers as a platform for drug delivery: possibilities and future trends. *Res J Pharm Dosage Forms Technol* 1:40
- Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 29:845–854
- Schrödinger Suite 2014-1 Protein Preparation Wizard, Epik, version 2.7; Schrödinger, LLC: New York, NY, USA, 2013
- Raies AB, Bajic VB (2016) In silico toxicology: computational methods for the prediction of chemical toxicity. *Wiley Interdiscip Rev Comput Mol Sci* 6:147–172
- Raizada A, Bandari A, Kumar B (2010) Polymers in drug delivery: a review. *Int J Pharm Res Dev* 2:9–20
- Rajamani R, Good AC (2007) Ranking poses in structure-based lead discovery and optimization: current trends in scoring function development. *Curr Opin Drug Discov Devel* 10:308–315
- Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7:95–99
- Rarey M, Kramer B, Lengauer T, Klebe GA (1996) Fast flexible docking method using an incremental construction algorithm. *J Mol Biol* 261:470–489
- Rask-Andersen M, Almén MS, Schiöth HB (2011) Trends in the exploitation of novel drug targets. *Nat Rev Drug Discov* 10:579–590
- Reddy PD, Swarnalatha D (2010) Recent advances in novel drug delivery systems. *Int J PharmTech Res* 2:2025–2027
- Retief JD (2000) Phylogenetic analysis using PHYLIP. *Methods Mol Evol* 132:143–158
- Ripphausen P, Nisius B, Peltason L, Bajorath J (2010) Quo vadis, virtual screening? A comprehensive survey of prospective applications. *J Med Chem* 53:8461–8467
- Rocchia W, Sridharan S, Nicholls A, Alexov E, Chiabrera A, Honig B (2002) Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. *J Comput Chem* 23:128–137
- Rohs R, Bloch I, Sklenar H, Shakked Z (2005) Molecular flexibility in ab-initio drug docking to DNA: binding-site and binding-mode transitions in all-atom Monte Carlo simulations. *Nucleic Acids Res* 33:7048–7057
- Rost B (2001) Review: protein secondary structure prediction continues to rise. *J Struct Biol* 134:204–218
- Rost B, Sander C (1995) Progress of 1D protein structure prediction at last. *Proteins* 23(3):295–300
- Rowe PH (2010) Statistical methods for continuous measured endpoints in in silico toxicology. In *in silico toxicology: Principles and applications* (pp. 228–251). Cambridge, UK: The Royal Society of Chemistry
- Roy H, Brahma CK, Kumar R, Nandi S (2013) Formulation of saquinavir mesylate loaded microparticle by counterion induced aggregation method: approach by hyperosmotic technique. *Drug Invention Today* 5:259–266
- Ruiz-Carmona S, Alvarez-Garcia D, Foloppe N (2014) rDock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS Comput Biol* 10:e1003571
- Rybak MJ, Hershberger E, Moldovan T, Gruz RG (2000) In vitro activities of daptomycin, vancomycin, linezolid, and quinupristin-dalfopristin against Staphylococci and Enterococci, including vancomycin intermediate and – resistant strains. *Antimicrob Agents Chemother* 44:1062–1066
- Rychlewski L, Jaroszewski L, Li W, Godzik A (2000) Comparison of sequence profiles: strategies for structural predictions using sequence information. *Protein Sci* 9:232–241

- Salum L, Polikarpov I, Andricopulo AD (2008) Structure-based approach for the study of estrogen receptor binding affinity and subtype selectivity. *J Chem Inf Model* 48:2243–2253
- Santos Filho OA, Alencastro RB (2003) Modelagem de proteínas por homologia. *Quím Nova* 26:253–259
- Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des* 27:221–234
- Schäffer AA, Aravind L, Madden TL, Shavirin S, Spouge JL, Wolf YI, Koonin EV, Altschul SF (2001) Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. *Nucleic Acids Res* 29:2994–3005
- Schapira M, Abagyan R, Totrov M (2003) Nuclear hormone receptor targeted virtual screening. *J Med Chem* 46:3045–3059
- Schmaljohann D (2006) Thermo- and pH-responsive polymers in drug delivery. *Adv Drug Deliv Rev* 58:1655–1670
- Schulz-Gasch T, Stahl M (2004) Scoring functions for protein-ligand interactions: a critical perspective. *Drug Discov Today Technol* 1:231–239
- Schwartz R, Istrail S, King J (2001) Frequencies of amino acid strings in globular protein sequences indicate suppression of blocks of consecutive hydrophobic residues. *Protein Sci* 10:1023–1031
- Schwede T, Kopp J, Guex N, Peitsch MC (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res* 31:3381–3385
- Sen S, Farooqui NA, Easwari TS, Roy B (2012) CoMFA-3D QSAR approach in drug design. *Int. J. Res. Dev. Pharm. L. Sci* 1:167–175
- Shaik S, Cohen S, Wang Y (2010) P450 enzymes: their structure, reactivity, and selectivity-modeled by QM/MM calculations. *Chem Rev* 110:949–1017
- Shan Y, Wang G, Zhou H (2001) Fold recognition and accurate query-template alignment by a combination of PSI-BLAST and threading. *Proteins: Struct Funct Bioinf* 42:23–37
- Shelley JC, Chollet A, Frye LL, Greenwood JR, Timlin MR, Uchimaya M (2007) Epik: a software program for pK (a) prediction and protonation state generation for drug-like molecules. *J Comput Aided Mol Des* 21:681–691
- Shi Y (2014) A glimpse of structural biology through X-ray crystallography. *Cell* 159:995–1014
- Shoichet BK, Stroud RM, Santi DV, Kuntz ID, Perry KM (1993) Structure-based discovery of inhibitors of thymidylate synthase. *Science* 259:1445–1450
- Silverman JA, Perlmutter NG, Shapiro HM (2003) Correlation of daptomycin bactericidal activity and membrane depolarization in *Staphylococcus aureus*. *Antimicrob Agents Chemother* 47:2538–2544
- Sippl MJ (1993) Recognition of errors in three-dimensional structures of proteins. *Proteins* 14:355–362
- Sippl MJ, Weitckus S (1992) Detection of native-like models for amino acid sequences of unknown three-dimensional structure in a data base of known protein conformations. *Proteins: Struct Funct Bioinf* 13:258–271
- Skolnick J, Kihara D, Zhang Y (2004) Development and large scale benchmark testing of the PROSPECTOR 3 threading algorithm. *Proteins* 56:502–518
- Sliwoski G, Kothiwale S, Meiler J, Lowe EW (2013) Computational methods in drug discovery. *Pharmacol Rev* 66:334–395
- Spellberg B, Powers JH, Brass EP, Miller LG, Edwards JE Jr (2004) Trends in antimicrobial drug development: implications for the future. *Clin Infect Dis* 38:1279–1286
- Springer J, Saffley M, Huber S, Troxclair D, Craver R, Newman WP (2010) Histopathological findings in fatal novel H1N1: an autopsy case series from September–November 2009 in New Orleans. *Louisiana J La State Med Soc* 162:88–91
- Srivastava S, Srivastava SK, Singh RB, Mishra S (2009) Computational characterization of proteins involved in banana (*Musa acuminata*) ripening. *Open Nutraceuticals J* 2:122–126
- Steenbergen JN, Alder J, Thorne GM, Tally FP (2005) Daptomycin: a lipopeptide antibiotic for the treatment of serious Gram-positive infections. *J Antimicrob Chemother* 55:283–288

- Sung JH, Srinivasan B, Esch MB, McLamb WT, Bernabini C, Shuler ML, Hickman JJ (2014) Using PBPK guided “Body-on-a-Chip” systems to predict mammalian response to drug and chemical exposure. *Exp Biol Med* 239:1225–1239
- Talele TT, Khedkar SA, Rigby AC (2010) Successful application of computer aided drug discovery: moving drugs from concept to the clinic. *Curr Top Med Chem* 10:127–141
- Tannor DJ (2008) Introduction to quantum mechanics: a time-dependent perspective. University Science Books, Sausalito
- ten Brink T, Exner TE (2010) pK(a) based protonation states and microspecies for protein-ligand docking. *J Comput Aided Mol Des* 24:935–942
- Teodorescu O, Galor T, Pillardy J, Elber R (2004) Enriching the sequence substitution matrix by structural information. *Proteins: Struct Funct Bioinf* 54:41
- Thomas CM, Nielsen KM (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat Rev Microbiol* 3:711–721
- Thompson DC, Humblet C, Joseph-McCarthy D (2008) Investigation of MM-PBSA rescoring of docking poses. *J Chem Inf Model* 48:1081–1091
- Tobi D, Elber R (2000) Distance-dependent, pair potential for protein folding: results from linear optimization. *Proteins: Struct Funct Bioinf* 41:40
- Tollman PA (2001) Revolution in R&D: how genomics and genetics are transforming the biopharmaceutical industry. Boston Consulting Group, Boston
- Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31:455–461
- Umasankar V, Gurunathan S (2015) Drug discovery: an Appraisal. *Int J Pharm Pharm Sci* 4:59–66
- Valerio LG Jr (2009) In silico toxicology for the pharmaceutical sciences. *Toxicol Appl Pharmacol* 241:356–370
- van der Kamp MW, Mulholland AJ (2013) Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry* 52:2708–2728
- Van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, HJC B (2005) GROMACS: fast, flexible, and free. *J Comput Chem* 26:1701–1718
- Venkatachalam CM, Jiang X, Oldfield T, Waldman M (2003) LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J Mol Graph Model* 21:289–307
- Venkatapathy R, Wang NCY (2013) Developmental toxicity prediction. In: Reisfeld B, Mayeno AN (eds) *Computational toxicology*, vol 930. Humana Press, New York, pp 305–340
- Verma RP, Hansch C (2009) Camptothecins: a SAR/QSAR study. *Chem Rev* 109:213–235
- Vyas V, Jain A, Jain A, Gupta A (2008) Virtual screening: a fast tool for drug design. *Sci Pharm* 76:333–360
- Vyas VK, Ukawala RD, Ghate M, Chintha C (2012) Homology modeling a fast tool for drug discovery: current perspectives. *Indian J Pharm Sci* 74(1):1–17
- Walsh C (2003) Antibiotics: actions, origins, resistance. American Society for Microbiology (ASM) Press, Washington, DC
- Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. *J Comput Chem* 25:1157–1174
- Wang Y, Shaikh SA, Tajkhorshid E (2010) Exploring transmembrane diffusion pathways with molecular dynamics. *Physiology* 25:142–154
- WaterMap (2014) New York, NY: Schrödinger, LLC
- Westbrook J, Feng Z, Jain S, Bhat TN, Thanki N, Veerasamy R, Gilliland GL, Bluhm WF, Weissing H, Greer DS, Bourne PE, Berman HM (2002) The protein data bank: unifying the archive. *Nucleic Acids Res* 30:245–248
- Wheeler SR, Kearney JB, Guardiola AR, Crews ST (2006) Single-cell mapping of neural and glial gene expression in the developing *Drosophila* CNS midline cells. *Dev Biol* 294:509–524
- Willett P (2006) Similarity-based virtual screening using 2D fingerprints. *Drug Discov Today* 11(23–24):1046–1053
- Wilson GL, Lill MA (2011) Integrating structure-based and ligand-based approaches for computational drug design. *Future Med Chem* 3:735–750

- Winkler DA (2002) The role of quantitative structure-activity relationships (QSAR) in biomolecular discovery. *Brief Bioinform* 3:73–86
- Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z, Woolsey J (2006) DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res* 34(Database issue):D668–D672
- Wolber G, Langer T (2005) LigandScout: 3-D pharmacophores derived from protein bound ligands and their use as virtual screening filters. *J Chem Inf Model* 45:160–169
- Wood MJ, Hirst JD (2005) Protein secondary structure prediction with dihedral angles. *Proteins* 59:476–481
- Woodfor N, Livermore DM (2009) Infections caused by Gram-positive bacteria: a review of the global challenge. *J Infect* 59:S4–S16
- Wootton M, Howe RA, Walsh TR, Bennett PM, Macgowan AI (2000) In-vitro activity of a range of old and new antimicrobials to hetero-vancomycin-intermediate *Staphylococcus aureus* (hVISA) [abstract 23131. In: Program and abstract of the 40th Interscience Conference on Antimicrobial agents and chemotherapy; September 17–20, Toronto, Canada
- Wu S, Skolnick J, Zhang Y (2007) Ab initio modeling of small proteins by iterative TASSER simulations. *BMC Biol* 5:17
- Xiang Z (2006) Advances in homology protein structure modeling. *Curr Protein Pept Sci* 7:217–227
- Xu Y, Xu D, Uberbacher EC (1998) An efficient computational method for globally optimal threading. *J Comput Biol* 5:597–614
- Yan R, Xu D, Yang J, Walker S, Zhang Y (2013) A comparative assessment and analysis of 20 representative sequence alignment methods for protein structure prediction. *Sci Rep* 3:2619
- Yang WW, Pierstorff E (2012) Reservoir-based polymer drug delivery systems. *J Lab Autom* 17:50–58
- Yang L, Wang W, Sun Q, Xu F, Niu Y, Wang C, Liang L, Xu P (2016) Development of novel proteasome inhibitors based on phthalazinone scaffold. *Bioorg Med Chem Lett* 26:2801–2805
- Yasuo K, Yamaotsu N, Gouda H, Tsujishita H, Hirono S (2009) Structure-based CoMFA as a predictive model – CYP2C9 inhibitors as a test case. *J Chem Inf Model* 49:853–864
- Yona G, Levitt M (2002) Within the twilight zone: a sensitive profile-profile comparison tool based on information theory. *J Mol Biol* 315:1257–1275
- Yoo SU, Krill SL, Wang Z, Telang C (2009) Miscibility/stability considerations in binary solid dispersion systems composed of functional excipients towards the design of multi-component amorphous systems. *J Pharm Sci* 98:4711–4723
- Young T, Abel R, Kim B, Berne BJ, Friesner RA (2007) Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding. *Proc Natl Acad Sci U S A* 104:808–813
- Yuriev E, Agostino M, Ramsland PA (2011) Challenges and advances in computational docking: 2009 in review. *J Mol Recognit* 24:149–164
- Zhang Y, Skolnick J (2004) Automated structure prediction of weakly homologous proteins on a genomic scale. *Proc Natl Acad Sci U S A* 101:7594–7599
- Zhang Y, Arakaki A, Skolnick J (2005) TASSER: an automated method for the prediction of protein tertiary structures in CASP6. *Proteins* 61:91–98
- Zhang W, Liu S, Zhou Y (2008) SP5: improving protein fold recognition by using torsion angle profiles and profile-based gap penalty model. *PLoS One* 3:e2325
- Zhang M, Li H, Lang B, O'Donnell K, Zhang H, Wang Z, Dong Y, Wu C, Williams RO (2012) Formulation and delivery of improved amorphous fenofibrate solid dispersions prepared by thin film freezing. *Eur J Pharm Biopharm* 82:534–544
- Zhao H, Caflisch A (2013) Discovery of ZAP70 inhibitors by high-throughput docking into a confirmation of its kinase domain generated by molecular dynamics. *Bioorg Med Chem Lett* 23:5721–5726
- Zhou H, Skolnick J (2007) Ab initio protein structure prediction using chunk-TASSER. *Biophys J* 93:1510–1518
- Zhou H, Zhou Y (2005) Fold recognition by combining sequence profiles derived from evolution and from depth-dependent structural alignment of fragments. *Proteins* 58:6321–6328