

Novel computational biology methods and their applications to drug discovery

Sharangdhar S. PHATAK^{1,2,*}, Hoang T. TRAN^{2,*}, Shuxing ZHANG (✉)²

¹ School of Biomedical Informatics, The University of Texas Health Science Center, 7000 Fannin Street, Houston, TX 77030, USA

² The Integrated Molecular Discovery Laboratory, Department of Experimental Therapeutics, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030, USA

© Higher Education Press and Springer-Verlag Berlin Heidelberg 2011

Abstract Computational biology methods are now firmly entrenched in the drug discovery process. These methods focus on modeling and simulations of biological systems to complement and direct conventional experimental approaches. Two important branches of computational biology include protein homology modeling and the computational biophysics method of molecular dynamics. Protein modeling methods attempt to accurately predict three-dimensional (3D) structures of uncrystallized proteins for subsequent structure-based drug design applications. Molecular dynamics methods aim to elucidate the molecular motions of the static representations of crystallized protein structures. In this review we highlight recent novel methodologies in the field of homology modeling and molecular dynamics. Selected drug discovery applications using these methods conclude the review.

Keywords computational biology, drug discovery, homology modeling, molecular dynamics, structure-based drug design

Introduction

The impact of computational methods in modern drug discovery process cannot be over emphasized. Early stages of drug discovery routinely use bioinformatics methods to efficiently analyze genomic data (Yao et al., 2009). Biomolecular modeling methods are being successfully applied to lead identification and optimization (Coumar et al., 2010; Jamieson et al., 2010; Muddassar et al., 2010; Murumkar et al., 2010; Ogata et al., 2010; Postigo et al., 2010; Talele et al., 2010). Pharmacokinetic/pharmacodynamic/pharmacometric methods model biological/physiological responses of drugs and play an important role in the decision making part of the drug development phase (Rajman, 2008). Last but not the least, computational methods help design optimized clinical trials (Holford et al., 2010). Computational methods are increasingly used to complement experimental

methods (Phatak et al., 2009) and have proved to reduce costs and time without compromising the quality of the drug discovery process. Such computational methods differ in their genesis (e.g. computer science, physics, mathematics) but share a common goal: solving biological problems. This commonality allowed researchers to encompass these methods in an interdisciplinary field termed computational biology. Computational biologists develop theoretical and/or knowledge-based predictive models of the biological system of interest among other applications (Kiran et al., 2008; Laursen, 2009). Protein modeling is one such important method that attempts to construct near-native 3D models of proteins that do not have an experimentally solved structure. The development of protein modeling methods is an actively pursued area of research because: (a) experimental methods may never provide structures for every protein (Baker and Sali, 2001), and (b) such models are required for the increasingly applied structure-based methods for drug discovery (Barcellos et al., 2008; Cavasotto and Phatak, 2009; Daga et al., 2010). However, these structures (modeled or experimental) are mere snapshots. The native state of a protein is not a static structure, but consists of an ensemble of structures that interconvert in a dynamic equilibrium. These spontaneous interconversions range from vibrations of

Received December 3, 2010; accepted January 4, 2011

Correspondence: Shuxing ZHANG

E-mail: shuzhang@mdanderson.org

* Both authors contributed equally to the paper

individual bonds to domain rearrangements. In addition to passive action, there is also evidence that these fluctuations actively drive processes including substrate binding and ion channel gating (Bahar et al., 2010). The widespread use of molecular dynamics in the study of biological molecules represents an acknowledgment of the critical importance of molecular motions in biological function. This review primarily provides a succinct background of homology modeling and molecular dynamics methods, their recent developments and selected applications in the drug discovery field.

Protein modeling

Overview

3D protein structures potentially rationalize the biology (molecular and system level) of a validated drug target (Cozzetto and Tramontano, 2008), and have been successfully applied to diverse structure-based drug design projects (Dessailly et al., 2009; Grant, 2009; Myler et al., 2009). However, the key to their applicability is their availability. Pragmatic approaches (e.g., focusing on solving representative protein structure per therapeutically important target family (Dessailly et al., 2009)) and technological innovations (X-ray crystallography and NMR techniques (Montelione and Szyperski, 2010)) from several structural genomics projects (Weigelt, 2010), though successful, have failed to narrow the ever increasing available-sequence:available-structure gap (Malmström and Goodlett, 2010). Computational modeling of proteins offers an alternative vista in the pursuit of 3D protein structures and may be considered reasonable substitutes until a corresponding experimental structure is determined.

Protein modeling can be broadly divided into two main categories: template-based and template-free modeling. Figure 1 demonstrates a representative workflow for protein modeling methods. Template-based models exploit the rationale that similar sequences share similar protein structures. Homology modeling (Xiang, 2006; Cavasotto and Phatak, 2009; Daga et al., 2010) is an avant-garde method in template-based modeling though threading (Peng and Xu, 2010; Trojanowski et al., 2010; Zhou and Skolnick, 2010)/fold-recognition approaches (Deschavanne and Tufféry, 2009; Lampros et al., 2009; Lobley et al., 2009; Yan et al., 2009; Ying et al., 2009) are being increasingly developed and applied. Homology modeling follows a well-established protocol to develop models for proteins with unknown structures (targets): identifying homologous protein structure(s) (template(s)), aligning the target and template sequences, building and subsequently refining models. Threading or fold recognition approaches exploit the finite number of protein structure folds in nature. Scoring functions are developed based on the known relationships between sequences and structures. Subsequently, models are built based on that

template which provides a maximum score post sequence alignment. This approach has shown promising results in critical assessment of techniques for protein structure prediction (CASP) competitions, particularly in its ability to detect remote homologs and target-sequence alignment (Kryshtafovych and Fidelis, 2009). It is routine to obtain models with 1–2 Å deviation from the experimental structure for closely homologous sequences (Ginalski, 2006; Burley et al., 2008). Automated servers based on template-based modeling methods now provide good quality models with minimal human intervention for simple cases (Kim et al., 2004; Grünberg et al., 2007; Bordoli et al., 2009; Hildebrand et al., 2009; Roy et al., 2010).

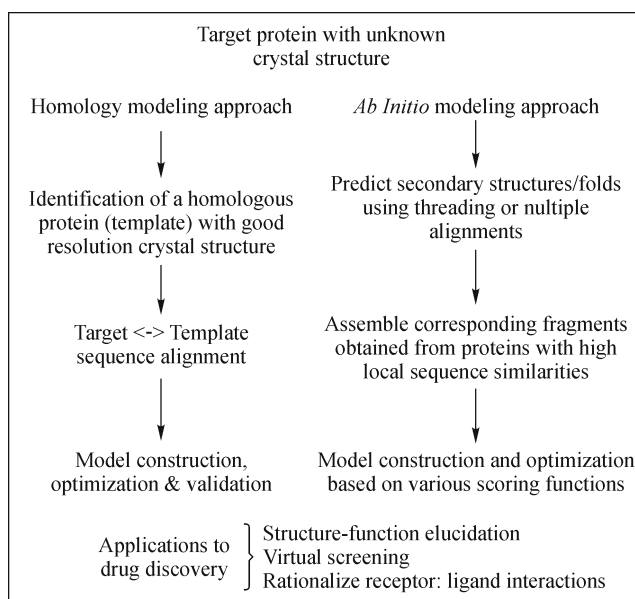


Figure 1 A representative workflow for protein modeling methods

In the absence of any template (homologs or not), *de novo* methods using knowledge and/or physics-based potentials are used to construct 3D structural models. Free, *ab initio* or *de novo* modeling is the umbrella terms for structure prediction methods that attempt to construct protein models using knowledge and/or physics-based approaches when no structural homolog exists. Rosetta (Das and Baker, 2008; Kaufmann et al., 2010) uses a knowledge-based approach to construct initial models, which in turn are refined by all-atom physics-based approaches (Bradley et al., 2005; Das et al., 2007). Threading assembling refinement (TASSER) from Wu et al. (2007) is another promising method which uses only knowledge-based approaches for *de novo* protein modeling. Presently, *de novo* model building exercises are limited to smaller proteins (120–150 amino acids) and rarely obtain models < 2 Å resolution (Zhang, 2008). However in either case the caveat is that protein modeling, by definition, is error-prone. This review will highlight selected recent comparative modeling and refinement methods.

Refinement and optimization of protein models

Comparative modeling often introduces errors due to: e.g., (a) choice of template(s) and sequence alignment (Larsson et al., 2008); (b) low sequence identity ($< 30\%$) between the target and template sequences (G-protein coupled receptors are an exception due to their structural signatures); and (c) flexible loop regions, particularly when involved in ligand binding. A plethora of bioinformatics approaches have been developed to minimize errors associated with the identification of homologous proteins from protein databases such as the Protein Data Bank (PDB) and subsequent sequence alignment. These include techniques based on pairwise sequence comparison (Dunbrack, 2006; Lu et al., 2008; Jayalakshmi et al., 2010), machine learning methods (Lee et al., 2008) e.g., hidden Markov models (Bystruff and Krogh, 2008; Karplus, 2009), neural networks (Mooney and Pollastri, 2009; Walsh et al., 2009), and support vector machines (Fariselli et al., 2007) among others. They attempt to maximize the sensitivity (i.e., the ability to identify distant relationships) and selectivity (i.e., to return a minimal set of false positives). Even then the constructed comparative model may be inaccurate.

Modeling errors, particularly incorrect side chain orientations in binding pockets, can negate their applicability in drug discovery applications. Two themes dominate the current methodologies of refining comparative models. In the first, known ligand information is used to optimize the pharmaceutically relevant binding pocket. In the MOBILE approach developed by Evers et al. (2003), spatial knowledge of ligands placed into an averaged binding site of crude ensemble of protein models is used as restraints to generate a second generation of refined homology models. The method was successfully applied to identify sub-micromolar inhibitors for the neurokinin-1 receptor (Evers and Klebe, 2004). Moro and colleagues used a soft-docking approach to generate initial ligand poses within the binding pocket and reconstructed the homology model of A3 adenosine receptors with optimized side-chain orientations to accommodate the ligand pose without the soft-docking approach (Moro et al., 2006). Continuing on this theme, but by incorporating a docking-based Monte Carlo approach and full receptor:ligand flexibility, Cavasotto et al. developed and applied the ligand-steered modeling method in multiple structure-based drug discovery projects targeting melanin concentrating hormone receptor 1 (Cavasotto et al., 2008) and cannabinoid receptor 2 (Diaz et al., 2009a, 2009b) (Fig. 2). Recently, Phatak et al. (2010) extended and benchmarked the ligand-steered modeling method by retrospectively cross modeling Class A G-protein coupled receptors of Beta2-adrenergic and Adenosine-2A. The models obtained near-native co-crystallized ligand poses, showed promising selectivity profiles and enrichment rates as compared to crystal structures and crude homology models respectively. Radestock et al.

(2008) manually refined the side-chain orientations to map the corresponding receptor:ligand interactions in the metabotropic glutamate receptor 5. These were then encoded as a simple binary scoring function and compared against conventional scoring functions in a retrospective virtual screening exercise. Neves et al. (2010) optimized receptor binding pockets of CXCR4, an anti HIV target, using diverse chemotype ligands in a cyclic combination of elastic network normal mode analysis for backbone/side chain sampling and full flexible receptor:ligand docking for model selection.

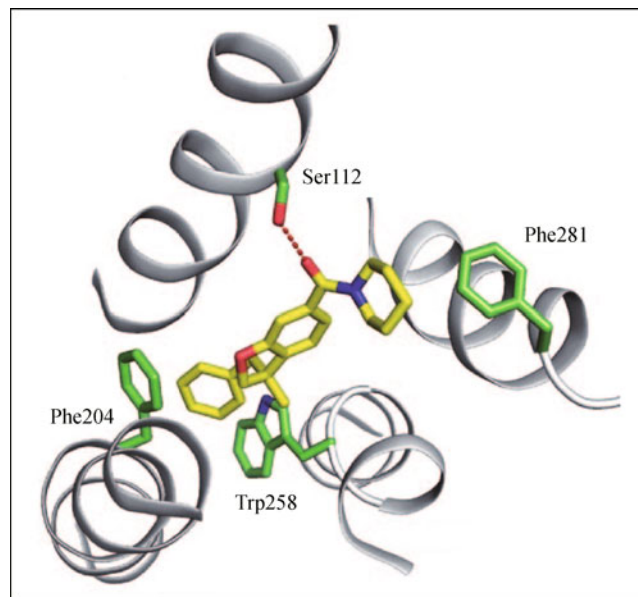


Figure 2 Cannabinoid receptor 2 complexed with N-alkyl-isatin acylhydrazone derivative (yellow carbon atoms) identified through protein homology modeling and structure-based library screening. Transmembrane regions are shown as white ribbons. Hydrogen bonds are represented as red dashes.

The second theme for refinement uses molecular dynamics (MD) methods. Michino et al. (2010) developed a two-step method for the seven transmembrane (7-TM) GPCRs termed foldGPCR. Using tertiary restraints from the template structure, and simulated annealing process the 7-TMs of beta-2 adrenergic receptor and its known ligand carazolol, is assembled in the first step. An ensemble of complexes is generated by random rotation of the helices and x -axis translation of the ligand. The averaged structure is then refined using a combination of generalized Born (GB) implicit membrane model and replica exchange sampling methods. The method retrospectively predicted low RMSD beta2-adrenergic model derived from a rhodopsin template (TM RMSD ~ 2 Å and ligand RMSD ~ 1 Å). Zhu et al. (2008) applied temperature-based replica exchange molecular dynamics (REMD) method to refine homology models. First, structural errors in side chains and loop regions of models were corrected using a variety of toolkits. Next, the models were refined using REMD protocol implemented in

the GROMACS package, where the replicas explored a range of temperature space to potentially allow the system to access conformational space that would be prohibited at standard temperatures. Finally, a combination of statistical potentials and scoring functions were used to identify near-native conformations. To offset the disadvantage of using a large number of replicas to cover a given temperature range, Kannan and Zacharias (2010) proposed a biased potential REMD flavor to refine protein models. The biasing potential for the phi and psi angles of the backbone dihedrals permits transitions and thus flexibility of the backbone by lowering energy barriers. This explicit solvent BP-REMD approach showed promising results for loop prediction for structures from the Rosetta decoy set at considerably shorter simulation time frames (~10 ns). Kimura et al. (2008) developed and validated protein models for chemokine receptor 2 (CCR-2 of the GPCR family) using a ligand-free pressure-based steered molecular dynamics method. The binding pocket was optimized by gradually increasing the radii of pre-populated Lennard-Jones particles during the course of the molecular dynamics simulation. To prevent any unreasonable shape of the pocket, the particles were tethered via harmonic bonds to four nearest neighbors and the backbone dihedrals of the transmembrane helices were restrained. Other approaches include the use of: low-frequency normal modes (Rai et al., 2010), a combination of normal modes and molecular dynamics (Stumpff-Kane et al., 2008), knowledge-based refinement methods (Chopra et al., 2010), and restraints derived from cryo electron microscopy maps (Zhu et al., 2010).

Applications of protein models to drug discovery projects

Homology models have been widely applied to a variety of structure-based drug discovery projects. Here, we list a few selected projects that have resulted in active hits against therapeutically relevant proteins. Table 1 shows other applications of homology models to drug discovery.

Emeria tenella CDK-related kinase 2 (EtCRK2) is a validated target against coccidiosis. Using an EtCRK2 homology model and chemoinformatic methods, Engels et al. (2010) identified 293 hits via a virtual screening

exercise (hit rate ~6%). Four compounds were classified as leads and one of them belonged to the benzimidazole carbonitrile class of compounds which were not previously documented as cyclin-dependent kinase activities. Sisay et al. (2010) used a comparative model of matriptase-2, a target for the treatment of systemic iron overload. Side chain conformations of the non conserved residues in the active pocket were manually adjusted by using low energy rotamers and loop regions were modeled to mimic the template protein matriptase-1. This rational medicinal chemistry effort based on pre-existing knowledge of known inhibitors and crystal structure complexes of other matriptase family proteins resulted in two small molecule inhibitors with K_i values of 170 and 460 nmol/L, respectively. NF-kappaB inducing kinase (NIK) is an understudied enzyme involved in the downstream of several tumor necrosis factor receptors implicated in the pathogenesis of rheumatoid arthritis. The lack of potent and selective inhibitors for NIK drove this study, where a previously built homology model of NIK by the same group (Mortier et al., 2010) was used to virtually screen compounds from the ZINC database (Irwin and Shoichet, 2005). The compounds were scored using three scoring functions and ranked based on a consensus ligand efficiency function. Top ranking compounds were visually inspected for the validated hydrogen bond to Leu472 and subsequently tested. One 4*H*-isoquinoline-1,3-dione scaffold-based compound and one of its 16 analogs were found to be micromolar hits (51–91 μ mol/L) (Mortier et al., 2010). Medina-Franco's group screened natural products from the ZINC library (Irwin and Shoichet, 2005) against the homology model of the catalytic domain of DNA methyltransferase, an anti-cancer target. A multi-step docking approach using three docking programs (AutoDock (Goodsell, Morris et al. 1996), Glide (www.schrodinger.com) and Gold (Verdonk et al., 2003)) was developed. 58 hits were identified from ~89K compounds. This was the first reported study involving the virtual screening of a large database of natural products (Medina-Franco et al., 2010). Homology models of histone deacetylase (HDAC) 1 and 6, another oncogenic target, were used to develop potent and selective inhibitors for HDAC 6. The models were developed using iTASSER platform. The structural differences between the

Table 1 Examples of applications of protein modeling to drug discovery and development.

Application	System	References
Structure-function relationship	Nicotinamide mononucleotide adenylyltransferase (NMNAT)	Brunetti et al., 2010
	Fas ligand (FasL) belonging to the tumor necrosis factor (TNF)	Obungu et al., 2009
	Nucleotide pyrophosphatase/phosphodiesterase (NPP)	Duan et al., 2010
	Cohesin chromosome maintenance protein	Kurkcuoglu and Bates, 2010
	<i>E. coli</i> ParE toxin	Barbosa et al., 2010
Rationalize protein:ligand interactions	Dopamine 3, Cannabinoid 2, 5-HT2A receptors	Diaz et al., 2009a, 2009b; Pecic et al., 2010; Kortagere et al., 2011
	Dengue virus NS2B/S3	Wichapong et al., 2010
	Chemokine receptors	Carter and Tebben, 2009

two isozymes (e.g. wider and shallower catalytic channel rim) were used to design carbazole hydroxamic acid based compounds with bulkier and short aromatic moieties to selectively target HDAC 6 (Butler et al., 2010). Several other examples of lead identification on a wide variety of protein targets have been reported elsewhere (Hu et al., 2009; Ostrov et al., 2009; Odell et al., 2010; Sierceki et al., 2010).

Molecular dynamics

Overview

In molecular dynamics simulations, a set of points in 3-dimensional space represents the configuration of the simulated system (e.g. a protein in water). The positions of these points, which represent the molecular conformations, are propagated in time. Snapshots of the system may be taken to form an ensemble of conformations representing possible molecular arrangements. An example of such an ensemble is illustrated in Fig. 3. These points are usually positioned on atomic sites or functional groups, and represent their molecular motions. The interactions between sites are represented by a function which gives the system energy as a function of all the site positions, and is usually called a “force field”. Force field implementations typically define parameters for each site. For instance, a site representing an atom may be described by a charge and van der Waals radius. These parameters must be determined, either using experimental data, or resulting from high-level quantum mechanics simulations. Methods for deriving more accurate force fields are an intense area of current research. Of note for drug discovery is the development of “polarizable” force fields (Jiao et al., 2008; Jiao et al., 2009), in which atomic parameters (specifically the charges) may change in response to changes in environment. This can be expected to improve predictions of binding free energy, where the molecular environment of a ligand may change greatly during binding. For the interested readers, more details of this procedure and other aspects of molecular dynamics can be found in various references (Allen and Tildesley, 1989; Leach, 2001; Frenkel and Smit, 2002). In this section, we describe a representative set of applications of molecular dynamics for therapeutic drug design.

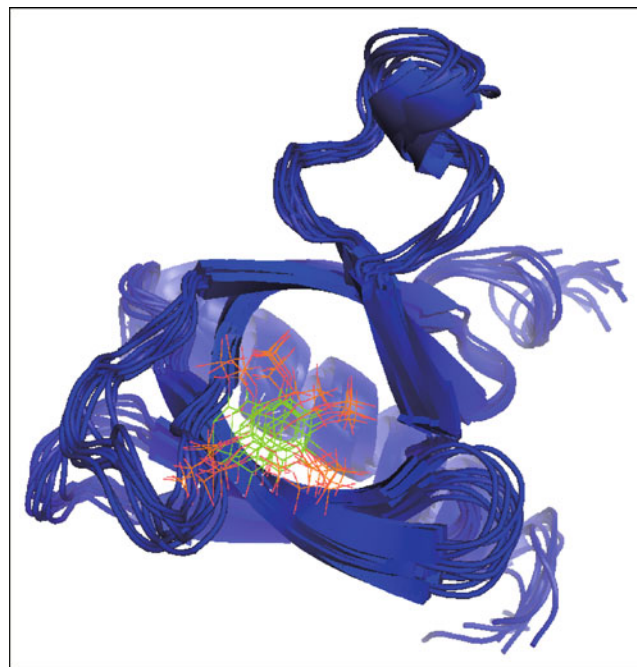


Figure 3 An example of using molecular dynamics to generate an ensemble of alternative receptor-ligand structures. Ten structures generated using molecular dynamics of the pleckstrin homology domain of protein kinase B (blue cartoon display) complexed with inositol 1,3,4,5-tetrakisphosphate (wire display) are shown.

Targeting receptor flexibility

Molecular dynamics may be useful for tackling the challenge of receptor flexibility in structure-based drug design (Marco and Gago, 2007; Morra et al., 2008). Virtual screening using rigid receptor docking potentially overlooks an abundance of drug candidates, so flexible docking is currently an area of intense research (Cozzini et al., 2008; B-Rao et al., 2009). Since molecular dynamics samples conformational flexibility, a number of algorithms based on molecular dynamics have been naturally developed. One such example is from McCammon et al., who have shown that incorporating a short molecular dynamics refinement step into an iterative docking method gives improved results and can find alternative docking poses which otherwise would be missed (Lin et al., 2002). A typical protocol for incorporating molecular dynamics into the docking procedure to obtain

Table 2 Applications of molecular dynamics to drug discovery and development

Application	System	References
Structure-function study	Toll-like receptor homologs	Kubarenko et al., 2007
Alternative ligand poses	Immunophilin FKBP	Fujitani et al., 2005
Alternative ligand poses	HIV-1 reverse transcriptase	Okumura et al., 2010
Receptor pocket flexibility	postsynaptic density-95/Dlg/ZO-1 (PDZ) domains	Gerek and Ozkan, 2010
Membrane-bound proteins	Various (reviews)	Lindahl and Sansom, 2008; Khalili-Araghi et al., 2009
Binding free energy	Immunophilin FKBP	Fujitani et al., 2005
Binding pathways	Imatinib and its targeting kinases c-Kit and Abl.	Yang et al., 2009

multiple poses (of receptor and ligand) is shown in Fig. 4. As a note, even if molecular dynamics simulation is not directly used in the docking process for virtual screening, concepts from molecular dynamics force field development have shown their influence on docking techniques. The problem of calculating the interaction free energy between ligand and receptor is a natural fit for concepts and techniques of force field development. This can be seen in the development of programs such as AutoDock (Goodsell et al., 1996) and APBS (Baker et al., 2001) showing the wide influence of molecular dynamics on the field of molecular modeling.

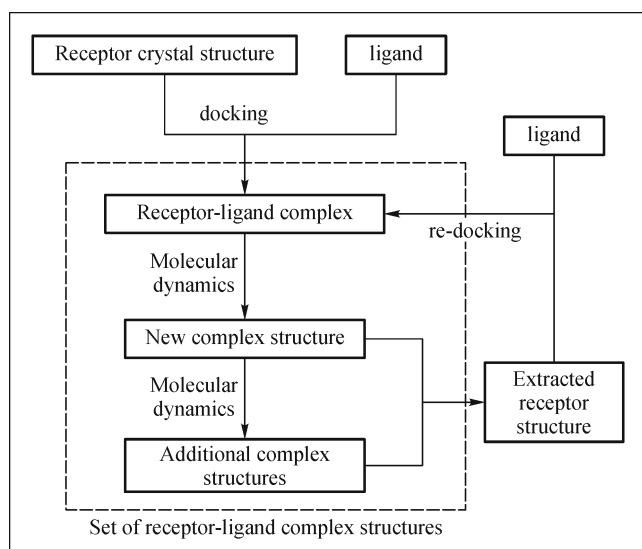


Figure 4 The generic strategy of using molecular dynamics simulations to generate alternative docking poses.

Performing standard molecular dynamics on a receptor may still be too computationally expensive for virtual screening, but variants of molecular dynamics have been developed to reduce the computational cost. One often-used technique is the temperature replica exchange method (Sugita and Okamoto, 1999). Briefly, several replicas of the system of interest are run in parallel using molecular dynamics at different temperatures. Periodic exchanges are performed between replicas such that the correct sampling is achieved at each temperature. This allows for a faster search of configurational space, and the parallel nature of this algorithm is well suited to modern cluster computing architecture. Examples of the application of replica exchange molecular dynamics include the work by Gerek and Ozkan (2010) who, in a ligand binding replica exchange simulation study of postsynaptic density-95/Dlg/ZO-1 PDZ domains, used dihedral restraints on the receptor to induce a bias toward binding site conformations, further reducing the simulation time while still allowing for binding site flexibility. Of course, replica exchange can also be used to efficiently sample ligand populations as well (Okumura et al., 2010). However, due to the computational expense of standard molecular dynamics,

its practical use in high-throughput virtual screening is limited to detailed studies of a limited number of selected compounds, late in the screening pipeline (Okimoto et al., 2010). This may change in the future with improvements in computing hardware.

An alternative approach to achieve biological timescale simulations is the development of MD specific computing hardware (Klepeis et al., 2009; Zwier and Chong, 2010). MDGRAPE-3 developed by RIKEN research is one of the first examples of such hardware. It reportedly performed up to 40 times faster as compared to single conventional processors (Kikugawa et al., 2009). Recently, the group of Shaw et al. (2007) developed specialized parallel computing hardware (Anton) to conduct microsecond simulations at greatly increased speeds. Of late, the inherent architectural features of general-purpose graphical processing units (e.g. high bandwidth memory systems, hardware multithreading, parallel data handling) are exploited for molecular dynamics and other modeling problems (Stone et al., 2010).

Membrane protein simulations

Approximately 30%–50% of current drug targets are membrane proteins (Gruber et al., 2010). However, the experimental structure determination of membrane proteins presents technical challenges (Cherezov et al., 2010) and the current number of unique structures determined is ~300, far behind the number for soluble proteins (White, 2004; Cherezov et al., 2010). Thus, post G-protein coupled receptor (GPCR) structure modeling with molecular dynamics simulations takes on a vital role in structure-function studies of membrane proteins. Molecular dynamics simulations of membrane proteins, e.g., GPCRs, offers an exciting avenue to elucidate details of protein-membrane interactions (Lindahl and Sansom, 2008; Khalili-Araghi et al., 2009). The potential barrier to these atomistic MD studies is the size of such systems and limited timescales (Lindahl and Sansom, 2008). Lately, coarse-grained approaches that aim to reduce the magnitude of interactions, increase the timesteps and speed up the simulation appear to be in vogue (Bond et al., 2008; Sansom et al., 2008). Specially developed interaction parameters and force fields (e.g. MARTINI (Marrink et al., 2007)), hardware and software (Phillips et al., 2005) are steadily contributing to the advances in membrane simulations. Please refer to the excellent reviews on this subject by Khalili-Araghi et al. (2009) and Pandit and Scott (2009).

Protein structure-function analysis

Despite the ever-increasing number of available protein structures, many systems of interest lack structures, either wholly or partially, limiting our ability to perform structure-function analysis. In particular, molecular docking depends heavily on the availability of accurate receptor structures. Perhaps not all residues are seen in a crystallography

experiment, or only the unbound form of a receptor is known. In these cases, molecular dynamics is a useful tool, and often the only reasonable choice, for obtaining a suitable receptor structure for docking (Marco and Gago, 2007). In this context, molecular dynamics is often combined with homology modeling to generate a reasonable structure suitable for further study. A standard protocol may involve, for instance, first generating a homology model followed by energy minimization and molecular dynamics using the GROMACS simulation package (Kutzner et al., 2007). For a good example of the application of this procedure, see the work by Kubarenko et al. (2007) on the generation of structures for toll-like receptor homologs and orthologs. Researchers at Merck used homology modeling, induced fit docking and molecular dynamics to design novel therapeutic agents for steroid hormone binding receptors and performed structure-function relationship studies (Cornell and Nam, 2009).

Free energy calculations

Free energy calculations can determine the difference of free energy between different states (such as the free and bound forms of a receptor-ligand system) or along a reaction path (such as the path for ligand binding). The theoretical foundations are well-established (Shirts and Pande, 2005; Deng and Roux, 2009) and the techniques have been implemented in molecular dynamics packages such as GROMACS (Kutzner et al., 2007). In principle, it is possible to directly calculate free energies of binding for ligand-receptor interactions, but only at great computational expense. One example is the determination of the free energies of binding of eight ligands to immunophilin FKBP using a massively parallel computing platform (Fujitani et al., 2005). The main difficulty lies in the entropic contribution to binding free energies, which requires heavy computational sampling of a large configurational space. Much of the size of this configurational space is due to the solvent (water), so improvement in computational efficiency of free energy calculations can be obtained by using continuum solvation models (which treat water only implicitly) combined with molecular dynamics (Hou et al., 2011). The computational expense has discouraged extensive application to drug design, but we may anticipate these methods will become increasingly important as computational power increases. In fact, this may already be the case, for example the study by Ge and Roux (2010) using all-atom molecular dynamics to calculate free energies of binding for several sparsomycin analogs to the bacterial ribosome.

Determination of binding pathways

It is often fruitful to study the action of known inhibitors to gain insight that can be used to design new inhibitors. This includes not only the binding mode of the ligand, but also the

specific pathway of binding. Such detailed studies can be easily performed with molecular dynamics simulations. For example, Yang et al. (2009) used steered molecular dynamics to directly compare two pathways for the dissociation of a type II kinase inhibitor Imatinib and its receptors c-Kit and Abl. In steered molecular dynamics, external forces are applied to the system, driving it along a pathway of interest. In this case, the ligand was forced to dissociate along two different pathways. Equilibrium information, in particular, the free energies for each pathway, can be recovered using the Jarzynski relation (Liphardt et al., 2002). In this case, the authors were able to determine that the ATP-channel pathway, even though it appears sterically restricted, was more favorable than the allosteric-pocket-channel pathway, due to favorable van der Waals interactions facilitated by the flexibility of residues lining the ATP channel. This information would not be possible to obtain by examination of static structures. In addition, this method has been extended to the comparison of the binding of a series of similar ligands (Colizzi et al., 2010) where the force profiles were shown to distinguish between active and inactive compounds.

Conclusions

Absence of an experimentally determined structure of a therapeutic target does not necessarily rule out structure-based drug discovery efforts. Protein modeling methods can now construct reasonably accurate models or at least the relevant binding sites of yet uncrystallized targets. Their increasing success in the drug discovery and design process provides further impetus to the protein modeling field. Future developments will aim at cases where there is low template: target sequence identity, loop modeling, and also modeling proteins based on just the sequence details. As computational power continues to rise, we may expect molecular dynamics to take on an increasing role in the development of therapeutic agents. Molecular dynamics is a very general simulation method that can potentially provide detailed information about a wide variety of dynamic systems, but can also be easily adapted for specific needs. The on-going development and testing of sophisticated algorithms for molecular dynamics promises to provide increasing and successful returns in the discovery and design of next generation therapeutics.

Acknowledgements

This work was supported in part by the US Department of Defense Concept Awards (BC085871), US National Institute of Health P41 grant (5P41GM079588-03), and Grant # IRG-08-061-01 from the American Cancer Society to SZ. SSP was supported by UTHealth Innovation for Cancer Prevention Research Pre-doctoral Fellowship, The University of Texas School of Public Health – Cancer Prevention and Research Institute of Texas grant # RP101503.

Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the Cancer Prevention and Research Institute of Texas.

Public access: For all publications, follow appropriate NIH public access rules (see <http://publicaccess.nih.gov>).

References

- Allen M P, Tildesley D J (1989). *Computer Simulations of Liquids*. New York, Oxford University Press
- Bahar I, Lezon T R, Bakan A, Shrivastava I H (2010). Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins. *Chem Rev*, 110(3): 1463–1497
- Baker D, Sali A (2001). Protein structure prediction and structural genomics. *Science*, 294(5540): 93–96
- Baker N A, Sept D, Joseph S, Holst M J, McCammon J A (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA*, 98(18): 10037–10041
- Barbosa L C, Garrido S S, Garcia A, Delfino D B, Marchetto R (2010). Function inferences from a molecular structural model of bacterial ParE toxin. *Bioinformatics*, 4(10): 438–440
- Barcellos G B, Pauli I, Caceres R A, Timmers L F, Dias R, de Azevedo W F Jr (2008). Molecular modeling as a tool for drug discovery. *Curr Drug Targets*, 9(12): 1084–1091
- Bond P J, Wee C L, Sansom M S (2008). Coarse-grained molecular dynamics simulations of the energetics of helix insertion into a lipid bilayer. *Biochemistry*, 47(43): 11321–11331
- Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T (2009). Protein structure homology modeling using SWISS-MODEL workspace. *Nat Protoc*, 4(1): 1–13
- Bradley P, Misura K M, Baker D (2005). Toward high-resolution de novo structure prediction for small proteins. *Science*, 309(5742): 1868–1871
- Brunetti L, Di Stefano M, Ruggieri S, Cimadamore F, Magni G (2010). Homology modeling and deletion mutants of human nicotinamide mononucleotide adenylyltransferase isozyme 2: new insights on structure and function relationship. *Protein Sci*, 19(12): 2440–2450
- Burley S K, Joachimiak A, Montelione G T, Wilson I A (2008). Contributions to the NIH-NIGMS protein structure initiative from the PSI production centers. *Structure*, 16(1): 5–11
- Butler K V, Kalin J, Brochier C, Vistoli G, Langley B, Kozikowski A P (2010). Rational design and simple chemistry yield a superior, neuroprotective HDAC6 inhibitor, tubastatin A. *J Am Chem Soc*, 132(31): 10842–10846
- Bystroff C, Krogh A (2008). Hidden Markov Models for prediction of protein features. *Methods Mol Biol*, 413: 173–198
- B-Rao C, Subramanian J, Sharma S D (2009). Managing protein flexibility in docking and its applications. *Drug Discov Today*, 14(7–8): 394–400
- Carter P H, Tebben A J (2009). Chapter 12. The use of receptor homology modeling to facilitate the design of selective chemokine receptor antagonists. *Methods Enzymol*, 461: 249–279
- Cavasotto C N, Orry A J, Murgolo N J, Czarniecki M F, Kocsi S A, Hawes B E, O'Neill K A, Hine H, Burton M S, Voigt J H, Abagyan R A, Bayne M L, Monsma F J Jr (2008). Discovery of novel chemotypes to a G-protein-coupled receptor through ligand-steered homology modeling and structure-based virtual screening. *J Med Chem*, 51(3): 581–588
- Cavasotto C N, Phatak S S (2009). Homology modeling in drug discovery: current trends and applications. *Drug Discov Today*, 14(13–14): 676–683
- Cherezov V, Abola E, Stevens R C (2010). Recent progress in the structure determination of GPCRs, a membrane protein family with high potential as pharmaceutical targets. *Methods Mol Biol*, 654: 141–168
- Chopra G, Kalisman N, Levitt M (2010). Consistent refinement of submitted models at CASP using a knowledge-based potential. *Proteins*, 78(12): 2668–2678
- Colizzi F, Perozzo R, Scapozza L, Recanatini M, Cavalli A (2010). Single-molecule pulling simulations can discern active from inactive enzyme inhibitors. *J Am Chem Soc*, 132(21): 7361–7371
- Cornell W, Nam K (2009). Steroid hormone binding receptors: application of homology modeling, induced fit docking, and molecular dynamics to study structure-function relationships. *Curr Top Med Chem*, 9(9): 844–853
- Coumar M S, Chu C Y, Lin C W, Shiao H Y, Ho Y L, Reddy R, Lin W H, Chen C H, Peng Y H, Leou J S, Lien T W, Huang C T, Fang M Y, Wu S H, Wu J S, Chittimalla S K, Song J S, Hsu J T, Wu S Y, Liao C C, Chao Y S, Hsieh H P (2010). Fast-forwarding hit to lead: aurora and epidermal growth factor receptor kinase inhibitor lead identification. *J Med Chem*, 53(13): 4980–4988
- Cozzetto D, Tramontano A (2008). Advances and pitfalls in protein structure prediction. *Curr Protein Pept Sci*, 9(6): 567–577
- Cozzini P, Kellogg G E, Spyraakis F, Abraham D J, Costantino G, Emerson A, Fanelli F, Gohlke H, Kuhn L A, Morris G M, Orozco M, Pertinhez T A, Rizzi M, Sotriffer C A (2008). Target flexibility: an emerging consideration in drug discovery and design. *J Med Chem*, 51(20): 6237–6255
- Daga P R, Patel R Y, Doerksen R J (2010). Template-based protein modeling: recent methodological advances. *Curr Top Med Chem*, 10(1): 84–94
- Das R, Baker D (2008). Macromolecular modeling with rosetta. *Annu Rev Biochem*, 77(1): 363–382
- Das R, Qian B, Raman S, Vernon R, Thompson J, Bradley P, Khare S, Tyka M D, Bhat D, Chivian D, Kim D E, Sheffler W H, Malmström L, Wollacott A M, Wang C, Andre I, Baker D (2007). Structure prediction for CASP7 targets using extensive all-atom refinement with Rosetta@home. *Proteins*, 69(S8): 118–128
- Deng Y, Roux B (2009). Computations of standard binding free energies with molecular dynamics simulations. *J Phys Chem B*, 113(8): 2234–2246
- Deschavanne P, Tufféry P (2009). Enhanced protein fold recognition using a structural alphabet. *Proteins*, 76(1): 129–137
- Dessailly B H, Nair R, Jaroszewski L, Fajardo J E, Kouranov A, Lee D, Fiser A, Godzik A, Rost B, Orengo C (2009). PSI-2: structural genomics to cover protein domain family space. *Structure*, 17(6): 869–881
- Diaz P, Phatak S S, Xu J, Astruc-Diaz F, Cavasotto C N, Naguib M (2009a). 6-Methoxy-N-alkyl isatin acylhydrazone derivatives as a novel series of potent selective cannabinoid receptor 2 inverse agonists: design, synthesis, and binding mode prediction. *J Med Chem*, 52(2): 433–444
- Diaz P, Phatak S S, Xu J, Fronczek F R, Astruc-Diaz F, Thompson C M,

- Cavasotto C N, Naguib M (2009b). 2,3-Dihydro-1-benzofuran derivatives as a series of potent selective cannabinoid receptor 2 agonists: design, synthesis, and binding mode prediction through ligand-steered modeling. *Chem Med Chem*, 4(10): 1615–1629
- Duan J, Wu J, Cheng Y, Duan R D (2010). Understanding the molecular activity of alkaline sphingomyelinase (NPP7) by computer modeling. *Biochemistry*, 49(42): 9096–9105
- Dunbrack R L Jr (2006). Sequence comparison and protein structure prediction. *Curr Opin Struct Biol*, 16(3): 374–384
- Engels K, Beyer C, Suárez Fernández M L, Bender F, Gassel M, Unden G, Marhöfer R J, Mottram J C, Selzer P M (2010). Inhibition of *Eimeria tenella* CDK-related kinase 2: From target identification to lead compounds. *Chem Med Chem*, 5(8): 1259–1271
- Evers A, Gohlke H, Klebe G (2003). Ligand-supported homology modelling of protein binding-sites using knowledge-based potentials. *J Mol Biol*, 334(2): 327–345
- Evers A, Klebe G (2004). Successful virtual screening for a submicromolar antagonist of the neurokinin-1 receptor based on a ligand-supported homology model. *J Med Chem*, 47(22): 5381–5392
- Fariselli P, Rossi I, Capriotti E, Casadio R (2007). The WWW of remote homolog detection: the state of the art. *Brief Bioinform*, 8(2): 78–87
- Frenkel D, Smit B (2002). *Understanding Molecular Simulations: From Algorithms to Applications*. San Diego, Academic Press
- Fujitani H, Tanida Y, Ito M, Jayachandran G, Snow C D, Shirts M R, Sorin E J, Pande V S (2005). Direct calculation of the binding free energies of FKBP ligands. *J Chem Phys*, 123(8): 084108
- Ge X, Roux B (2010). Absolute binding free energy calculations of sparsomycin analogs to the bacterial ribosome. *J Phys Chem B*, 114(29): 9525–9539
- Gerek Z N, Ozkan S B (2010). A flexible docking scheme to explore the binding selectivity of PDZ domains. *Protein Sci*, 19(5): 914–928
- Ginalski K (2006). Comparative modeling for protein structure prediction. *Curr Opin Struct Biol*, 16(2): 172–177
- Goodsell D S, Morris G M, Olson A J (1996). Automated docking of flexible ligands: applications of AutoDock. *J Mol Recognit*, 9(1): 1–5
- Grant M A (2009). Protein structure prediction in structure-based ligand design and virtual screening. *Comb Chem High Throughput Screen*, 12(10): 940–960
- Gruber C W, Muttenthaler M, Freissmuth M (2010). Ligand-based peptide design and combinatorial peptide libraries to target G protein-coupled receptors. *Curr Pharm Des*, 16(28): 3071–3088
- Grünberg R, Nilges M, Leckner J (2007). Biskit—a software platform for structural bioinformatics. *Bioinformatics*, 23(6): 769–770
- Hildebrand A, Remmert M, Biegert A, Söding J (2009). Fast and accurate automatic structure prediction with HHpred. *Proteins*, 77(S9 Suppl 9): 128–132
- Holford N, Ma S C, Ploeger B A (2010). Clinical trial simulation: a review. *Clin Pharmacol Ther*, 88(2): 166–182
- Hou T, Wang J, Li Y Y, Wang W (2011). Assessing the performance of the MM/PBSA and MM/GBSA Methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model*, 51(1): 69–82
- Hu H, He L Y, Gong Z, Li N, Lu Y N, Zhai Q W, Liu H, Jiang H L, Zhu W L, Wang H Y (2009). A novel class of antagonists for the FFAs receptor GPR40. *Biochem Biophys Res Commun*, 390(3): 557–563
- Irwin J J, Shoichet B K (2005). ZINC—a free database of commercially available compounds for virtual screening. *J Chem Inf Model*, 45(1): 177–182
- Jamieson C, Basten S, Campbell R A, Cumming I A, Gillen K J, Gillespie J, Kazemier B, Kiczun M, Lamont Y, Lyons A J, Maclean J K, Moir E M, Morrow J A, Papakosta M, Rankovic Z, Smith L (2010). A novel series of positive modulators of the AMPA receptor: discovery and structure based hit-to-lead studies. *Bioorg Med Chem Lett*, 20(19): 5753–5756
- Jayalakshmi R, Natarajan R, Vivekanandan M, Natarajan G S (2010). Alignment-free sequence comparison using N-dimensional similarity space. *Curr Comput Aided Drug Des*, 6(4): 290–296
- Jiao D, Golubkov P A, Darden T A, Ren P (2008). Calculation of protein-ligand binding free energy by using a polarizable potential. *Proc Natl Acad Sci USA*, 105(17): 6290–6295
- Jiao D, Zhang J, Duke R E, Li G, Schnieders M J, Ren P (2009). Trypsin-ligand binding free energies from explicit and implicit solvent simulations with polarizable potential. *J Comput Chem*, 30(11): 1701–1711
- Kannan S, Zacharias M (2010). Application of biasing-potential replica-exchange simulations for loop modeling and refinement of proteins in explicit solvent. *Proteins*, 78(13): 2809–2819
- Karplus K (2009). SAM-T08, HMM-based protein structure prediction. *Nucleic Acids Res*, 37(Suppl 2): W492–W497
- Kaufmann K W, Lemmon G H, Deluca S L, Sheehan J H, Meiler J (2010). Practically useful: what the Rosetta protein modeling suite can do for you. *Biochemistry*, 49(14): 2987–2998
- Khalili-Araghi F, Gumbart J, Wen P C, Sotomayor M, Tajkhorshid E, Schulten K (2009). Molecular dynamics simulations of membrane channels and transporters. *Curr Opin Struct Biol*, 19(2): 128–137
- Kikugawa G, Apostolov R, Kamiya N, Taiji M, Himeno R, Nakamura H, Yonezawa Y (2009). Application of MDGRAPE-3, a special purpose board for molecular dynamics simulations, to periodic biomolecular systems. *J Comput Chem*, 30(1): 110–118
- Kim D E, Chivian D, Baker D (2004). Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res*, 32(Suppl 2): W526–W531
- Kimura S R, Tebben A J, Langley D R (2008). Expanding GPCR homology model binding sites via a balloon potential: A molecular dynamics refinement approach. *Proteins*, 71(4): 1919–1929
- Kiran M, Coakley S, Walkinshaw N, McMinn P, Holcombe M (2008). Validation and discovery from computational biology models. *Biosystems*, 93(1–2): 141–150
- Klepeis J L, Lindorff-Larsen K, Dror R O, Shaw D E (2009). Long-timescale molecular dynamics simulations of protein structure and function. *Curr Opin Struct Biol*, 19(2): 120–127
- Kortagere S, Cheng S Y, Antonio T, Zhen J, Reith M E, Dutta A K (2011). Interaction of novel hybrid compounds with the D3 dopamine receptor: Site-directed mutagenesis and homology modeling studies. *Biochem Pharmacol*, 81(1): 157–163
- Kryshtafovych A, Fidelis K (2009). Protein structure prediction and model quality assessment. *Drug Discov Today*, 14(7–8): 386–393
- Kubarenko A, Frank M, Weber A N (2007). Structure-function relationships of Toll-like receptor domains through homology modelling and molecular dynamics. *Biochem Soc Trans*, 35(Pt 6): 1515–1518
- Kurcuoglu O, Bates P A (2010). Mechanism of cohesin loading onto chromosomes: a conformational dynamics study. *Biophys J*, 99(4):

- 1212–1220
- Kutzner C, van der Spoel D, Fechner M, Lindahl E, Schmitt U W, de Groot B L, Grubmüller H (2007). Speeding up parallel GROMACS on high-latency networks. *J Comput Chem*, 28(12): 2075–2084
- Lampros C, Papaloukas C, Exarchos K, Fotiadis D I, Tsalikakis D (2009). Improving the protein fold recognition accuracy of a reduced state-space Hidden Markov model. *Comput Biol Med*, 39(10): 907–914
- Larsson P, Wallner B, Lindahl E, Elofsson A (2008). Using multiple templates to improve quality of homology models in automated homology modeling. *Protein Sci*, 17(6): 990–1002
- Laursen L (2009). Computational biology: Biological logic. *Nature*, 462(7272): 408–410
- Leach A R (2001). *Molecular Modeling: Principles and Applications*. Dorset, Pearson Education Ltd.
- Lee M M, Bundschuh R, Chan M K (2008). Distant homology detection using a LEngth and SStructure-based sequence Alignment Tool (LESTAT). *Proteins*, 71(3): 1409–1419
- Lin J H, Perryman A L, Schames J R, McCammon J A (2002). Computational drug design accommodating receptor flexibility: the relaxed complex scheme. *J Am Chem Soc*, 124(20): 5632–5633
- Lindahl E, Sansom M S (2008). Membrane proteins: molecular dynamics simulations. *Curr Opin Struct Biol*, 18(4): 425–431
- Liphardt J, Dumont S, Smith S B, Tinoco I Jr, Bustamante C (2002). Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. *Science*, 296(5574): 1832–1835
- Lobley A, Sadowski M I, Jones D T (2009). pGenTHREADER and pDomTHREADER: new methods for improved protein fold recognition and superfamily discrimination. *Bioinformatics*, 25(14): 1761–1767
- Lu G, Zhang S, Fang X (2008). An improved string composition method for sequence comparison. *BMC Bioinformatics*, 9(Suppl 6): S15
- Malmström L, Goodlett D R (2010). Protein structure modeling. *Methods Mol Biol*, 673: 63–72
- Marco E, Gago F (2007). Overcoming the inadequacies or limitations of experimental structures as drug targets by using computational modeling tools and molecular dynamics simulations. *Chem Med Chem*, 2(10): 1388–1401
- Marrink S J, Risselada H J, Yefimov S, Tieleman D P, de Vries A H (2007). The MARTINI force field: coarse grained model for biomolecular simulations. *J Phys Chem B*, 111(27): 7812–7824
- Medina-Franco J L, Lopez-Vallejo F, Kuck D, Lyko F (2010). Natural products as DNA methyltransferase inhibitors: a computer-aided discovery approach. *Mol Divers*,
- Michino M, Chen J, Stevens R C, Brooks C L 3rd (2010). FoldGPCR: structure prediction protocol for the transmembrane domain of G protein-coupled receptors from class A. *Proteins*, 78(10): 2189–2201
- Montelione G T, Szyperski T (2010). Advances in protein NMR provided by the NIGMS Protein Structure Initiative: impact on drug discovery. *Curr Opin Drug Discov Devel*, 13(3): 335–349
- Mooney C, Pollastri G (2009). Beyond the Twilight Zone: automated prediction of structural properties of proteins by recursive neural networks and remote homology information. *Proteins*, 77(1): 181–190
- Moro S, Deflorian F, Bacilieri M, Spalluto G (2006). Ligand-based homology modeling as attractive tool to inspect GPCR structural plasticity. *Curr Pharm Des*, 12(17): 2175–2185
- Morra G, Meli M, Colombo G (2008). Molecular dynamics simulations of proteins and peptides: from folding to drug design. *Curr Protein Pept Sci*, 9(2): 181–196
- Mortier J, Frederick R, Ganef C, Remouchamps C, Talaga P, Pochet L, Wouters J, Piette J, Dejardin E, Masereel B (2010). Pyrazolo[4,3-c] isoquinolines as potential inhibitors of NF-kappaB activation. *Biochem Pharmacol*, 79(10): 1462–1472
- Mortier J, Masereel B, Remouchamps C, Ganef C, Piette J, Frederick R (2010). NF-kappaB inducing kinase (NIK) inhibitors: identification of new scaffolds using virtual screening. *Bioorg Med Chem Lett*, 20(15): 4515–4520
- Muddassar M, Jang J W, Hong S K, Cho Y S, Kim E E, Keum K C, Oh T, Cho S N, Pae A N (2010). Identification of novel antitubercular compounds through hybrid virtual screening approach. *Bioorg Med Chem*, 18(18): 6914–6921
- Murumkar P R, Zambre V P, Yadav M R (2010). Development of predictive pharmacophore model for in silico screening, and 3D QSAR CoMFA and CoMSIA studies for lead optimization, for designing of potent tumor necrosis factor alpha converting enzyme inhibitors. *J Comput Aided Mol Des*, 24(2): 143–156
- Myler P J, Stacy R, Stewart L, Staker B L, Van Voorhis W C, Varani G, Buchko G W (2009). The Seattle Structural Genomics Center for Infectious Disease (SSGCID). *Infect Disord Drug Targets*, 9(5): 493–506
- Neves M A, Simões S, Sáe Melo M L (2010). Ligand-guided optimization of CXCR4 homology models for virtual screening using a multiple chemotype approach. *J Comput Aided Mol Des*, 24(12): 1023–1033
- Obungu V H, Gelfanova V, Rathnachalam R, Bailey A, Sloan-Lancaster J, Huang L (2009). Determination of the mechanism of action of anti-FasL antibody by epitope mapping and homology modeling. *Biochemistry*, 48(30): 7251–7260
- Odell L R, Howan D, Gordon C P, Robertson M J, Chau N, Mariana A, Whiting A E, Abagyan R, Daniel J A, Gorgani N N, Robinson P J, McCluskey A (2010). The pthaladyns: GTP competitive inhibitors of dynamin I and II GTPase derived from virtual screening. *J Med Chem*, 53(14): 5267–5280
- Ogata K, Isomura T, Kawata S, Yamashita H, Kubodera H, Wodak S J (2010). Lead generation and optimization based on protein-ligand complementarity. *Molecules*, 15(6): 4382–4400
- Okimoto N, Futatsugi N, Fuji H, Suenaga A, Morimoto G, Yanai R, Ohno Y, Narumi T, Taiji M (2010). High-performance drug discovery: Computational screening by combining docking and molecular dynamics simulations. *PLoS Comp Bio*, 5(10): e1000528
- Okumura H, Gallicchio E, Levy R M (2010). Conformational populations of ligand-sized molecules by replica exchange molecular dynamics and temperature reweighting. *J Comput Chem*, 31(7): 1357–1367
- Ostrov D A, Magis A T, Wronski T J, Chan E K, Toro E J, Donatelli R E, Sajek K, Haroun I N, Nagib M I, Piedrahita A, Harris A, Holliday L S (2009). Identification of enoxacin as an inhibitor of osteoclast formation and bone resorption by structure-based virtual screening. *J Med Chem*, 52(16): 5144–5151
- Pandit S A, Scott H L (2009). Multiscale simulations of heterogeneous model membranes. *Biochim Biophys Acta*, 1788(1): 136–148
- Pecic S, Makkar P, Chaudhary S, Reddy B V, Navarro H A, Harding W W (2010). Affinity of aporphines for the human 5-HT2A receptor:

- insights from homology modeling and molecular docking studies. *Bioorg Med Chem*, 18(15): 5562–5575
- Peng J, Xu J (2010). Low-homology protein threading. *Bioinformatics*, 26(12): i294–i300
- Phatak S S, Clifford C S, et al (2009). High-throughput and in silico screening in drug discovery. *Exp. Opin. Drug Disc*, 4(9): 947–959
- Phatak S S, Gatica E A, Cavasotto C N (2010). Ligand-steered modeling and docking: a benchmarking study in class a g-protein-coupled receptors. *J Chem Inf Model*, 50(12): 2119–2128
- Phillips J C, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel R D, Kalé L, Schulten K (2005). Scalable molecular dynamics with NAMD. *J Comput Chem*, 26(16): 1781–1802
- Postigo M P, Guido R V, Oliva G, Castilho M S, da R Pitta I, de Albuquerque J F, Andricopulo A D (2010). Discovery of new inhibitors of *Schistosoma mansoni* PNP by pharmacophore-based virtual screening. *J Chem Inf Model*, 50(9): 1693–1705
- Radestock S, Weil T, Renner S (2008). Homology model-based virtual screening for GPCR ligands using docking and target-biased scoring. *J Chem Inf Model*, 48(5): 1104–1117
- Rai B K, Tawa G J, Katz A H, Humblet C (2010). Modeling G protein-coupled receptors for structure-based drug discovery using low-frequency normal modes for refinement of homology models: application to H3 antagonists. *Proteins*, 78(2): 457–473
- Rajman I (2008). PK/PD modelling and simulations: utility in drug development. *Drug Discov Today*, 13(7–8): 341–346
- Roy A, Kucukural A, Zhang Y (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc*, 5(4): 725–738
- Sansom M S, Scott K A, Bond P J (2008). Coarse-grained simulation: a high-throughput computational approach to membrane proteins. *Biochem Soc Trans*, 36(Pt 1): 27–32
- Shaw D E, Chao J C, Eastwood M P, Gagliardo J, Grossman J P, Ho C R, Ierardi D J, Kolossváry I, Klepeis J L, Layman T, McLeavey C, Deneroff M M, Moraes M A, Mueller R, Priest E C, Shan Y, Spengler J, Theobald M, Towles B, Wang S C, Dror R O, Kuskin J S, Larson R H, Salmon J K, Young C, Batson B, Bowers K J (2007). Anton, a special-purpose machine for molecular dynamics simulation. *ACM SIGARCH Computer Architecture News*, 35(2): 1–12
- Shirts M R, Pande V S (2005). Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration. *J Chem Phys*, 122(14): 144107
- Sierecki E, Sinko W, McCammon J A, Newton A C (2010). Discovery of small molecule inhibitors of the PH domain leucine-rich repeat protein phosphatase (PHLPP) by chemical and virtual screening. *J Med Chem*, 53(19): 6899–6911
- Sisay M T, Steinmetzer T, Stirnberg M, Maurer E, Hammami M, Bajorath J, Gütschow M (2010). Identification of the first low-molecular-weight inhibitors of matriptase-2. *J Med Chem*, 53(15): 5523–5535
- Stone J E, Hardy D J, Ufimtsev I S, Schulten K (2010). GPU-accelerated molecular modeling coming of age. *J Mol Graph Model*, 29(2): 116–125
- Stumpff-Kane A W, Maksimiak K, Lee M S, Feig M (2008). Sampling of near-native protein conformations during protein structure refinement using a coarse-grained model, normal modes, and molecular dynamics simulations. *Proteins*, 70(4): 1345–1356
- Sugita Y, Okamoto Y (1999). Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett*, 314(1–2): 141–151
- Talele T T, Arora P, Kulkarni S S, Patel M R, Singh S, Chudayeu M, Kaushik-Basu N (2010). Structure-based virtual screening, synthesis and SAR of novel inhibitors of hepatitis C virus NS5B polymerase. *Bioorg Med Chem*, 18(13): 4630–4638
- Trojanowski S, Rutkowska A, Kolinski A (2010). TRACER. A new approach to comparative modeling that combines threading with free-space conformational sampling. *Acta Biochim Pol*, 57(1): 125–133
- Verdonk M L, Cole J C, Hartshorn M J, Murray C W, Taylor R D (2003). Improved protein-ligand docking using GOLD. *Proteins*, 52(4): 609–623
- Walsh I, Baù D, Martin A J, Mooney C, Vullo A, Pollastri G (2009). Ab initio and template-based prediction of multi-class distance maps by two-dimensional recursive neural networks. *BMC Struct Biol*, 9(1): 5
- Weigelt J (2010). Structural genomics-impact on biomedicine and drug discovery. *Exp Cell Res*, 316(8): 1332–1338
- White S H (2004). The progress of membrane protein structure determination. *Protein Sci*, 13(7): 1948–1949
- Wichapong K, Pianwanit S, Sippl W, Kokpol S (2010). Homology modeling and molecular dynamics simulations of Dengue virus NS2B/NS3 protease: insight into molecular interaction. *J Mol Recognit*, 23(3): 283–300
- Wu S, Skolnick J, Zhang Y (2007). Ab initio modeling of small proteins by iterative TASSER simulations. *BMC Biol*, 5(1): 17
- Xiang Z (2006). Advances in homology protein structure modeling. *Curr Protein Pept Sci*, 7(3): 217–227
- Yan R X, Si J N, Wang C, Zhang Z (2009). DescFold: a web server for protein fold recognition. *BMC Bioinformatics*, 10(1): 416
- Yang L J, Zou J, Xie H Z, Li L L, Wei Y Q, Yang S Y (2009). Steered molecular dynamics simulations reveal the likelier dissociation pathway of imatinib from its targeting kinases c-Kit and Abl. *PLoS One*, 4(12): e8470
- Yao L, Evans J A, Rzhetsky A (2009). Novel opportunities for computational biology and sociology in drug discovery. *Trends Biotechnol*, 27(9): 531–540
- Ying Y, Huang K, Campbell C (2009). Enhanced protein fold recognition through a novel data integration approach. *BMC Bioinformatics*, 10(1): 267
- Zhang Y (2008). Progress and challenges in protein structure prediction. *Curr Opin Struct Biol*, 18(3): 342–348
- Zhou H, Skolnick J (2010). Improving threading algorithms for remote homology modeling by combining fragment and template comparisons. *Proteins*, 78(9): 2041–2048
- Zhu J, Cheng L, Fang Q, Zhou Z H, Honig B (2010). Building and refining protein models within cryo-electron microscopy density maps based on homology modeling and multiscale structure refinement. *J Mol Biol*, 397(3): 835–851
- Zhu J, Fan H, Periole X, Honig B, Mark A E (2008). Refining homology models by combining replica-exchange molecular dynamics and statistical potentials. *Proteins*, 72(4): 1171–1188
- Zwier M C, Chong L T (2010). Reaching biological timescales with all-atom molecular dynamics simulations. *Curr Opin Pharmacol*, 10(6): 745–752