Quarterly Reviews of Biophysics **45**, 3 (2012), pp. 301–343. © Cambridge University Press 2012
doi:10.1017/S0033583512000066 Printed in the United States of America

Protein flexibility in docking and surface mapping

Katrina W. Lexa and Heather A. Carlson*

Department of Medicinal Chemistry, University of Michigan, 428 Church Street, Ann Arbor, MI 48109-1065, USA

Abstract. Structure-based drug design has become an essential tool for rapid lead discovery and optimization. As available structural information has increased, researchers have become increasingly aware of the importance of protein flexibility for accurate description of the native state. Typical protein–ligand docking efforts still rely on a single rigid receptor, which is an incomplete representation of potential binding conformations of the protein. These rigid docking efforts typically show the best performance rates between 50 and 75%, while fully flexible docking methods can enhance pose prediction up to 80–95%. This review examines the current toolbox for flexible protein–ligand docking and receptor surface mapping. Present limitations and possibilities for future development are discussed.

1. Introduction 302

2. Protein flexibility 303
2.1. Conformational variability 303
2.2. Allostery 304

3. Protein–ligand binding 304
3.1. The cross-docking problem 304
3.2. Sources of structural information 305

4. Early protein–ligand flexibility in docking 308
4.1. Soft docking 308
4.2. Relaxation methods 308
4.3. Docking with flexible side chains 308

5. Induced fit docking 313
5.1. Conformational generation on the fly 313

6. Ensemble docking 314
6.1. Grid-based ensembles 314
6.2. Structure-based ensembles 322

7. *De novo* approaches through fragment-based drug design 324

8. Limitations of structure-based drug design 335

9. Future directions 336

10. Acknowledgements 336

11. References 337

* Author for correspondence: Heather A. Carlson. Email: carlsonh@umich.edu

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

1. Introduction

Structure-based drug design (SBDD) involves the use of three-dimensional (3D) structural data to advance lead identification and subsequent optimization for drug discovery. The exponential growth of the Protein Data Bank (PDB) and improvement in homology modeling techniques makes SBDD applicable to an ever-growing number of pharmaceutically relevant targets with an available 3D structure. SBDD studies are generally centered upon at least one of the following three goals: prediction of binding modes, prediction of binding affinities and prediction of novel binding partners. Depending on these goals and available data, SBDD can involve different approaches, which are often separated into docking techniques and *de novo* design. Docking is implemented to predict the binding mode and affinity of small molecules. When docking is applied to a large compound database, it is referred to as virtual screening (VS) and often centered on the prediction of new ligands with high affinity for a target molecule to enrich the compound set for experimental testing. *De novo* design is performed with the intent of predicting new compounds in novel chemical space. Fragment-mapping techniques that use functional groups to probe binding sites are often applied for this type of approach. Over the past 20 years, these SBDD studies have produced viable leads, enabling the development of successful clinical drugs and leading to the extensive implementation of SBDD in medicinal chemistry research (Jorgensen, 2004; Taft *et al.* 2008; Talele *et al.* 2010).

X-ray crystallography and NMR studies have clearly demonstrated conformational differences between many receptors' holo (bound) and apo (unbound) states. Sampling ligand conformations is straightforward; most SBDD protocols now include ligand flexibility (with on-the-fly sampling being superior to a rigid set of pre-generated conformations), yet this is insufficient for the most accurate results. Despite data demonstrating the influence of protein flexibility on ligand binding, most SBDD efforts still rely upon a static receptor structure because of the resources required to account for its many degrees of freedom. Although a few proteins can have their binding potential represented with a single, fixed conformation, for most systems the information presented by a rigid structure is simply inadequate.

In a comparative study of 10 docking programs and 37 scoring functions, no single method outperformed the others when performing docking of diverse compounds to a set of eight rigid proteins (Warren *et al.* 2006). The scoring functions were unable to accurately predict binding affinity or relatively rank the compounds. Ginalski and co-workers evaluated the binding predictions and scoring results for 1300 protein–ligand complexes from PDBbind 2007 with Surflex, LigandFit, Glide, GOLD, FlexX, eHiTS, and AutoDock (Plewczynski *et al.* 2011). The authors found that the programs achieved a mean $RMSD_{Top\text{-}Score}$ that ranged from 2.77 Å (GOLD) to 4.37 Å (FlexX) for the docked poses, but those top scoring poses did not typically match the best RMSD pose (matching occurred in 40–60% of cases), and none of the scoring functions were able to achieve a reasonable correlation between the best pose score and the experimental activity (Spearman correlation ranged from 0.07 (LigandFit) to 0.39 (eHiTS)).

Several recent articles and reviews have been published, which discuss the deficiencies of existing scoring functions for docking (Englebienne & Moitessier, 2009; Huang *et al.* 2010; Leach *et al.* 2006; Moitessier *et al.* 2008; Mukherjee *et al.* 2010; Warren *et al.* 2006). Importantly, Englebienne and Moitessier's careful study of 209 protein–ligand complexes and 18 scoring functions demonstrated that the accuracy of certain scoring functions was negatively impacted by protein flexibility and solvation. Mukherjee *et al.* (2010) evaluated the relationship between success, sampling failure, and scoring failure in static, fixed anchor, and anchor-and-grow

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

docking outcomes with DOCK. They showed that scoring failures increased as sampling errors decreased, and scoring failures appeared to peak at an RMSD between 1.5 and 2.0 Å. Klebe and co-workers hypothesized that there is an intimate link between docking and scoring, postulating that correct prediction of accurate binding geometries will simultaneously solve the scoring problem (Klebe, 2006; Velec *et al.* 2005; Zentgraf *et al.* 2007). Therefore, while we acknowledge the limitations of current scoring functions, we focus on the variety of techniques for incorporating protein flexibility in SBDD.

A number of reviews have included some consideration of the impact of protein flexibility on drug discovery (Alonso *et al.* 2006; Bottegoni, 2011; Cozzini *et al.* 2008; Durrant & McCammon, 2010; Fuentes *et al.* 2011; Henzler & Rarey, 2010; Klebe, 2006; Leach *et al.* 2006; Rao *et al.* 2009; Sotriffer, 2011; Sousa *et al.* 2006; Teague, 2003; Teodoro & Kavraki, 2003; Totrov & Abagyan, 2008). Many of these reviews concentrated on only a subset of protein flexibility methods or emphasized a particular technique. The individual roles of sampling protein flexibility and accurate scoring upon docking success and failure remain unclear. Without a standard for reporting methodological outcomes or a consistent set of test cases, it is difficult to tease apart the relative influence of conformational searching versus interaction scoring. Furthermore, it may be difficult, if not impossible, to judge the extent to which conformational sampling appropriately models the true potential energy landscape. Here, we try to address these difficult questions through the examination of a variety of techniques that account for flexibility in protein–ligand binding, not only studies where the primary focus is on docking but also studies that concentrate on identifying potential binding sites with small molecules or fragments.

2. Protein flexibility

2.1 Conformational variability

Our understanding of receptor–ligand binding has significantly advanced from the original lock-and-key model proposed by Fischer (1890). Early experimental studies showed that the act of ligand binding influences the protein conformation, referred to as conformational induction or induced fit (Koshland, 1958). Another model of ligand binding is conformational selection, wherein the ligand chooses a binding partner from among available states in the conformational ensemble, thereby shifting the population distribution (Berger *et al.* 1999; Foote & Milstein, 1994; Kumar *et al.* 2000; Monod *et al.* 1965; Tsai *et al.* 1999a, b, 2001). Recently, several papers have been published examining the evidence for whether receptor binding occurs because of induced fit or conformational selection. Sullivan & Holyoak (2008) studied the kinetics of phosphoenolpyruvate carboxykinase and concluded that induced fit and conformational selection were not mutually exclusive, rather they were complementary avenues for binding. Weikl & von Deuster (2009) developed equations of binding kinetics using four states, the dominant apo and holo protein conformations without a bound ligand and those same conformations with a ligand bound, to describe the protein–ligand complex. They used the equilibrium constants from this model to distinguish between the induced and selected fit. Hammes *et al.* (2009) examined the binding pathways for dihydrofolate reductase (DHFR) and flavodoxin and determined that a mixed binding mechanism was most likely. They noted that the relative importance of induced fit or conformational selection for a particular case could be analyzed by comparing the flux through the reaction path of ligand-binding and conformational change. Both mechanisms of binding will produce the same result; it is important only that some

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

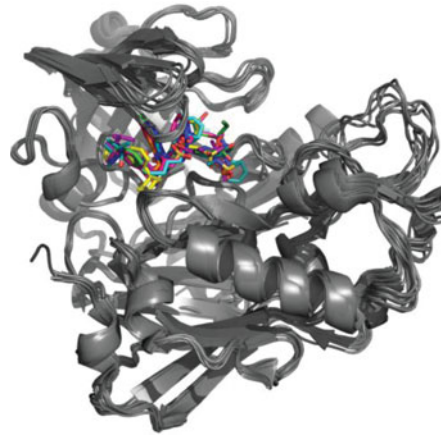


Fig. 1. The conformational variation in holo- and apo structures of beta-secretase 1 illustrates structural differences frequently seen across multiple crystal structures of the same protein.

mechanism of receptor conformational change be incorporated in docking simulations. This is particularly useful in SBDD because it implies that computationally inexpensive methods that include protein flexibility should correctly predict binding modes.

2.2 Allostery

Protein–ligand flexibility upon binding is crucial for proper understanding of allosteric regulation. Nussinov and co-workers postulated that most proteins exist in an ensemble of states, and thus most proteins have the potential for allostery (Gunasekaran *et al.* 2004). Based on the experimental literature, they demonstrated that the binding of an allosteric ligand shifted the population of conformational substates, thereby influencing the ability of other ligands to bind an alternate site.

Drug resistance is a growing problem that calls for new approaches to drug therapy. Exploring allosteric control in protein targets provides exciting opportunities for finding new modes of action and hence, overcoming emergent resistance, and developing cocktails of drugs to improve treatment.

3. Protein–ligand binding

3.1 The cross-docking problem

Research into protein flexibility and allostery has lent support to the importance of representing multiple states in binding studies (Fig. 1). Mobley & Dill (2009) noted that binding free energy (ΔG_{bind}) and entropy are influenced by the shape and width of the entire conformational landscape, rather than a single rigid pose. Murray *et al.* (1999) examined approaches for using rigid receptors in docking studies. When the authors attempted to dock a known ligand into a protein structure solved in the presence of a different ligand (referred to as cross-docking), they found that the active site was biased toward to the native ligand (Fig. 2). A variety of differences in the surface of the binding site were identified for the same protein solved with different ligands or in the absence of a ligand. Movement was observed in the backbone, side chain (both dependent

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

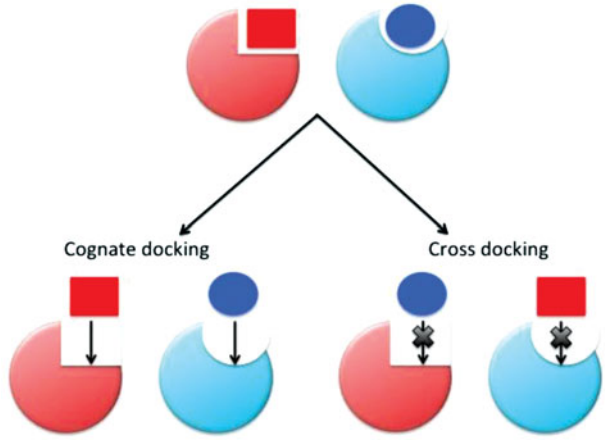


Fig. 2. The same protein crystallized with different ligands. This simple two ligand-one protein set demonstrates the influence of structure on the shape of the binding pocket.

and independent of the backbone Cα), and active site metals. As a consequence, active sites were biased toward a particular ligand type, negatively impacting docking efforts. The resultant mis-docking could not be overcome without accounting for critical conformational shifts. Najmanovich *et al.* (2000) examined the percentage of residues that actually undergo a change upon ligand binding, based on two non-redundant datasets containing paired holo- and apo-protein structures from the PDB. No significant correlation was found between backbone movement and side-chain flexibility, but Lys, Arg, Gln, and Met were identified as the residues most likely to demonstrate side-chain rearrangement upon ligand binding. Najmanovich *et al.* (2000) showed that 60–70% of the binding site undergoes some changes in the side-chain orientation. Taken together, these studies clearly illustrated that the lack of protein flexibility in typical SBDD efforts severely limits the identification of true ligands and accurate docked poses.

An underlying assumption in rigid docking efforts is that the complexed state is the lowest free energy state. However, the ligand-bound state is not always the lowest energy state for either the protein or the ligand. Smith *et al.* (1994) were first to examine conformational similarity, they compared the four available structures of interleukin-4, two from NMR and two from crystallography. They found that differences in the experimental methods clearly affected the structure, and that the two NMR structures were more similar to the crystal structures than to each other. Further, they noted that the exposed regions of the surface, particularly the highly flexible loops, showed the largest conformational difference among the structures. Eyal *et al.* (2005) evaluated a set of 659 structure pairs from the PDB and showed that experimental variation in protein structures certainly exists, as a result of refinement, crystal packing, and/or crystallization conditions. Chopra *et al.* (2008) employed Molecular Dynamics (MD) on a set of 75 proteins to explore the role of solvent in structural determination and concluded that solvent effects have an enormous influence on the final refined protein structure.

3.2 Sources of structural information

Use of a set of similar conformations only generates a finite amount of information on potential binding partners. Additional receptor space should be explored, especially in scenarios where alternate binding modes are of interest. X-ray crystallography, NMR structures, homology

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

models, normal mode analysis (NMA), and molecular mechanics simulations are all potential sources of structural diversity. Advances in technology have allowed for significant increases in available structural information for a vast number of receptor targets. The PDB now contains over 76 000 structures (as of 9/27/11), up from 13 605 structures in 2000 (Berman *et al.* 2000). Several research groups, as well as thousands of computer users, have dedicated their computer power to distributed structure prediction of proteins with Folding@HOME, POEM@HOME, and Rosetta@HOME. Martin-Renom *et al.* (2000) estimated that almost one-third of all known protein sequences can have their structure predicted via homology modeling efforts. Schafferhans & Klebe (2001) found that accuracy increases when the results of several different homology modeling programs are combined. While predictive models are not as accurate as structures solved by x-ray crystallography or NMR, they enable the study of targets that are difficult to determine experimentally. No consensus has been reached on the best source of data for including structural flexibility in SBDD, and a number of studies clearly demonstrate the difficulty and variability in assessing appropriate conformations (Bolstad & Anderson, 2008; Damm & Carlson, 2007; Laughton *et al.* 2009; Philippopoulos & Lim, 1999; Yang *et al.* 2007). Bolstad and Anderson (2008) found that their docking protocol correctly placed all forty tested LcDHFR ligands into the crystal structure, while the scoring function correctly ranked only 80% of the ligands. Including an NMR structure, homology model, or an ensemble of structures increased the number of incorrectly placed ligands while decreasing or maintaining the correct ranking of affinities. The authors found a similar trend when docking 19 ligands to hDHFR. This suggested that both sampling and scoring failures might lead to inaccuracies within a study. Sampling failures may occur without careful selection of flexibility information, but scoring may also negatively impact the results if the function cannot correctly classify binding poses.

The following sections are organized based on different techniques for including protein flexibility in docking and *de novo* design (Table 1). Soft docking allows limited discovery of new conformation space while docking with flexible side-chain samples observed conformational space. More recent techniques incorporate ensembles of structures to allow for a greater measure of native flexibility or to reveal new conformations. By relying on several protein conformations, new chemical space can be explored, allosteric sites can be discovered, and more accurate binding conformations can be generated. The accepted metric for a well-docked pose is $\leqslant 2.0$ Å RMSD from the experimentally determined structure. Protein flexibility can be incorporated into a binding-site model before docking is initiated or it can be allowed post-docking through refinement of the bound complex. Docking can be performed against an average structure or ensemble, a set of receptor conformations (multiple receptor conformations (MRC)), or with conformations generated 'on the fly' (induced-fit docking (IFD)).

Experimental methods for SBDD inherently include protein–ligand flexibility. However, lead development and design that is based solely on the experimental data for ligand binding can be time-consuming and prohibitively expensive since every lead optimization step must be tested. Fragment-based drug design (FBDD) has enabled the application of experimental methods to searching the interaction potential along the protein surface with small organic probes. The use of fragment-based studies in SBDD enables the identification of ligands with high affinity for targets that have been traditionally difficult to characterize. These insights have greatly influenced computational approaches to protein flexibility, where fragment mapping can be applied to flexible receptors with functional group probes to find binding hot spots and identify specific interaction potentials. The use of molecular dynamics and/or probe minimizations to search for important binding sites with small probes has been frequently applied in SBDD projects because

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 1. *Techniques for incorporating full protein flexibility into docking approaches*

Approach	Method for incorporation	Advantages	Limitations
Refinement	Flexibility introduced after docking through reduced or all-atom modeling with molecular dynamics or Monte Carlo minimization	Fast docking to rigid receptor enables searching through vast compound libraries	Unlikely to generate as much structural diversity as the other methods, hard to move beyond known binding space
Average or unified structure	Ensemble averaging through use of a unified structure or grid representation Can also occur through the selection of conformational subsets from a rotameric library May also involve generation of receptor conformations based on ligand poses	Can find novel binding mechanisms, orphan sites, and explore new receptor conformations	Discovery of 'paradoxical inhibitors' that bind only to averaged conformation but not a native structure
Serial docking	Docking performed iteratively to a rigid ensemble of structures, conformational variation of the receptor ensemble typically comes from the inclusion of several x-ray crystallography, NMR, homology model, PCA-derived, NMA-derived or MD-derived structures	Can allow for discovery of novel binding modes	Ensemble generation and parallel docking can be time-consuming, not usually appropriate for screening large libraries Structural variation can increase false positives
Conformations on the fly	Receptor conformational changes are explicitly modeled during docking	Allows receptor conformation to change during interaction with ligand for optimal binding, can be quite accurate	Can generate conformations that are not experimentally accessible Can be quite time-intensive, not necessarily appropriate for virtual screening

the results enable the development of a pharmacophore model for use in virtual screening studies.

4. Early protein–ligand flexibility in docking

4.1 Soft docking

It was not until the mid-1990s that researchers began to tackle the problem of receptor degrees of freedom in SBDD. The original technique involved soft docking, which accounted for minimal backbone movement, and side-chain flexibility through attenuation of the Lennard–Jones repulsion term between the rigid protein and ligand (Jiang & Kim, 1991). This allowed for some penetration between the ligand and protein, and was followed by a rigid-body minimization. Soft docking increased the number of high scoring hits compared to the use of a single structure for docking. A study of T4 lysozyme and aldose reductase showed that soft docking was superior to docking to a static structure (Ferrari *et al.* 2004). However, the authors also found that serial docking to a structural ensemble of rigid receptors identified six novel compounds targeting aldose reductase that soft docking had ranked poorly. Of these six compounds, four were experimentally verified and one had a low micromolar IC_{50} value. Ferrari *et al.* conjectured that the differences in ranking were due to the differences in the scoring function applied for soft versus serial docking.

Soft docking is a fast and easy method for including some protein flexibility into docking studies, and as a result is it included as a step in many of the methods presented below.

4.2 Relaxation methods

Refinement of the docked complex is another simple approach that adjusts for protein flexibility by modeling induced-fit effects. The incorporation of protein flexibility on the 'back end' can only be done when the docking technique was based on an all-atom structure; it cannot be performed on a grid, or in any other situation where the protein is not explicitly represented. Monte Carlo (MC) or MD simulations are a popular choice for refinement because they enable the optimization of docked poses, investigation of solvent effects, examination of the kinetic stability, and prediction of ΔG_{bind} from physics-based scoring functions. Refinement is frequently performed as a final step in many of the docking approaches discussed here; Table 2 lists protein–ligand docking methods that limit the inclusion of full flexibility to the refinement stage (Mizutani *et al.* 2006; Pencheva *et al.* 2008; Taylor *et al.* 2003).

4.3 Docking with flexible side chains

Early efforts at on-the-fly docking focused on the side chains, with the use of a rotamer library based on backbone dihedral angles to describe protein flexibility. In 1993, rotamer libraries were first used to predict side-chain conformations while studying the protein folding problem (Dunbrack & Karplus, 1993). In 1994, the use of side-chain rotamer libraries was extended to protein–ligand docking with the examination of trypsin-benzamidine and antibody McPC 603-phosphocholine binding through a modified version of AMBER 4.0 (Leach). The study restricted the protein backbone completely while the side chains within 10 Å of the ligand were allowed for sampling a set of discrete rotameric states. Leach found that the presence

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 2. *Studies including full receptor flexibility after docking* (*post-docking refinement*)

Method	Target	Flexibility	Results	Caveats	Author
FDS	15 cases from GOLD test set	Rigid docking directed by hydrogen bonding Results clustered (clique finding technique), side-chain flexibility for ~5 poses allowed via MC with GB/SA, soft potential & SA	Rigid docking reproduced crystal pose for 13 cases Flexible side chains reproduced crystal pose for 11 cases Accounted for continuum solvation With refinement, RMSD of pose closest to crystal conformation (majority ranked 1) was 0.78 to 3.81 Å to the crystal structure Failures were attributed to both sampling and scoring errors	No desolvation penalty Minimal conformation change Required 30–40 h for a single run	Taylor *et al.* (2003)
ADAM/ BLUTO	18 cases for native docking; 22 for cross-docking	Docked to binding site as vdW-offset grid Post-docking optimization of ligand and protein side chains (within 7 Å) of the binding site	Top-ranked poses for thymidine kinase from ADAM/BLUTO were more accurate (RMSD 0.52–1.89 Å) than rigid docking results from GOLD (RMSD 0.49–3.11 Å), DOCK (RMSD 0.82–9.62 Å), Surflex (RMSD 0.74–3.84 Å), Glide (RMSD 0.35–4.22 Å), FlexX (RMSD 0.78–13.30 Å), and ADAM (RMSD 0.49–3.11 Å) RMSD of top-ranked model for all cases from ADAM/BLUTO ranged from 0.43 to 2.66 Å for cross-docking RMSD of top-ranked model for all cases from ADAM alone ranged from 0.67 to 6.31 Å for cross-docking	Most relevant for studying local changes in binding site conformation	Mizutani *et al.* (2006)
AMMOS	Estrogen Receptor, Neuramindase	Five options to refine docked complex, from full to zero protein flexibility Minimized results and reranked Assessed impact of AMMOS force field minimization versus MM94 and Tripos FF on results for four known ligands	Initial docking in DOCK6 RMSD of refined structure to the crystal conformation ranged from 1.03 to 2.12 Å for AMMOS, 1.01 to 2.39 Å for MMFF94s, and 1.17 to 1.57 Å for Tripos FF Enrichment increased by AMMOS refinement from 40 to 60% overall, with actives retrieved from the top 3–5% of the data set	Extent of minimization was very limited (2×500 cycles) Only identified minor conformational shifts	Pencheva *et al.* (2008)

Table 3a. *Docking studies that included side-chain flexibility through rotamer libraries. Caveats for each method are similar in that flexibility of the side chains alone has limited success in describing receptor motion for most binding events and the implementation of a rotamer library limits the conformational space that can be sampled*

Method	Target	Flexibility	Results	Author
Rotamer Library	Trypsin, McPC 603	Rotameric states sampled for all side chains within 10 Å of binding site	Limited by the lack of a solvation term RMSD of closest structure to the bound pose from the crystal was 0.7 Å for trypsin, 0.8 Å for McPC 603; in neither case was the closest pose also the lowest energy structure Did not recover a good correlation between binding energy and RMSD to crystal structure	Leach (1994)
SOFTSPOTS/ PLASTIC	Thymidylate synthase	Identified variation based on structural comparison (or binding site analysis) Disregarded polar residues Retained hydrophobic residues and loops for potential adaptation Subjected 3 residues to rotamer variation Minimized docked pose	Minimized, remodeled complex achieved reasonable energy scores for two potent inhibitors Found −51.5 kcal/mol for ligand BW1843U89 and −49.7 kcal/mol for CB3717 for cross-docking with side-chain flexibility Found −32.8 kcal/mol for CB3717 and −44.7 kcal/mol for BW1843U89 with rigid docking Scores from native docking were −56.5 kcal/mol for BW1843U89 and −52.9 kcal/mol for CB3717	Anderson *et al.* (2001)

of the ligand revealed additional accessible conformational states by modulating the protein's potential energy surface. Similar discrete sampling methods for modeling side-chain flexibility in docking include SOFTSPOTS/PLASTIC (Anderson *et al.* 2001), as presented in Table 3a.

Side-chain motion upon ligand binding has also been approached through continuum sampling methods. Alberts *et al.* (2005a) found that random movement of side chains was preferable to a rotamer library for successful docking results. Techniques that apply some form of continuum sampling to side-chain motion are listed in Table 3b, including AutoDock 4.0 (Morris *et al.* 2009), Dynasite/GOLD (Kallblad & Dean, 2003), FlexX (Rarey *et al.* 1996), ICM/MC (Abagyan *et al.* 1994), Mining Minima (Kairys & Gilson, 2002), Skelgen (Alberts *et al.* 2005a, b), and SLIDE (Schnecke & Kuhn, 2000; Zavodszky *et al.* 2002; Zavodszky & Kuhn, 2005). A rotamer library can be limited by the information it contains, while continuum sampling may locate unrealistic conformational states.

Allowing side-chain flexibility is less resource intensive than full flexibility methods and it enabled some conformational variability through the exploration of low-energy orientations of side chains. However, incorporating side-chain flexibility has failed in cases of proteins with

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 3b. *Docking studies where selected side-chain flexibility was accounted for through continuum sampling. This approach to flexibility at the side-chain positions may explore new space but cannot describe large-scale motion*

Method	Target	Flexibility	Results	Author
ICM/MC	Leucine zipper helices	Applied internal coordinated modeling (ICM), where random changes are applied to the conformation through biased probability random step, a Brownian-like step, or alteration of a single angle 600 K MC minimization of side chains	Lowest-energy conformation of leucine zipper had RMSD of 1.18 Å to crystal pose	Abagyan *et al.* (1994)
Mining Minima	Set of 18 protein–ligand crystal structures	On-the-fly optimization of the 5 nearest side chains to the x-ray ligand, rest of the receptor represented as grid Accelerated narrowing of sampling range for side-chain dihedrals and required 2 kcal/mol energy cutoff between side chain and ligand atoms	RMSD from flexible side-chain docking for lowest energy pose ranged from 0.34 to 2.32 Å to the crystal orientation for eight cases Rigid docking RMSD of lowest energy pose was 0.41–7.31 Å Scoring was inadequate since success was greatest when the first 100 structures were assessed for the lowest-energy conformation within 1.5 Å	Kairys & Gilson (2002)
Skelgen	MMP-1, Acetylcholinesterase	Random transitions of side-chain χ angles were found to be preferable to a rotamer library due to the variation in composition and construction of rotamer libraries	RMSD of pose closest to crystal conformation (majority ranked 1) were 1.0–1.3 Å (native), 1.3–1.4 Å (non-native), and 1.4 Å (apo): flexible RMSD of pose closest to crystal conformation (majority ranked 1) was 0.7–1.0 Å (native), 4.4–4.5 Å (non-native), and 5.6 Å (apo): rigid Best score differed from best pose	Alberts *et al.* (2005a)

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 3b (*cont.*)

Method	Target	Flexibility	Results	Author
SLIDE	20 cases	Ligand anchor fragment, induced fit/side-chain rotation based on mean-field theory Binding site represented as interaction points Docked to the apo pose Included active site solvation Minimal rotation hypothesis: side chains will move as little as necessary to form complex	Specifically intended for systems without large global rearrangements For 63 ligands, best RMSD to the crystal orientation was 0.30–2.48 Å (flexible docking), 0.31–2.16 Å with 33 failures (rigid docking) Sampling failures in rigid docking were corrected with small side-chain rotations	Zavodszky & Kuhn (2005)
AutoDock 4.0	188 native cases, 87 HIVp cross-docking cases	User-selected side chains granted flexibility through torsional rotations, remaining receptor represented as a grid	Docked small native molecules well Successful for most large inhibitors of HIVp, failed relative to rigid docking for small inhibitors in $>50\%$ of the cases (3.5 Å)	Morris *et al.* (2009)

large-scale hinge or loop rearrangements or even backbone-dependent movement of side chains, neither of which can be taken into account by discrete or continuum sampling.

5. Induced fit docking

5.1 Conformational generation on the fly

Significant progress has been made since the initial implementation of flexible side chains in docking. Sampling motion of the side chains on the fly increases the potential energy space but can still overlook global conformational shifts. Allowing for conformational changes during docking can be a highly accurate technique for modeling bound poses of protein–ligand complexes. IFD is important because it can allow the docking simulation to search new conformational space; however, this sampling of receptor and ligand degrees of freedom is also quite computationally intensive, which limits its application in large-scale virtual screening studies.

The first on-the-fly method to expand beyond conformational sampling of side chains without relying upon a composite representation was based on DOCK 4.0 and ITScore: the ensemble docking algorithm (Huang & Zou, 2007). The authors included the conformational ensemble as an adjustable variable in the IFD algorithm as opposed to performing serial docking across an entire ensemble. In addition to the usual six degrees of freedom for ligand flexibility, the authors added a protein conformation term, which allowed the protocol to simultaneously optimize the ligand conformation and choose the best protein structure for binding. In order to initially orient the ligand, a reference conformation comprising the largest possible binding site was generated from the combination of existing conformations. Ten separate protein ensembles were used for development and were obtained from the set by Cavasotto & Abagyan (2004), the set by Claussen *et al.* (2001), and two HIV-1 protease (HIVp) ensembles. Huang and Zou achieved an average success rate of 92% based on both pose and energy score, where the energy score was required to be within 1.0 kcal/mol of the native score or better and the five highest-ranked poses were included using a threshold for success of 2.5 Å RMSD from the experimental pose. They noted that docking accuracy based on binding pose alone was insufficient, as demonstrated by the influence of the binding energy upon results with rigid docking. Their ensemble method ran for a similar amount of time as one rigid docking calculation and was 12.9 times faster than serial docking (data averaged over ten systems, with an average of 10.5 structures).

RosettaLigand was introduced as an extension of the protein–protein docking program, RosettaDock (Meiler & Baker, 2006). In this paper, side chains were repacked during docking through the implementation of a rotamer library, while the vdW repulsion term was softened and ligand conformations were perturbed with MC. Their benchmark set included 100 complexes for native docking and 20 for cross-docking. They found a success rate (RMSD <2.0 Å) of 80% for native docking and 70% for cross-docking; the authors noted that their RMSD metric accounted for hydrogen atoms and side chains in the binding site, thus capturing small changes in conformation overlooked by other methods. RosettaLigand was updated to include backbone sampling through a gradient-based minimization of the torsion angles for both protein and ligand as the final step in their docking procedure (Davis & Baker, 2009; Davis *et al.* 2009). Although RosettaLigand has generated good results, with protein flexibility included, it requires 40–80 processor hours to generate a bound pose for the receptor–ligand complex and therefore is too time-consuming for use with large compound libraries.

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Other IFD methods for including protein flexibility include modeling flexibility for a limited number of receptor residues, Table 4a: GLIDE/PRIME (Sherman *et al.* 2006a, b), SCaRE (Bottegoni *et al.* 2008); discrete sampling, Table 4b: FlexX-Ensemble (Claussen *et al.* 2001), a modified DOCK protocol (Wei *et al.* 2004), Fleksy (Nabuurs *et al.* 2007), FLIPDock (Zhao & Sanner, 2007, 2008), FITTED (Corbeil *et al.* 2007, 2008; Corbeil & Moitessier, 2009), 4D Docking (Bottegoni *et al.* 2009); and continuum sampling, Table 4c: F-DycoBlock (Zhu *et al.* 2001), PC-RELAX (Zacharias, 2004, 2008), RosettaLigand (Davis & Baker, 2009), and implicit MD (Huang & Wong, 2009).

6. Ensemble docking

Ensemble docking differs from IFD in that protein flexibility is accounted for prior to the actual docking. Although the studies by Huang and Zou as well as Bottegoni *et al.* used a pre-existing ensemble of conformations, they sampled those conformations on the fly, and so they are included above in the IFD section. Two different types of methods exist for representing receptor flexibility during docking: grid-based ensembles and structure-based ensembles. Frequently, alternate protein conformations are represented on a two-body potential grid, enabling fast, inexpensive docking simulations. Flexibility can also be incorporated into binding predictions through the sequential docking to structures in a conformational ensemble or docking to an averaged/united receptor structure. Due to the time required for initial ensemble generation and repeated docking runs to a set of static structures, sequential docking is typically the most computationally intensive approach. However, it avoids the discovery of receptor–ligand complexes that are physically impossible (paradoxical ligands), which are sometimes seen in the results from docking to an average structure or grid.

6.1 Grid-based ensembles

The first docking method to use a composite grid was performed through DOCK3.5 in order to evaluate the impact of representing conformational variability as a structural ensemble on binding pose results (Knegtel *et al.* 1997). With HIVp, ras p21, retinol binding protein, and uteroglobin as their test cases, the authors compared the capacity of standard grids to energy- and geometry-weighted average grids to reproduce known binding poses and affinities. They found that docking scores were sensitive to the grid spacing and threshold parameters used to define the composite grid. The performance of the geometry-weighted grid was less dependent on the protein conformations that were available than the energy-weighted, to the extent that the geometry-weighted grid placed known ligands for HIVp within the top 21%, while the energy-weighted grid placed them in the top 33%. This occurred because the geometry-weighted grid ignored the repulsive potentials between flexible atoms, while the energy-weighted grid simply smoothed the repulsive potential, which can result in the retrieval of paradoxical inhibitors.

The incorporation of protein flexibility through docking to an interaction grid that represents the receptor ensemble is a common approach for SBDD, and additional methods are presented in Table 5, including those based on AutoDock 3.0 (Osterberg *et al.* 2002), Flog (Broughton, 2000), GRID/CPCA (Kastenholz *et al.* 2000), IFREDA (Cavasotto & Abagyan, 2004), ISCD (Zentgraf *et al.* 2006), and sets of consensus structures (van Westen *et al.* 2010).

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 4a. *Docking studies that modeled 2–3 residues to represent receptor flexibility during the performance of IFD*

Method	Target	Flexibility	Results	Caveats	Author
GLIDE/ PRIME	21 cases	Soft-potential docking to rigid receptor Ala replacement $\leqslant$3 residues for initial search Top 20 complexes refined via side-chain conformational search and minimization Redocked all complexes within 30 kcal/mol of minimum	Avg RMSD = 5.5 Å for rigid docking, 1.4 Å for IFD IFD failures mainly attributed to poor quality data for co-crystallized ligand	Very time intensive, limited to local motion	Sherman *et al.* (2006a, b)
SCaRE	16 cross-docking pairs	Optimal results with Ala replacement of two neighboring side-chain pairs Ligand docked to gapped conformation Clustered, optimized, and refined complex with original side chains	With pocket boundaries = all residues $\leqslant$5 Å from crystal pose and 1.0 kcal/mol vdW softening, best RMSD of docked pose to crystal structure (90% with rank 1) was 0.6–2.0 Å With pocket boundaries defined by Pocketome Gaussian Convolution algorithm, best RMSD of docked pose (80% with rank 1) was 0.3–2.0 Å and one failure Accurate prediction of binding pose and rank was more affected by scoring than by sampling	Limited view of active site flexibility Time consuming	Bottegoni *et al.* (2008)

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 4b. *Docking studies that applied discrete sampling of flexibility through panels of receptor structures during the performance of IFD*

Method	Target	Flexibility	Results	Caveats	Author
Flex X-Ensemble	105 cases	Averaged highly conserved regions Retained orientations of flexible regions as a set	FlexX-Ensemble yielded 66.7% success compared to 63.3% with FlexX within an RMSD of <2 Å for the first ten solutions (out of 40 hits) Scoring could not always identify the lowest RMSD complex within the top ten ranked solutions	May find ligands that are not compatible with the 'real' ensemble	Claussen *et al.* (2001)
Modified DOCK	T4 lysozyme mutant	Created a composite receptor, including rigid and flexible regions Calculated interaction energy between ligand and flexible regions independently Created average receptor representation for VS from results	Found 18 new hits Internal energy of conformer important for final ranking Conformational ensemble retrieved 79% of ligands in the top 1.5% of the database, compared to the single holo conformation that retrieved 77%, and the apo conformation that retrieved 54% Found that the receptor conformational energy was important to success	Conformational changes caused by ligand binding not fully predicted by method	Wei *et al.* (2004)
Ensemble as docking variable	10 cases	Conformational ensemble included as a variable with DOCK4.0 Representative for each side chain within the active site was based on the greatest distance from the reference sphere points by SPHGEN Optimized bound complex with SIMPLEX	Success defined as RMSD from crystal pose of $\leqslant$2.5 Å and an energy score $>$ native docking Ensemble docking had the same speed as rigid docking Ensemble docking had 67–100% success, rigid docking had 23–87% success, sequential docking had 33–100% success based on pose and score	Use of a unified representation of the receptor can result in the identification of high-scoring false positives	Huang & Zou (2007)
Fleksy	35 cases	Generated ensemble from backbone-dependent rotamer exploration with a soft potential Used FlexX-Ensemble with soft docking Flexible optimization of complex with Yasara	Cross-docking with Fleksy gave RMSD to the crystal pose (rank 1) of 0.5–5.3 Å and FlexX found 1.2–9.3 Å Successful cross-docking by Fleksy in 78% of all three sets, compared to 44% success for FlexX (rigid docking) Docking failures related to sampling problems in ligand placement, rotatable bonds, and receptor conformation	Cannot handle large changes in backbone conformation No consideration of solvent effects	Nabuurs *et al.* (2007)

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 4b (*cont.*)

Method	Target	Flexibility	Results	Caveats	Author
Flip DOCK	HIVp, Protein Kinase A	Flexibility Tree data structure represented conformational subspace Docking performed with AutoDock Divide-and-conquer genetic algorithm (GA) search performed substantially better than simple GA in best of 10 and average of 10 docking runs to 2 HIVp structures Also critical side chains in the active site were sampled (based on a rotamer library) during optimization	Performed 400 dockings (5 runs per complex) with 96.25% success (RMSD <2.0 Å) In a later paper (2008), the authors successfully docked 22/25 'tough' cross-docking cases (those that had failed rigid docking) from four proteins by making 3–6 side chains flexible in FLIPDock Failure possibly caused by sampling, side chains could adopt a less than favorable position due to a steric clash with the ligand Failure also attributed to inability of scoring function to distinguish alternate binding modes	Each degree of freedom must be selected by hand Side chains selected as critical for binding were not sufficient for complete success	Zhao & Sanner (2007)
FITTED 2.6	18 cases	Docked against rigid protein, MRC, or flexible protein with a modified GA for the receptor chromosome Allowed switching between conformations, side chains in binding site, and/or water positions	Success rate based on RMSD of docked pose to within 2.5 Å of crystal structure: 79% for rigid native-, 56% for rigid cross-, 67% for flexible-docking, and 67% for MRC Notable speed increase over previous versions	Accuracy decreased between FITTED 1.5 and 2.6 Explained as being due to omega-generated ligand structures and a different test set	Corbeil & Moitessier (2009) (earlier versions: Corbeil *et al.* 2007, 2008)
4D Docking	267 non-redundant structures	Generated MRC through EN-NMA Ensemble assembled onto 4D grid based on binding potential and superposition During docking, could switch receptor conformation as well as ligand conformation	4D docking 73% success rate with 3–8 conformers when the cognate receptor was not included For the same scenario, MRC had 71.1% success rate 4D docking was 4x faster than MRC Inadequate sampling occurred in 12.4% of the set with 4D docking, compared to 2.7% of the time for MRC Incorrect scoring resulted in the wrong pose being top-scored in 17.67% of the cases for MRC, and 10.3% for 4D docking	Performance decreased with >8 conformers 4D docking marginally less successful than MRC	Bottegoni *et al.* (2009)

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 4c. *Docking studies that applied molecular mechanics or minimization to incorporate receptor flexibility during IFD*

Method	Target	Flexibility	Results	Caveats	Author
F-Dyco block	HIVp, COX-2	Split ligands into fragments Performed LES and multiple-copy MD with a dynamic connecting algorithm Grid approximation or three ways to manage receptor flexibility	Recovered the ligand in 3 of 3 results for HIVp when full flexibility was allowed (except secondary-structure hydrogen bonds), compared to only recovering complex for 1 ligand with backbone-restrained flexibility Recovered the ligand in 7 of 76 results for COX-2	Simple clustering Dependent on threshold value and cluster position Grid negatively influenced incorporation of flexibility	Zhu *et al.* (2001)
PC-RELAX	apo FKBP12 rotamase; set of CDK2 structures	ENM-derived soft mode docking with four flexibility options: none, backbone minimization along NM, side-chain rotamers, or both	Apo structure gave best RMSD to crystal structure (rank ranges 1–7) of 0.8–1.8 Å for fully flexible versus (rank ranges 3–15) 0.8–4.3 Å for rigid Flexibility markedly improved ligand placement for certain cases, without influencing rank	Time intensive relative to rigid docking, cannot capture highly flexible motion	May & Zacharias (2008)
Rosetta ligand	JNK3 kinase and urokinase, 96 ligands from SAMPL-1	Assessed ligand complementarity with shape matching Included receptor flexibility through rotamer library for side-chain repacking during MC minimization (six cycles) Gradient-based minimization of ligand pose and receptor torsions	Average RMSD of top pose closest to crystal conformation was 0.45–6.89 Å (flexible docking) compared to 0.47–9.47 Å (rigid docking) for JNK3 kinase Average RMSD of top pose was 0.43–7.93 Å (flexible docking) compared to 0.52–8.27 Å (rigid docking) for urokinase Flexible docking never found 4/62 pairs and only 84% of the top-ranked compounds were within 2.0 Å (JNK3 kinase), never found 4/34 pairs and only 44% of the top-ranked were within 2.0 Å (urokinase) Docking failure sometimes caused by receptor sampling or the absence of water	Not as easily applicable to virtual screening	Davis *et al.* (2009)

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 4c (*cont.*)

Method	Target	Flexibility	Results	Caveats	Author
Implicit solvent MD	six protein kinases and phosphatases	Docked via 40 cycles of simulated annealing for each protein–peptide complex Restrained protein Cαs, water oxygens, ATP, and adenosine Used only one ns for simulations	Best results were achieved with docking via the distance-dependent dielectric solvation model and scoring according to GB molecular volume (GBMV), based on total system energy RMSD to crystal (rank 1) was 0.60–1.16 Å RMSD to extended pose (rank 1) was 1.30–4.09 Å	Scoring based on one solvation model alone was not always correct Sampling based on GBMV resulted in more failures	Huang & Wong (2009)

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 5. *Studies including receptor flexibility through grid-based ensemble docking*

Method	Target	Flexibility	Results	Caveats	Author
DOCK3.5 composite grids	Four cases	Geometry-averaged and energy-weighted grids	Geometry-averaged grids gave results that were less influenced by input conformation Geometry-weighted average gave RMSD of 0.4–1.6 Å Energy-weighted average gave RMSD of 0.4–4.0 Å	Recovers local minima as well as the active site May identify paradoxical inhibitors	Knegtel *et al.* (1997)
Flog	*L. casei* DHFR, murine COX2	Used snapshots from short MD (76ps) to generate averaged or weight-averaged grid for docking Weighted-average grid calculated areas of low structural variability among the structures Optimized the ligand in the binding site	Weighted-average grids performed better than the grid of the average or static crystal structure, selecting eight leads within the top 1% of the database, compared to six (crystal) and seven (average), with 16 possible Average and weight-averaged grids required less than half the CPU hours with optimization than using the crystal structure grid	Automated process, likely to result in incorrect results for certain complexes Could identify paradoxical binders	Broughton (2000)
GRID/ CPCA	3 serine proteases	Docking performed to crystal structures with GRID Scaling weight normalized probes Consensus PCA gave contour plots of MIF	Validated as a potential tool for enhancing ligand selectivity to a specific target Able to predict important cation-pi and hydrophobic interactions	Contour plots Only gives probe position, not orientation	Kastenholz *et al.* (2000)
AutoDock 3.0	HIVp (21)	Combined receptor conformations onto interaction energy grid Compared mean, minimum, clamped, and energy-weighted grids Retained structural waters	RMSD in energies was 1.34 kcal/mol for clamped, 1.43 kcal/mol for energy weighted, 7.69 kcal/mol for mean, 6.07 kcal/mol for minimum grids Weight-averaged grids performed best, retrieving the correct conformation 87% of the cases	Limited ability to map changes in conformation Could identify paradoxical binders	Osterberg *et al.* (2002)
IFREDA	PK (33), 1000 compound virtual screen	Multiple conformers generated by repeated flexible docking with known ligands Serially docked ligands to grid Merged results from conformations, then selected the best rank for each ligand	Average accuracy of 70% for cross-docking, with a threshold of 2.5 Å RMSD Only 49% of the ligands within the top 10% rank were also within 2.5 Å of the native state	Unable to map backbone shifts Inadequate solvation model	Cavasotto & Abagyan (2004)

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 5 (*cont.*)

Method	Target	Flexibility	Results	Caveats	Author
ISCD	Aldose reductase, thrombin, trypsin	Three single confomer grids were joined Repulsive layers between the individual grids	Top 15 ranked ligands had RMSD to the crystal structure <1.4 Å for aldose reductase Joined grids for thrombin and trypsin, docking found that seven or nine ligands preferentially docked to thrombin (cluster rank 1) The two ligands selective for trypsin bound trypsin (cluster rank 1) Computationally inexpensive relative to typical ensemble docking Potential conflict between desirability of full convergence to global minimum structure and lower-ranked states	Limited capacity for structural ensemble Could identify paradoxical inhibitors Difficult to know the appropriate search protocol beforehand	Zentgraf *et al.* (2006)
Consensus structures	HIV-RT (47)	Normalized B-factors, ligand-induced displacement, and consensus grids represented binding cavity	Identified two novel interaction sites Results supported by experimental work by Sweeney *et al.* (2008) Proposed combining theories of NNRTI activity	Requires available experimental structural information for consensus calculation	van Westen *et al.* (2010)

6.2 Structure-based ensembles

The multiple copy simultaneous search (MCSS) methodology was the earliest computational approach for mapping binding sites with functional group probes (Miranker & Karplus, 1991). The authors simultaneously minimized or quenched the probes by MD to the binding site of the hemagglutinin protein to identify favorable minima and found some correspondences between the positions of the minima and the functional groups on the ligand, sialic acid. While the authors noted that their method could account for a limited amount of side-chain flexibility by including multiple copies of the side chains, the published work was performed against a rigid structure.

Stultz and Karplus (1999) explored the influence of protein flexibility on the results from MCSS where five different protocols for placing two functional group probes, methanol and methyl ammonium, were used to search the binding surface of HIVp. MCSS calculations were performed for *1*: a minimized crystal structure, *2*: a conformation generated from quenched MD of the initial structure, *3*: an unrestrained structure, *4*: the output of *1* subjected to quenched MD with functional groups restrained and multiple side-chain copies, and *5*: a different rigid crystal structure of HIVp. The authors found that their results with protocols *4* and *5* yielded more favorable interaction energies than the reference protocol *1*. Protocol *4* gave valuable information for the ligand design, while the different low-energy minima from protocol *5* supported the idea of performing MCSS against an ensemble of structures and comparing the results; therefore, the authors suggested a combination of *4* and *5*. However, one of the limitations of MCSS and other ensemble methods is that they do not account for entropy or solvation during surface mapping, which hampers discrimination between druggable and irrelevant minima. Furthermore, as Schubert & Stultz (2009) point out in their review of MCSS, functionality maps are difficult to combine across different structures of a protein.

An original technique for structure-based ensemble docking was introduced by Carlson *et al.* (2000) as the dynamic pharmacophore model (Carlson *et al.* 1999), later termed the multiple protein structure (MPS) method. Two pharmacophores for HIV integrase were generated, a static model based on a rigid crystal structure and a dynamic model based on MD snapshots initiated from the same crystal structure. The MPS method mapped the dynamic binding surface of the protein with common functional groups by running a series of multi-unit search for interacting conformers (MUSIC) simulations (Fig. 3), wherein hundreds of probes were used to flood the protein surface and then minimized by MC sampling. During MUSIC simulations, all of the probe–probe interactions were ignored, which allowed probes to interact with the rugged binding potential of the protein and cluster into hot-spot positions along the protein surface. These cluster sites were used to develop a pharmacophore for virtual screening, and where the rigid model was unable to locate any of the test case inhibitors, the dynamic pharmacophore model not only identified known inhibitors, but it also predicted new inhibitors that were confirmed by experiment. The MPS method is one of the few ensemble-based methods that has been successfully used to find novel leads with demonstrated activity (Bowman *et al.* 2007; Damm *et al.* 2008).

McCammon and co-workers developed the relaxed complex scheme (RCS) as a technique for incorporating receptor flexibility prior to docking (Lin *et al.* 2002, 2003; McCammon, 2005). Initially the method was developed to generate conformational ensembles of the *apo* state through MD, which were then used in AutoDock to screen compound mini-libraries and score the receptor–ligand complexes. The 2003 study by Lin *et al.* examined FKBP-12 by RCS and implemented a final refinement step with MM/PBSA (Kollman *et al.* 2000) to yield final results.

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press



Fig. 3. MUSIC simulations of benzene probes (gray) to the surface of an HIVp monomer (purple). The initial stages are universal to mapping methods, but the combination of the structures in the fourth and final frames are unique to the MPS method. Five hundred probes were flooded onto the protein surface, minimized, and clustered together. Parent probes were identified for each monomer conformation, the HIVp monomers were aligned through a weighted-RMSD function, and then the parent probes were clustered together and retained when at least half of the protein conformations contained a probe at that site. This resulted in low-energy probe clusters (colored red through purple in the final frame) of benzene that could be combined with ethane and methanol clusters to develop a pharmacophore model.

Further updates to this method have included extension to virtual screening and reduction of the conformational ensemble to a representative configurational set (Amaro *et al.* 2008). RCS was successfully used to identify cryptic binding pockets and/or lead inhibitors of HIVp (Perryman *et al.* 2006), avian influenza neuraminidase (Cheng *et al.* 2008), acetylcholine binding protein (Babakhani *et al.* 2009), DNA polymerase β (Barakat & Tuszynski, 2010), MDM2/MDMX (Barakat *et al.* 2010), and cruzain (Durrant *et al.* 2010). Cheng *et al.* noted that conformational sampling markedly influenced the re-ranking of compounds based on interactions gained or lost compared to the crystal structure. The authors noted that the most dominant structures did not have the lowest binding energy, but given the limitations of MD, they believed that the dominant clusters were good representatives of the ensemble. Babakhani *et al.* suggested that both the scoring function and the conformational variability contributed to several failures in reproducing known binding modes.

To determine the best process for developing a structural ensemble, Rueda *et al.* (2010) used ICM to examine the same set of 1068 conformations from 99 pharmaceutically relevant proteins that was used in the 4D docking study. The authors judged performance in virtual screening based on the area under the curve (AUC) of a ROC plot and found that holo conformations yielded improved AUC values compared to apo conformations. It was interesting that the authors found that ensembles of holo plus apo conformations did not significantly improve results compared to holo-only ensembles. In cases without a holo structure, the authors recommended the use of an apo structure with the largest pocket volume. The authors proposed the use of a ligand-guided approach to find the optimal protein conformation, but stated that docking to an ensemble would be an acceptable substitute in the absence of ligand activity data. Several groups have found that there tends to be a single receptor conformation that grants the best performance in docking studies (Barril & Morley, 2005; Rueda *et al.* 2010).

NMA has been frequently used as a tool for examining protein flexibility relevant to ligand binding (Cavasotto *et al.* 2005; Keseru & Kolossvary, 2001; May & Zacharias, 2005, 2008;

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

May *et al.* 2008; Zacharias, 2004; Zacharias & Sklenar, 1999; Zavodszky & Kuhn, 2005). It is important to note that low-frequency modes identify large-scale motions, while high-frequency modes identify small-scale motion. Rueda *et al.* (2009) showed that they could use NMA on elastic network models for all heavy atoms of the receptor in order to derive an ensemble of protein conformations and improve results from cross-docking in ICM. Fourteen proteins, each with two structures and their cognate ligands, were used as the benchmark test set. Optimal results involved the generation of 100 different conformations from NMA on all heavy atoms within 10 Å of the ligand-binding site. Larger conformational ensembles, such as those with 200 members, negatively affected false positive rates. Another NMA-based method for conformational selection was based on cyclin dependent kinase-2 (CDK-2), because it had a sufficient amount of available experimental data on ligand binding for validation of the NMA protocol (Sperandio *et al.* 2010). Unlike most other NMA-based methods, Sperandio *et al.* included all of the protein atoms in the calculation in order to better represent conformational changes and found that too many protein conformations negatively impacted docking results by recovering additional false positives. They suggested addressing this by developing scoring protocols specifically tailored to the use of a conformational ensemble.

Of the variety of techniques that have been developed to improve binding predictions for protein–ligand complexes, the use of a conformational ensemble is perhaps the most common approach. Representative cases and novel methods based on structure-based ensemble docking are listed in Tables 6a–6d. Crystal structure and NMR-based ensemble methods include docking to a large set of conformations (Barril & Morley, 2005; Craig *et al.* 2010), and ICM (Rueda *et al.* 2010), as listed in Table 6a. Procedures that rely upon molecular dynamics to provide receptor flexibility include MCSS (Stultz & Karplus, 1999), MPS (Carlson *et al.* 1999, 2000), MD + LigBuilder (Mustata & Briggs, 2002; Mustata *et al.* 2003), RCS (Lin *et al.* 2003), MD + AutoDock (Frembgen-Kesner & Elcock, 2006), Flo98 (Popov *et al.* 2007), a reduced receptor ensemble (Landon *et al.* 2008), REMD/PLOP (Wong & Jacobson, 2008), the ensemble reduction method (Bolstad & Anderson, 2009), enhanced molecular docking (Kranjc *et al.* 2009) and MD + Glide (Nichols *et al.* 2011), as listed in Table 6b. Ensembles have also been generated for docking through NMA, as demonstrated by FIRST/ROCK/SLIDE (Zavodszky *et al.* 2004), low-frequency NMA (Cavasotto *et al.* 2005), ENM (Lindahl & Delarue, 2005), EN–NMA (Rueda *et al.* 2009), and all atom NMA (Sperandio *et al.* 2010), as listed in Table 6c. Finally, ligand-induced flexibility has also been used as another source for conformers, as shown in Table 6d and applied by MCSA-PCR (Ota & Agard, 2001), Skelgen (Firth-Clark *et al.* 2006), QSiCR (Subramanian *et al.* 2006, 2008), PhE/SVM (Leong, 2007; Leong & Chen, 2008; Leong *et al.* 2009), structure prediction (Bisson *et al.* 2007), and Surflex (Jain, 2003, 2007, 2009). Although listed in Table 5 due to its incorporation of protein flexibility within an interaction grid, IFREDA also relies upon ligand-induced conformations for contributing flexibility.

7. *De novo* approaches through fragment-based drug design

Several interesting experimental methods that naturally account for protein flexibility have greatly influenced computational SBDD. The burgeoning field of FBDD has allowed for the identification of low-molecular weight binders, which allow for the identification of difficult binding sites (Erlanson *et al.* 2004). First introduced by the multiple solvent crystal structure (MSCS) (Mattos & Ringe, 1996) and SAR-by-NMR (Shuker *et al.* 1996) techniques, FBDD can be performed via NMR or x-ray crystallography, and provides a different set of hits from

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 6a. *Studies including full receptor flexibility through collection of a crystal or NMR structural ensemble*

Method	Target	Flexibility	Results	Caveats	Author
Docking to a structural ensemble	Hsp90 (149), CDK2 (49)	Compared use of a structural ensemble to rigid docking 57 Hsp90 ligands and 34 CDK2 ligands	For CDK2, single best cavity gave 64% success, six best cavities gave 94%, full ensemble gave 76% For Hsp90, the use of pharmacophoric restraints improved docking, with multiple cavity docking having 86% average success compared to single cavity docking having 57% average success Experimental poses were retrieved but not among the top-rank complexes	Incorrectly assumed negligible difference in conformational free energy Use of multiple cavities for virtual screening was limited by false positives	Barril & Morley (2005)
Docking to a structural ensemble	BACE, cAbl	Enrichment from ensemble docking compared to rigid docking for choosing inhibitors over decoys Used five different strategies for combining receptor structures into an ensemble Docked in Glide	AUC for rigid BACE structures ranged from 0.688 to 0.778 Construction of an ensemble from receptors which demonstrated optimal success during rigid docking against different chemotypes yielded the best performing ensemble	Technique was based on improving enrichment against known inhibitors rather than exploring conformational space Did not account for desolvation or energy of the receptor	Craig *et al.* (2010)
ICM	99 cases	Studied the impact of using holo versus apo-protein conformations for virtual screening based on AUC scores	Found that there is one receptor conformation within the ensemble that gives the optimal performance Ensemble docking performed marginally better than a single average structure: 0.88 ± 0.18 mean AUC versus 0.78 ± 0.22 Holo conformations outperformed apo conformations in most cases	Significant overlap in the AUC range for ensemble docking compared to an average structure	Rueda *et al.* (2010)

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 6b. *Studies including full receptor flexibility through a structural ensemble previously generated by MD*

Method	Target	Flexibility	Results	Caveats	Author
MCSS	HIVp	Used quenched MD to generate new conformations or mapped a different conformation from the reference Flooded protein surface with functional group probes to identify regions of consensus	Local optimization of selected probes from MCSS improved interaction energy Protocols 2 and 3 yielded functionality maps that were the most divergent from the reference Protocol 4 yielded the most low-energy minima	Recovers local minima in addition to the binding site Very minor amount of protein flexibility included	Stultz & Karplus (1999)
MPS/dynamic pharmacophore	HIV integrase	Used MD to generate new conformations from a crystal structure Flooded probes to protein surface and determined consensus clusters	Flexible model recovered known inhibitors from a virtual screen Rigid model did not recover any known inhibitors Identified novel inhibitors for binding the target with experimental verified activity	No desolvation penalty Can recover local minima	Carlson *et al.* (1999, 2000)
MD + LigBuilder	Alanine Racemase	Dynamic pharmacophore modeling similar to MPS, but can simultaneously search with multiple probes LigBuilder mapped surface properties of each conformation (11 total) after MD	Dynamic pharmacophore model identified 34 hits out of the set of 43 known binders, compared to the 27 identified by the static model	Very brief phase of MD for conformation generation	Mustata & Briggs (2002)
RCS	FKBP-12	Conformations generated by MD Rapid docking in AutoDock Refinement of high scoring complexes with MM/PBSA, which was necessary to recover the crystal pose	Correctly ranked the crystallographic pose as the highest rank with MM/PBSA for all three small molecules Noted that docking results were highly sensitive to protein conformation	Can be time-consuming Very short MD simulation (2 ns) for conformer generation Free energies were distributed within 2–3 kcal/mol, potentially resulting in misranking a binder as weak/potent	Lin *et al.* (2003)

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 6b (*cont.*)

Method	Target	Flexibility	Results	Caveats	Author
Explicit solvent MD/AutoDock	p38 MAPK (5000 MD snapshots)	Performed 10 MD simulations of 30 ns at 1000 K, continued 3 runs to 60 ns upon conformational divergence from the initial structure Harmonic restraint on all heavy atoms, excluding the activation loop and nearby residues Applied AutoDock using snapshots from single simulation	Found 2 novel conformations of DFG motif Successfully docked 5 known inhibitors Conformers from simulation identified cryptic binding sites Docking failures were judged to be the result of conformational sampling	Extensive MD simulation (390 ns total) render this technique impractical for some SBDD applications	Frembgen-Kesner & Elcock (2006)
Flo98	*C. hominis* DHFR, *T. gondii* DHFR	Simulated annealing of active site MC search for optimal bound pose Averaged energy of 25 lowest energy protein-ligand complexes Developed homology model for TgDHFR	Calculated binding affinity was 72.9% correlated to experiment (ChDHFR) Identified alternate binding site Correlation between docking and activity (TgDHFR) was 50.2% by R^2 Found that an averaged energy value outperformed individual values, by reducing error in pose evaluation by the scoring function	Enables limited flexibility of binding site residues only during global MC docking	Popov *et al.* (2007)
Reduced Receptor Ensemble	Avian flu neuramindase	Similar to MPS Generated MD ensemble Flooded ensemble with probes (CS-Map), created pharmacophore	Predicted novel hot spots potentially relevant for *de novo* ligand design Proposed a new inhibitor class	No experimental support of sites	Landon *et al.* (2008)
REMD/PLOP	Six cases	REMD used to generate low-energy loop conformations After clustering, loops refined with PLOP Lowest energy conformer used for docking with GLIDE	RMSD of docked pose between holo and predicted structure was 1.4–12.5 Å RMSD of docked pose between crystal and holo structure was 1.0–2.5 Å Failures resulted from sampling deficiencies and structural features unrelated to the loop Scoring favored closed conformation loops	Limited by efficiency of REMD for generating loop conformations	Wong & Jacobson (2008)

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 6b (*cont.*)

Method	Target	Flexibility	Results	Caveats	Author
Ensemble reduction method	DHFR	Generated MD ensemble Used relative difference of interatomic distances for ensemble structures compared to an experimentally determined holo structure to define optimal conformations	Found that conservation of essential distances between the residues that were important for binding was best for selecting a representative ensemble Pruned the structural ensemble by 50–75% while retaining docking accuracy Some structures scored poses well but placed ligands incorrectly	Maintaining conserved core distances may limit exploration of important large-scale shifts No correlation between relative difference and docking score or correct pose	Bolstad & Anderson (2009)
Enhanced molecular docking	Human prion protein	Used 20-ns MD to generate 20 conformations for docking along with NMR structure Docked with AutoDock and GOLD Clustered results, ran 10 ns of MD, then six independent metadynamics simulations to find the free energies of binding/dissociation	Calculated dissociation dG was 7.8–8.6 kcal/mol while experimental dissociation dG was 7.5 kcal/mol Predicted multiple binding sites Affinities agreed with NMR experiment	Computationally intensive	Kranjc *et al.* (2009)
Explicit solvent MD/Glide	Reverse transcriptase (10 000), W191 G (7500)	Generated a conformational ensemble from MD of holo- and apo-crystal structures Docked ligand/decoy set to conformers in Glide	Found that MD could be used to move a conformation into a predictive range for docking MD and AUC were not correlated Identified a correlation between the average predictive power and the average flexibility of the binding site, such that highly flexible sites had less utility for docking Concluded that a broadly applicable protocol for the application of a structural ensemble to docking is still a distant goal	No single feature can be used to pick out conformations May require extensive knowledge of the system	Nichols *et al.* (2011)

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 6c. *Studies including full receptor flexibility through a normal mode analysis, or other non-MD-based conformational generation, prior to docking*

Method	Target	Flexibility	Results	Caveats	Author
FIRST/ ROCK/ SLIDE	CypA, estrogen receptor	Analyzed flexibility potential (FIRST) and sampled rotational bonds to generate receptor ensemble (ROCK) SLIDE accounted for side-chain flexibility and solvation	Representative ensemble similar to NMR structures Captured important hydrogen bonds present in receptor–ligand complex	Compared score distributions, no individual assessment of relative likelihood/energy for generated conformers	Zavodszky *et al.* (2004)
Low-frequency Cα NMA	cAPK Kinase	Used NMA to generate alternative conformations based on relevance to loop plasticity Minimized side chains while bound to non-native ligand	Improved docking and enrichment scores relative to rigid docking RMSD to the native pose for rank 1 was 0.6–4.0 Å (one outlier at 7.0 Å) for docking to NMA conformer and 0.4–10.2 Å for docking to crystal structure Docking failures due to the scoring function; all compounds docked correctly to the apo structure ensemble, but only 50% were found in the top 10% of scores	Used known binders, which may bias the available space for lead discovery	Cavasotto *et al.* (2005)
Elastic network model	Six protein-ligand cases	Low frequency NMA to optimize docked receptor–ligand complexes and identify new conformational space Used eigenvectors of *apo* protein	Improved the global coordinate RMSD by $\leqslant$3.2 Å of the predicted complex to the experimental pose Refinement with restraints on mode amplitude performed much better than refinement with no restraints or by MD/NM minimization	Most successful for systems where flexibility can be represented well in only a few modes	Lindahl & Delarue (2005)
EN-NMA	14 cases	Used NMA to generate 100 distinct conformers for heavy atoms with 10 Å of the binding site Geometric clustering of side-chain heavy atoms to reduce overlap (threshold of 0.6 Å)	Ensemble of 66 ± 24 conformers optimal From NMA-based conformers, able to correctly select 20 of 28 cross-docking cases among the top 5 results for each run Failures due to sampling, neither cross-docking nor EN-NMA could model 'induced fit' variation between conformers	Experimental structure necessary in ensemble <25 Large conformational ensemble (200) can increase false positive rates	Rueda *et al.* (2009)

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 6c (*cont.*)

Method	Target	Flexibility	Results	Caveats	Author
All atom NMA	CDK2	Used all of the protein atoms for conformational selection Pursued relevant structures from the 20–25 lowest modes Required a deviation of at least 1 Å, but no more than 8 Å, from the minimized crystal structure to retain new conformations	Over all cases, successful docking occurred for 54.3% (holo-crystal structure), 58.1% (apo-crystal structure), 42.8% (minimized apo-crystal structure), 70.4% (best NMA structure), 55.7% (second-best NMA structure), and 55.2% (third-best NMA structure)	Large conformation set can result in the discovery of many false positives NMA is not applicable to small local motions or large domain shifts	Sperandio *et al.* (2010)

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 6d. *Studies including full receptor flexibility through ligand-induced flexibility*

Method	Target	Flexibility	Results	Caveats	Author
MCSA–PCR	Major histocompatibility complex	Ligand grown into binding pocket with each SA cycle Generated pseudo-crystallographic density map	Side-chain RMSD of VSV8 peptides from explicit simulation was 2.15 Å for N/C-termini restraints, 1.95 Å for CA restraints, and 1.78 Å for backbone restraints Side-chain RMSD of VSV8 peptides from *in vacuo* simulation with dielectric of $4r$ was 2.84 Å for N/C-termini restraints, 2.61 Å for CA restraints, and 2.41 Å for backbone restraints	Requires a large computational effort, not suitable for VS	Ota & Agard (2001)
Skelgen	Estrogen receptor-α (7)	*De novo* design through serial-dock Ligands generated from fragment design Two different pharmacophoric constraints applied	Selected top 25 complexes for each constraint set, yielding 350 compounds for testing Identified four novel lead compounds Best had $IC_{50} = 340$ nm	Dependent on pharmacophoric constraints Time and material intensive	Firth-Clark *et al.* (2006)
QSiCR	CDK2, p38 MAPK	Used known binders to generate protein flexibility Built conformers from structural fingerprints (MACCS) and topological details	Average $R^2 = 0.71$ for active site size and distances of CDK2, including mutant structures Average $R^2 = 0.72$ for predicted backbone conformational changes of p38 MAPK Conformational variability negatively affected results	Does not fully represent available interaction space Highly dependent on ligand training set	Subramanian *et al.* (2006, 2008)
Predictive modeling	Androgen receptor	Predicted receptor conformation from backbone conformation Optimized side chains using biased probability MC Repeated selection of receptor–ligand complex ~15 times to improve discrimination of agonist and antagonist binding	Potentially identified candidates for drug repurposing with antiandrogenic activity, based on cross-docking and competitive binding assays High-scoring ligands appeared promising, but their poses differed from orientations of known drugs	Time intensive Requires *a priori* knowledge and similar protein structures for initial predictive modeling	Bisson *et al.* (2007)

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Table 6d (*cont.*)

Method	Target	Flexibility	Results	Caveats	Author
PhE/SVM	Ether-à-go-go Related Gene	Receptor flexibility through a pharmacophore ensemble from 26 training set and 13 test set compounds Potential to enable protein plasticity even when structural information is inadequate for other methods	Predicted IC_{50} values compared to observed IC_{50} values had R^2 of 81%	Cannot identify large structural shifts	Leong (2007)
Surflex	85 cognate cases, eight cross-docking cases	Updated to allow structural ensembles and post-docking refinement Created an idealized ligand to model the active site Ligand was built into the active site through incremental construction/ combination	Cognate docking 76% had top score within 2.0 Å Cross-docking with multiple structures and pocket adaption: 60.9% average success after docking, 66.7% after refinement of pocket protons and rescoring, 75.5% after using the best of two pose families Failure was discussed in terms of ligand sampling	Moderate computational expense For some cases, using ligand sub-fragments to guide docking could have a negative impact	Jain (2009)

high-throughput screening (HTS) experiments. Screening proteins with a fragment collection both verifies the druggability of a target system and identifies fragments that can be linked and optimized to develop a viable lead compound (Wiesmann *et al.* 2004). Importantly, hits from FBDD have been experimentally validated and identified novel compounds shown to work in the clinic (Howard *et al.* 2009; Murray & Rees, 2009).

The MSCS technique was originally introduced as a crystallography tool for the characterization of the binding potential of pancreatic elastase. Other groups have also used MSCS to probe ligand interactions in thermolysin (English *et al.* 1999, 2001), subtilisin (Fitzpatrick *et al.* 1993; Schmitke *et al.* 1997; 1998), RNase A (Fedorov *et al.* 1996; Dechene *et al.* 2009), p53 core (Ho *et al.* 2006), lysozyme (Wang *et al.* 1998), and H-Ras (Buhrman *et al.* 2003). During crystallization, the water and organic solvent acted as probes of binding affinity because the organic solvent, which was chosen to represent a common functional group, displaces bound waters only at locations with favorable affinity for the particular interaction type presented by the protein. MSCS can be performed with a variety of probes, thus enabling the results from the protein crystallography experiment to be superimposed in order to identify regions of consensus binding.

When Joseph-McCarthy *et al.* (1996) compared their results from studies with MCSS to results from MSCS, they found that it was important for traditional MCSS to consider occupancy at a first site in order to accurately predict occupancy at a second binding site. English *et al.* (2001) compared the use of MSCS, GRID, and MCSS through the exploration of the binding surface of thermolysin with acetone, acetonitrile, isopropanol, and phenol acting as functional group probes. The authors found that in comparison with MSCS, the MCSS, and GRID methods overestimated electrostatic interactions, retrieved an overabundance of local energy minima, and could not provide detailed descriptions of the interactions between the protein and probes.

Although both SAR-by-NMR and MSCS allow for a larger amount of protein flexibility then current methods for FBDD do, they are also both material and time-intensive. Furthermore, as discussed by English *et al.* (2001), MSCS can be quite slow and probe choice is limited by the fragility of protein crystals. Computational solvent mapping is a complementary approach to experimental fragment-binding studies; however, most computational approaches do not account for either the impact of protein flexibility or proper solvation effects, which can lead to poor mapping results. Over the past two years, several groups have developed new methodology for including flexibility and solvent competition in computational fragment mapping. While the results of any fragment-based study can be translated into a consensus pharmacophore model and used in virtual screening applications, the primary objective of MSCS and computational solvent mapping has been the correct identification of potential binding sites.

Locus Pharmaceuticals developed a method based on a grand-canonical (GC) MC simulation for pharmacophore development (Clark *et al.* 2006, 2009; Moore, 2005). Both studies from Clark *et al.* were performed against a static structure. Their later study concentrated on deriving accurate ΔG_{bind} , while their earlier study used ten simulated annealing runs of GC-MC to explore the protein surface of thermolysin with 2 probes and of T4 lysozyme with 14 probes (2006). Based on the simulated annealing runs, GB/SA was used to predict binding affinities and Clark *et al.* found that their results retrieved some of the hot spots observed in the published MSCS data. Moore discussed the use of torsion-space dynamics to generate protein flexibility in combination with general outcomes from the use of a 500-member fragment library to search target surfaces with GC-MC; however, no specific details were given.

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Vadja and co-workers developed a fast algorithm for searching protein surfaces with small organic probes that was based on a fast Fourier Transform correlation approach (FTMAP) (Brenke *et al.* 2009). In a manner akin to MPS, billions of solvent probes are minimized to the protein surface and then clustered based on a simple greedy algorithm. The probe clusters demonstrate the position and orientation of proposed hot spots along the target, where the largest cluster is defined as the maximal hot-spot location; the second largest is the second most-important site; and etc. Although FTMAP is another method for computational solvent mapping that does not inherently include protein flexibility, receptor dynamics could be modeled through serial searches over multiple conformations.

A novel method for probe mapping incorporated solvent competition and protein flexibility through MD (Seco *et al.* 2009). Based on MSCS, the authors used isopropyl alcohol (IPA) and water together to perform a single MD simulation over 16 ns. The use of binary solvent MD inherently accounted for solvation effects and protein flexibility. This method was applied to three proteins with experimental MSCS results and five pharmaceutically relevant receptors that have not been studied by MSCS; simulation probe occupancy was used to differentiate binding sites and predict ΔG_{bind}. This method was much more computationally demanding than other computational solvent-mapping techniques such as FTMAP (Brenke *et al.* 2009), but may be more competent at distinguishing between druggable and non-druggable sites because of the use of flexibility data. Yang & Wang (2010) used this same solvent-mapping technique to study hot-spot mapping against thermolysin together with a double-decoupling method, which provided a more rigorous calculation of ΔG_{bind} at potential hot spots. It is important to note that the analysis of probe occupancy to locate binding sites necessarily assumes that Boltzmann sampling occurred during simulation; therefore, careful selection of the trajectory length required for convergence is exceedingly important but frequently neglected in method development.

Guvench & MacKerell (2009) published a similar method, Site Identification by Ligand Competitive Saturation (SILCS), where small molecule probes were used to examine the binding surface. Their technique used all-atom MD simulations of protein in a box of propane, benzene, and water as the explicit solvent probes. Ten independent simulations over 5 ns were generated for analysis and the mapping results were represented on a 1 Å grid of volume occupancy. SILCS was capable of reproducing known binding interactions for the test case, BCL-6 oncoprotein, and a follow-up study further validated the approach on seven proteins from five different protein families (Raman *et al.* 2011). In their follow-up study, MacKerell and co-workers extended the simulation time to 10 runs of 20 ns each and scored their occupancy results based on a normalized calculation of the observed:expected occupancy, which they termed grid free energy (GFE). When compared to the positions of known ligands, their GFE was able to select for the crystallographic pose of the bound ligand in several protein families. However, in the only available figure of an entire protein surface, SILCS is clearly shown to preferentially map many irrelevant minima before identifying the binding site.

We have developed a computational method that achieves hot-spot mapping results that are similar to available experimental data; Mixed-solvent Molecular Dynamics (MixMD). Our MPS method (Carlson *et al.* 2000) for pharmacophore development demonstrated success in mapping protein systems for drug design (Bowman *et al.* 2007; Damm *et al.* 2008) and MixMD complements MPS while simultaneously allowing protein flexibility and probe competition with water. MixMD was inspired by the MSCS technique but incorporates more explicit conformational sampling. Since MSCS results identified specific electron density that can be attributed to the placement of organic solvent along the protein surface, we specifically compare grids of solvent

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

occupancy from simulation to that of experimental electron density. Using many, short MD simulations of protein in 50% weight/weight mixtures of acetonitrile and water, we successfully validated MixMD based on the available MSCS structures (Lexa & Carlson, 2011). Similar results were obtained with isopropanol (unpublished data).

8. Limitations of structure-based drug design

Most published methods for flexible protein–ligand docking are based on limited sets for benchmarking results. The use of a large test set for methods development is crucial for demonstrating performance across a range of different targets. Furthermore, docking to a flexible protein is more resource and time-intensive than docking to a rigid receptor and in many cases ensemble docking is unrealistic for screening huge compound libraries. Through the careful development of methods for docking and scoring, the computational requirements for accurate simulation may be sufficiently lessened to allow for more widespread implementation of flexibility.

The use of multiple conformations for docking is limited in that several protein conformations may decrease the selectivity of lead compounds by increasing the false positive rate. Using combinations of features from different conformations may also lead to the creation of a ligand with high affinity for an average receptor structure that is not experimentally accessible, a so-called 'paradoxical inhibitor'. As a result, there is the potential for introduction of bias through user intervention in many of these methods for flexible docking.

Due to the rugged landscape of most proteins, not every conformation that is included in a low-energy ensemble will adequately represent its true binding potential. This has led many recent studies to dedicate their focus to identifying the optimal method for selection of only the relevant protein conformations. For this process of conformer selection and weighting to succeed, it is crucial that the internal energy of the individual protein conformations be included in the scoring process (though this is more difficult to properly calculate). Furthermore, it is necessary that the scoring functions used in fully flexible procedures be as accurate as possible so as to provide the most physically realistic results.

Within the field of SBDD, it remains unclear whether sampling or scoring is the limiting factor in our ability to correctly predict binding modes and affinities. Hampering our ability to address the clear need for improved methods is an unfortunate lack of careful documentation in the literature covering docking or scoring failures for a given study. Careful examination of the techniques presented in this review yielded 30 papers with enough detail to allow assessment of the possible factors resulting in inaccurate results. Of these papers, 14 leaned toward scoring being the dominant factor, 11 toward sampling, four were evenly split, and one toward poor crystallographic density. This depicts the unfortunate scenario where both sampling and scoring can negatively impact success in the field. The results from several papers convincingly argue that the combination of sampling and scoring can be applied to predict better complexes than those identified without flexibility. Beyond applying RMSD metrics to experimental structures, there is no extensive evidence for whether the current sampling approaches model changes in conformation accurately; while some results successfully predict known experimental poses, other results conflict with, or have not been corroborated by, experimental data. It is important that consistent data and analysis be presented in the future so that these sources of error in SBDD can be better understood.

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Experimental FBDD approaches such as SAR-by-NMR and MSCS are time-consuming, costly, and dependent upon the receptor size and its potential for crystallization. On the other hand, computational FBDD approaches can be just as time-consuming and are heavily dependent on proper parameterization and analysis. The trend toward estimating binding affinity from computational solvent grids has shown only partial success when compared with experimental affinities. Computational FBDD shows great promise as a tool for including flexibility and probe-water competition explicitly in a 'docking' experiment; however, implementation of these methods requires careful consideration of the underlying chemistry to ensure reliable results.

9. Future directions

Considerable progress in the field of SBDD is still required before achieving the ultimate goal: accurately predicting the native state conformation of novel receptor–ligand complexes. Since there is a strong dependency upon the test set for both docking and scoring functions, the use of the high-quality test sets is essential for method development and the sets from Murray and co-workers have given researchers a resource for building docking and scoring functions that are applicable across a wide range of protein targets (Hartshorn *et al.* 2007; Verdonk *et al.* 2008). Some of the congeneric series in the CSAR-NRC set developed in our group are also very valuable to this effort (Dunbar *et al.* 2011; Smith *et al.* 2011). As technology for structure determination advances, the structures of many more pharmaceutically relevant receptors can be solved. Additionally, as structural information grows, so too does our ability to develop reliable homology models that closely mimic the native conformation of the protein; thus allowing the continued exploration of new targets for drug discovery. However, the successful use of homology models in SBDD is obviously wedded to accuracy in methods for full flexibility in docking and scoring.

Computational methods for SBDD aim to use docking and *de novo* methods to shorten the period of time required for lead discovery and optimization. As we improve our ability to represent protein flexibility in the presence of potential binders, structurally diverse binding sites can be better characterized in order to improve predictions of cross-reactive or promiscuous ligands. Additionally, allowing conformational variability in SBDD can assist particular goals in lead optimization, such as broad-spectrum treatments or improved selectivity for a specific homologue.

Unfortunately, the timeline for SBDD programs in industry requires that results be quickly obtained in order to contribute to lead development. Techniques for incorporating protein flexibility are quite exciting, but they often increase computation time as well as the number of hits. For fully flexible methodology to become more widely applicable, scoring functions must be devised that can identify the optimal conformers for the receptor and decrease false positive hits. Although many VS papers still measure accuracy and success in terms of recovery rather than experimentally validated leads, scientific journals are beginning to require experimental evidence to support VS success. Hopefully, this trend will inspire the development of more accurate scoring functions.

10. Acknowledgements

This work has been supported by the National Institutes of Health (GM65372). KWL acknowledges Rackham Graduate School, the Pharmacological Sciences Training Program (GM07767), and the American Foundation for Pharmaceutical Education for funding.

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

11. References

Abagyan, R., Totrov, M. & Kuznetsov, D. A. (1994). ICM – a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. *Journal of Computational Chemistry* **15**, 488–506.

Alberts, I. L., Todorov, N. P. & Dean, P. M. (2005a). Receptor flexibility in de novo ligand design and docking. *Journal of Medicinal Chemistry* **48**, 6585–6596.

Alberts, I. L., Todorov, N. P., Kallblad, P. & Dean, P. M. (2005b). Ligand docking and design in a flexible receptor site. *QSAR and Combinatorial Science* **24**, 503–507.

Alonso, H., Bliznyuk, A. A. & Gready, J. E. (2006). Combining docking and molecular dynamic simulations in drug design. *Medicinal Research Reviews* **26**, 531–568.

Amaro, R. E., Baron, R. & McCammon, J. A. (2008). An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. *Journal of Computer-Aided Molecular Design* **22**, 693–705.

Anderson, A. C., O'Neil, R. H., Surti, T. S. & Stroud, R. M. (2001). Approaches to solving the rigid receptor problem by identifying a minimal set of flexible residues during ligand docking. *Chemistry and Biology* **8**, 445–457.

Babakhani, A., Talley, T. T., Taylor, P. & McCammon, J. A. (2009). A virtual screening study of the acetylcholine binding protein using a relaxed-complex approach. *Computational Biology and Chemistry* **33**, 160–170.

Barakat, K., Mane, J., Friesen, D. & Tuszynski, J. (2010). Ensemble-based virtual screening reveals dual-inhibitors for the p53-MDM2/MDMX interactions. *Journal of Molecular Graphics and Modelling* **28**, 555–568.

Barakat, K. & Tuszynski, J. (2010). Relaxed complex scheme suggests novel inhibitors for the lyase activity of DNA polymerase beta. *Journal of Molecular Graphics and Modelling* **29**, 702–716.

Barril, X. & Morley, S. D. (2005). Unveiling the full potential of flexible receptor docking using multiple crystallographic structures. *Journal of Medicinal Chemistry* **48**, 4432–4443.

Berger, C., Weber-Bornhauser, S., Eggenberger, J., Hanes, J., Pluckthun, A. & Bosshard, H. R. (1999). Antigen recognition by conformational selection. *FEBS Letters* **450**, 149–153.

Berman, H. M., Bhat, T. N., Bourne, P. E., Feng, Z., Gilliland, G., Weissig, H. & Westbrook, J. (2000). The Protein Data Bank and the challenge of structural genomics. *Nature Structural Biology* **7** (Suppl), 957–959.

Bisson, W. H., Cheltsov, A. V., Bruey-Sedano, N., Lin, B., Chen, J., Goldberger, N., May, L. T., Christopoulos, A., Dalton, J. T., Sexton, P. M., Zhang, X. K. & Abagyan, R. (2007). Discovery of antiandrogen activity of nonsteroidal scaffolds of marketed drugs. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 11927–11932.

Bolstad, E. S. & Anderson, A. C. (2008). In pursuit of virtual lead optimization: the role of the receptor structure and ensembles in accurate docking. *Proteins* **73**, 566–580.

Bolstad, E. S. & Anderson, A. C. (2009). In pursuit of virtual lead optimization: pruning ensembles of receptor structures for increased efficiency and accuracy during docking. *Proteins* **75**, 62–74.

Bottegoni, G. (2011). Protein–ligand docking. *Frontiers in Bioscience* **16**, 2289–2306.

Bottegoni, G., Kufareva, I., Totrov, M. & Abagyan, R. (2008). A new method for ligand docking to flexible receptors by dual alanine scanning and refinement (SCARE). *Journal of Computer-Aided Molecular Design* **22**, 311–325.

Bottegoni, G., Kufareva, I., Totrov, M. & Abagyan, R. (2009). Four-dimensional docking: a fast and accurate account of discrete receptor flexibility in ligand docking. *Journal of Medicinal Chemistry* **52**, 397–406.

Bowman, A. L., Nikolovska-Coleska, Z., Zhong, H., Wang, S. & Carlson, H. A. (2007). Small molecule inhibitors of the MDM2-p53 interaction discovered by ensemble-based receptor models. *Journal of the American Chemical Society* **129**, 12809–12814.

Brenke, R., Kozakov, D., Chuang, G. Y., Beglov, D., Hall, D., Landon, M. R., Mattos, C. & Vajda, S. (2009). Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. *Bioinformatics* **25**, 621–627.

Broughton, H. B. (2000). A method for including protein flexibility in protein–ligand docking: improving tools for database mining and virtual screening. *Journal of Molecular Graphics and Modelling* **18**, 247–257, 302–244.

Buhrman, G., de Serrano, V. & Mattos, C. (2003). Organic solvents order the dynamic switch II in Ras crystals. *Structure* **11**, 747–751.

Carlson, H. A., Masukawa, K. M. & McCammon, J. A. (1999). Method for including the dynamic fluctuations of a protein in computer-aided drug design. *Journal of Physical Chemistry A* **103**, 10213–10219.

Carlson, H. A., Masukawa, K. M., Rubins, K., Bushman, F. D., Jorgensen, W. L., Lins, R. D., Briggs, J. M. & McCammon, J. A. (2000). Developing a dynamic pharmacophore model for HIV-1 integrase. *Journal of Medicinal Chemistry* **43**, 2100–2114.

Cavasotto, C. N. & Abagyan, R. A. (2004). Protein flexibility in ligand docking and virtual screening to protein kinases. *Journal of Molecular Biology* **337**, 209–225.

Cavasotto, C. N., Kovacs, J. A. & Abagyan, R. A. (2005). Representing receptor flexibility in ligand docking through relevant normal modes. *Journal of the American Chemical Society* **127**, 9632–9640.

Cheng, L. S., Amaro, R. E., Xu, D., Li, W. W., Arzberger, P. W. & McCammon, J. A. (2008). Ensemble-based virtual screening reveals potential

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

novel antiviral compounds for avian influenza neuraminidase. *Journal of Medicinal Chemistry* **51**, 3878–3894.

Chopra, G., Summa, C. M. & Levitt, M. (2008). Solvent dramatically affects protein structure refinement. *Proceedings of the National Academy of Sciences USA* **105**, 20239–20244.

Clark, M., Guarnieri, F., Shkurko, I. & Wiseman, J. (2006). Grand canonical Monte Carlo simulation of ligand-protein binding. *Journal of Chemical Information and Modeling* **46**, 231–242.

Clark, M., Meshkat, S. & Wiseman, J. S. (2009). Grand canonical free-energy calculations of protein-ligand binding. *Journal of Chemical Information and Modeling* **49**, 934–943.

Claussen, H., Buning, C., Rarey, M. & Lengauer, T. (2001). FlexE: efficient molecular docking considering protein structure variations. *Journal of Molecular Biology* **308**, 377–395.

Corbeil, C. R., Englebienne, P. & Moitessier, N. (2007). Docking ligands into flexible and solvated macromolecules. 1. Development and validation of FITTED 1.0. *Journal of Chemical Information and Modeling* **47**, 435–449.

Corbeil, C. R., Englebienne, P., Yannopoulos, C. G., Chan, L., Das, S. K., Bilimoria, D., L'Heureux, L. & Moitessier, N. (2008). Docking ligands into flexible and solvated macromolecules. 2. Development and application of FITTED 1.5 to the virtual screening of potential HCV polymerase inhibitors. *Journal of Chemical Information and Modeling* **48**, 902–909.

Corbeil, C. R. & Moitessier, N. (2009). Docking ligands into flexible and solvated macromolecules. 3. Impact of input ligand conformation, protein flexibility, and water molecules on the accuracy of docking programs. *Journal of Chemical Information and Modeling* **49**, 997–1009.

Cozzini, P., Kellogg, G. E., Spyrakis, 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. *Journal of Medicinal Chemistry* **51**, 6237–6255.

Craig, I. R., Essex, J. W. & Spiegel, K. (2010). Ensemble docking into multiple crystallographically derived protein structures: an evaluation based on the statistical analysis of enrichments. *Journal of Chemical Information and Modeling* **50**, 511–524.

Damm, K. L. & Carlson, H. A. (2007). Exploring experimental sources of multiple protein conformations in structure-based drug design. *Journal of the American Chemical Society* **129**, 8225–8235.

Damm, K. L., Ung, P. M., Quintero, J. J., Gestwicki, J. E. & Carlson, H. A. (2008). A poke in the eye: inhibiting HIV-1 protease through its flap-recognition pocket. *Biopolymers* **89**, 643–652.

Davis, I. W. & Baker, D. (2009). ROSETTALIGAND docking with full ligand and receptor flexibility. *Journal of Molecular Biology* **385**, 381–392.

Davis, I. W., Raha, K., Head, M. S. & Baker, D. (2009). Blind docking of pharmaceutically relevant compounds using ROSETTALIGAND. *Protein Science* **19**, 1998–2002.

Dechene, M., Wink, G., Smith, M., Swartz, P. & Mattos, C. (2009). Multiple solvent crystal structures of ribonuclease A: an assessment of the method. *Proteins* **76**, 861–881.

Dunbar, J. B., Smith, R. D., Yang, C.-Y., Ung, P. M. U., Lexa, K. W., Khazanov, N. A., Stuckey, J. A., Wang, S. & Carlson, H. A. (2011). CSAR benchmark exercise of 2010: selection of the protein-ligand complexes. *Journal of Chemical Information and Modeling* **51**, 2036–2046.

Dunbrack, R. L., Jr. & Karplus, M. (1993). Backbone-dependent rotamer library for proteins. Application to side-chain prediction. *Journal of Molecular Biology* **230**, 543–574.

Durrant, J. D., Keranen, H., Wilson, B. A. & McCammon, J. A. (2010). Computational identification of uncharacterized cruzain binding sites. *PLoS Neglected Tropical Diseases* **4**, e676.

Durrant, J. D. & McCammon, J. A. (2010). Computer-aided drug-discovery techniques that account for receptor flexibility. *Current Opinion in Pharmacology* **10**, 770–774.

Englebienne, P. & Moitessier, N. (2009). Docking ligands into flexible and solvated macromolecules. 4. Are popular scoring functions accurate for this class of proteins? *Journal of Chemical Information and Modeling* **49**, 1568–1580.

English, A. C., Done, S. H., Caves, L. S. D., Groom, C. R. & Hubbard, R. E. (1999). Locating interaction sites on proteins: the crystal structure of thermolysin soaked in 2% to 100% is opropanol. *Proteins-Structure Function and Genetics* **37**, 628–640.

English, A. C., Groom, C. R. & Hubbard, R. E. (2001). Experimental and computational mapping of the binding surface of a crystalline protein. *Protein Engineering* **14**, 47–59.

Erlanson, D. A., McDowell, R. S. & O'Brien, T. (2004). Fragment-based drug discovery. *Journal of Medicinal Chemistry* **47**, 3463–3482.

Eyal, E., Gerzon, S., Potapov, V., Edelman, M. & Sobolev, V. (2005). The limit of accuracy of protein modeling: influence of crystal packing on protein structure. *Journal of Molecular Biology* **351**, 431–442.

Fedorov, A. A., Joseph-McCarthy, D., Fedorov, E., Sirakova, D., Graf, I. & Almo, S. C. (1996). Ionic interactions in crystalline bovine pancreatic ribonuclease A. *Biochemistry* **35**, 15962–15979.

Ferrari, A. M., Wei, B. Q., Costantino, L. & Shoichet, B. K. (2004). Soft docking and multiple receptor

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

conformations in virtual screening. *Journal of Medicinal Chemistry* **47**, 5076–5084.

Firth-Clark, S., Willems, H. M., Williams, A. & Harris, W. (2006). Generation and selection of novel estrogen receptor ligands using the de novo structure-based design tool, SkelGen. *Journal of Chemical Information and Modeling* **46**, 642–647.

Fischer, E. (1890). Synthese des traubenzuckers. *Berichte der deutschen chemischen Gesellschaft* **23**, 799–805.

Fitzpatrick, P. A., Steinmetz, A. C., Ringe, D. & Klibanov, A. M. (1993). Enzyme crystal structure in a neat organic solvent. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 8653–8657.

Foote, J. & Milstein, C. (1994). Conformational isomerism and the diversity of antibodies. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 10370–10374.

Frembgen-Kesner, T. & Elcock, A. H. (2006). Computational sampling of a cryptic drug binding site in a protein receptor: explicit solvent molecular dynamics and inhibitor docking to p38 MAP kinase. *Journal of Molecular Biology* **359**, 202–214.

Fuentes, G., Dastidar, S. G., Madhumalar, A. & Verma, C. S. (2011). Role of protein flexibility in the discovery of new drugs. *Drug Development Research* **72**, 26–35.

Gunasekaran, K., Ma, B. & Nussinov, R. (2004). Is allostery an intrinsic property of all dynamic proteins? *Proteins* **57**, 433–443.

Guvench, O. & MacKerell, A. D., Jr. (2009). Computational fragment-based binding site identification by ligand competitive saturation. *PLoS Computational Biology* **5**, e1000435.

Hammes, G. G., Chang, Y. C. & Oas, T. G. (2009). Conformational selection or induced fit: a flux description of reaction mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13737–13741.

Hartshorn, M. J., Verdonk, M. L., Chessari, G., Brewerton, S. C., Mooij, W. T., Mortenson, P. N. & Murray, C. W. (2007). Diverse, high-quality test set for the validation of protein–ligand docking performance. *Journal of Medicinal Chemistry* **50**, 726–741.

Henzler, A. M. & Rarey, M. (2010). In pursuit of fully flexible protein–ligand docking: modeling the bilateral mechanism of binding. *Molecular Informatics* **29**, 164–173.

Ho, W. C., Luo, C., Zhao, K. H., Chai, X. M., Fitzgerald, M. X. & Marmorstein, R. (2006). High-resolution structure of the p53 core domain: implications for binding small-molecule stabilizing compounds. *Acta Crystallographica Section D Biological Crystallography* **62**, 1484–1493.

Howard, S., Berdini, V., Boulstridge, J. A., Carr, M. G., Cross, D. M., Curry, J., Devine, L. A., Early, T. R., Fazal, L., Gill, A. L., Heathcote, M., Maman, S., Matthews, J. E., McMenamin, R. L., Navarro, E. F., O'Brien, M. A., O'Reilly, M., Rees, D. C., Reule, M., Tisi, D., Williams, G., Vinkovic, M. & Wyatt, P. G. (2009). Fragment-based discovery of the pyrazol-4-yl urea (AT9283), a multitargeted kinase inhibitor with potent aurora kinase activity. *Journal of Medicinal Chemistry* **52**, 379–388.

Huang, S. Y., Grinter, S. Z. & Zou, X. (2010). Scoring functions and their evaluation methods for protein–ligand docking: recent advances and future directions. *Physical Chemistry Chemical Physics* **12**, 12899–12908.

Huang, Z. & Wong, C. F. (2009). Docking flexible peptide to flexible protein by molecular dynamics using two implicit-solvent models: an evaluation in protein kinase and phosphatase systems. *Journal of Physical Chemistry B* **113**, 14343–14354.

Huang, S. Y. & Zou, X. (2007). Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. *Proteins* **66**, 399–421.

Jain, A. N. (2003). Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *Journal of Medicinal Chemistry* **46**, 499–511.

Jain, A. N. (2007). Surflex-Dock 2.1: robust performance from ligand energetic modeling, ring flexibility, and knowledge-based search. *Journal of Computer-Aided Molecular Design* **21**, 281–306.

Jain, A. N. (2009). Effects of protein conformation in docking: improved pose prediction through protein pocket adaptation. *Journal of Computer-Aided Molecular Design* **23**, 355–374.

Jiang, F. & Kim, S. H. (1991). "Soft docking": matching of molecular surface cubes. *Journal of Molecular Biology* **219**, 79–102.

Jorgensen, W. L. (2004). The many roles of computation in drug discovery. *Science* **303**, 1813–1818.

Joseph-McCarthy, D., Fedorov, A. A. & Almo, S. C. (1996). Comparison of experimental and computational functional group mapping of an RNase A structure: implications for computer-aided drug design. *Protein Engineering* **9**, 773–780.

Kairys, V. & Gilson, M. K. (2002). Enhanced docking with the mining minima optimizer: acceleration and side-chain flexibility. *Journal of Computational Chemistry* **23**, 1656–1670.

Kallblad, P. & Dean, P. M. (2003). Efficient conformational sampling of local side-chain flexibility. *Journal of Molecular Biology* **326**, 1651–1665.

Kastenholz, M. A., Pastor, M., Cruciani, G., Haaksma, E. E. & Fox, T. (2000). GRID/CPCA: a new computational tool to design selective ligands. *Journal of Medicinal Chemistry* **43**, 3033–3044.

Keseru, G. M. & Kolossvary, I. (2001). Fully flexible low-mode docking: application to induced fit in HIV integrase. *Journal of the American Chemical Society* **123**, 12708–12709.

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Klebe, G. (2006). Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery Today* **11**, 580–594.

Knegtel, R. M., Kuntz, I. D. & Oshiro, C. M. (1997). Molecular docking to ensembles of protein structures. *Journal of Molecular Biology* **266**, 424–440.

Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A. & Cheatham, T. E. (2000). Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of Chemical Research* **33**, 889–897.

Koshland, D. E. (1958). Application of a theory of enzyme specificity to protein synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **44**, 98–104.

Kranjc, A., Bongarzone, S., Rossetti, G., Biarnes, X., Cavalli, A., Bolognesi, M. L., Roberti, M., Legname, G. & Carloni, P. (2009). Docking ligands on protein surfaces: the case study of prion protein. *Journal of Chemical Theory and Computation* **5**, 2565–2573.

Kumar, S., Ma, B., Tsai, C. J., Sinha, N. & Nussinov, R. (2000). Folding and binding cascades: dynamic landscapes and population shifts. *Protein Science* **9**, 10–19.

Landon, M. R., Amaro, R. E., Baron, R., Ngan, C. H., Ozonoff, D., McCammon, J. A. & Vajda, S. (2008). Novel druggable hot spots in avian influenza neuraminidase H5N1 revealed by computational solvent mapping of a reduced and representative receptor ensemble. *Chemical Biology and Drug Design* **71**, 106–116.

Laughton, C. A., Orozco, M. & Vranken, W. (2009). COCO: a simple tool to enrich the representation of conformational variability in NMR structures. *Proteins* **75**, 206–216.

Leach, A. R. (1994). Ligand docking to proteins with discrete side-chain flexibility. *Journal of Molecular Biology* **235**, 345–356.

Leach, A. R., Shoichet, B. K. & Peishoff, C. E. (2006). Prediction of protein-ligand interactions. Docking and scoring: successes and gaps. *Journal of Medicinal Chemistry* **49**, 5851–5855.

Leong, M. K. (2007). A novel approach using pharmacophore ensemble/support vector machine (phe/svm) for prediction of hERG liability. *Chemical Research in Toxicology* **20**, 217–226.

Leong, M. K. & Chen, T. H. (2008). Prediction of cytochrome P450 2B6-substrate interactions using pharmacophore ensemble/support vector machine (PhE/SVM) approach. *Medicinal Chemistry* **4**, 396–406.

Leong, M. K., Chen, Y. M., Chen, H. B. & Chen, P. H. (2009). Development of a new predictive model for interactions with human cytochrome P450 2A6 using pharmacophore ensemble/support vector machine (PhE/SVM) approach. *Pharmaceutical Research* **26**, 987–1000.

Lexa, K. W. & Carlson, H. A. (2011). Full protein flexibility is essential for proper hot-spot mapping. *Journal of the American Chemical Society* **133**, 200–202.

Lin, J. H., Perryman, A. L., Schames, J. R. & McCammon, J. A. (2002). Computational drug design accommodating receptor flexibility: the relaxed complex scheme. *Journal of the American Chemical Society* **124**, 5632–5633.

Lin, J. H., Perryman, A. L., Schames, J. R. & McCammon, J. A. (2003). The relaxed complex method: accommodating receptor flexibility for drug design with an improved scoring scheme. *Biopolymers* **68**, 47–62.

Lindahl, E. & Delarue, M. (2005). Refinement of docked protein–ligand and protein–DNA structures using low frequency normal mode amplitude optimization. *Nucleic Acids Research* **33**, 4496–4506.

Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F. & Sali, A. (2000). Comparative protein structure modeling of genes and genomes. *Annual Reviews of Biophysics and Biomolecular Structure* **29**, 291–325.

Mattos, C. & Ringe, D. (1996). Locating and characterizing binding sites on proteins. *Nature Biotechnology* **14**, 595–599.

May, A., Sieker, F. & Zacharias, M. (2008). How to efficiently include receptor flexibility during computational docking. *Current Computer-Aided Drug Design* **4**, 143–153.

May, A. & Zacharias, M. (2005). Accounting for global protein deformability during protein–protein and protein–ligand docking. *Biochimica and Biophysica Acta* **1754**, 225–231.

May, A. & Zacharias, M. (2008). Protein–ligand docking accounting for receptor side chain and global flexibility in normal modes: evaluation on kinase inhibitor cross docking. *Journal of Medicinal Chemistry* **51**, 3499–3506.

McCammon, J. A. (2005). Target flexibility in molecular recognition. *Biochimica and Biophysica Acta* **1754**, 221–224.

Meiler, J. & Baker, D. (2006). ROSETTALIGAND: protein-small molecule docking with full side-chain flexibility. *Proteins* **65**, 538–548.

Miranker, A. & Karplus, M. (1991). Functionality maps of binding sites: a multiple copy simultaneous search method. *Proteins* **11**, 29–34.

Mizutani, M. Y., Takamatsu, Y., Ichinose, T., Nakamura, K. & Itai, A. (2006). Effective handling of induced-fit motion in flexible docking. *Proteins* **63**, 878–891.

Mobley, D. L. & Dill, K. A. (2009). Binding of small-molecule ligands to proteins: "What you see" is not always "what you get". *Structure* **17**, 489–498.

Moitessier, N., Englebienne, P., Lee, D., Lawandi, J. & Corbeil, C. R. (2008). Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. *British Journal of Pharmacology* **153** (Suppl. 1), S7–26.

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Monod, J., Wyman, J. & Changeux, J. P. (1965). On the nature of allosteric transitions: a plausible model. *Journal of Molecular Biology* **12**, 88–118.

Moore, W. R., Jr. (2005). Maximizing discovery efficiency with a computationally driven fragment approach. *Current Opinion in Drug Discovery and Development* **8**, 355–364.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *Journal of Computational Chemistry* **30**, 2785–2781.

Mukherjee, S., Balius, T. E. & Rizzo, R. C. (2010). Docking validation resources: protein family and ligand flexibility experiments. *Journal of Chemical Information and Modeling* **50**, 1986–2000.

Murray, C. W., Baxter, C. A. & Frenkel, A. D. (1999). The sensitivity of the results of molecular docking to induced fit effects: application to thrombin, thermolysin and neuraminidase. *Journal of Computer-Aided Molecular Design* **13**, 547–562.

Murray, C. W. & Rees, D. C. (2009). The rise of fragment-based drug discovery. *Nature Chemistry* **1**, 187–192.

Mustata, G. I. & Briggs, J. M. (2002). A structure-based design approach for the identification of novel inhibitors: application to an alanine racemase. *Journal of Computer-Aided Molecular Design* **16**, 935–953.

Mustata, G. I., Soares, T. A. & Briggs, J. M. (2003). Molecular dynamics studies of alanine racemase: a structural model for drug design. *Biopolymers* **70**, 186–200.

Nabuurs, S. B., Wagener, M. & De Vlieg, J. (2007). A flexible approach to induce fit docking. *Journal of Medicinal Chemistry* **50**, 6507–6518.

Najmanovich, R., Kuttner, J., Sobolev, V. & Edelman, M. (2000). Side-chain flexibility in proteins upon ligand binding. *Proteins* **39**, 261–268.

Nichols, S. E., Baron, R., Ivetac, A. & McCammon, J. A. (2011). Predictive power of molecular dynamics receptor structures in virtual screening. *Journal of Chemical Information and Modeling* **51**, 1439–1446.

Osterberg, F., Morris, G. M., Sanner, M. F., Olson, A. J. & Goodsell, D. S. (2002). Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins* **46**, 34–40.

Ota, N. & Agard, D. A. (2001). Binding mode prediction for a flexible ligand in a flexible pocket using multiconformation simulated annealing pseudo crystallographic refinement. *Journal of Molecular Biology* **314**, 607–617.

Pencheva, T., Lagorce, D., Pajeva, I., Villoutreix, B. O. & Miteva, M. A. (2008). AMMOS: automated molecular mechanics optimization tool for *in silico* screening. *BMC Bioinformatics* **9**, 438.

Perryman, A. L., Lin, J. H. & McCammon, J. A. (2006). Optimization and computational evaluation of a series of potential active site inhibitors of the V82F/I84V drug-resistant mutant of HIV-1 protease: an application of the relaxed complex method of structure-based drug design. *Chemical Biology and Drug Design* **67**, 336–345.

Philippopoulos, M. & Lim, C. (1999). Exploring the dynamic information content of a protein NMR structure: comparison of a molecular dynamics simulation with the NMR and X-ray structures of *Escherichia coli* ribonuclease HI. *Proteins* **36**, 87–110.

Plewczynski, D., Lazniewski, M., Augustyniak, R. & Ginalski, K. (2011). Can we trust docking results? Evaluation of seven commonly used programs on PDBbind database. *Journal of Computational Chemistry* **32**, 742–755.

Popov, V. M., Yee, W. A. & Anderson, A. C. (2007). Towards *in silico* lead optimization: scores from ensembles of protein/ligand conformations reliably correlate with biological activity. *Proteins* **66**, 375–387.

Raman, E. P., Yu, W., Guvench, O. & Mackerell, A. D. (2011). Reproducing crystal binding modes of ligand functional groups using site-identification by ligand competitive saturation (SILCS) simulations. *Journal of Chemical Information and Modeling* **51**, 877–896.

Rao, C. B., Subramanian, J. & Sharma, S. D. (2009). Managing protein flexibility in docking and its applications. *Drug Discovery Today* **14**, 394–400.

Rarey, M., Kramer, B., Lengauer, T. & Klebe, G. (1996). A fast flexible docking method using an incremental construction algorithm. *Journal of Molecular Biology* **261**, 470–489.

Rueda, M., Bottegoni, G. & Abagyan, R. (2009). Consistent improvement of cross-docking results using binding site ensembles generated with elastic network normal modes. *Journal of Chemical Information and Modeling* **49**, 716–725.

Rueda, M., Bottegoni, G. & Abagyan, R. (2010). Recipes for the selection of experimental protein conformations for virtual screening. *Journal of Chemical Information and Modeling* **50**, 186–193.

Schafferhans, A. & Klebe, G. (2001). Docking ligands onto binding site representations derived from proteins built by homology modelling. *Journal of Molecular Biology* **307**, 407–427.

Schmitke, J. L., Stern, L. J. & Klibanov, A. M. (1997). The crystal structure of subtilisin Carlsberg in anhydrous dioxane and its comparison with those in water and acetonitrile. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 4250–4255.

Schmitke, J. L., Stern, L. J. & Klibanov, A. M. (1998). Comparison of x-ray crystal structures of an acylenzyme intermediate of subtilisin Carlsberg formed in anhydrous acetonitrile and in water. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 12918–12923.

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Schnecke, V. & Kuhn, L. A. (2000). Virtual screening with solvation and ligand-induced complementarity. *Perspectives in Drug Discovery and Design* **20**, 171–190.

Schubert, C. R. & Stultz, C. M. (2009). The multi-copy simultaneous search methodology: a fundamental tool for structure-based drug design. *Journal of Computer-Aided Molecular Design* **23**, 475–489.

Seco, J., Luque, F. J. & Barril, X. (2009). Binding site detection and druggability index from first principles. *Journal of Medicinal Chemistry* **52**, 2363–2371.

Sherman, W., Beard, H. S. & Farid, R. (2006a). Use of an induced fit receptor structure in virtual screening. *Chemical Biology and Drug Design* **67**, 83–84.

Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A. & Farid, R. (2006b). Novel procedure for modeling ligand/receptor induced fit effects. *Journal of Medicinal Chemistry* **49**, 534–553.

Shuker, S. B., Hajduk, P. J., Meadows, R. P. & Fesik, S. W. (1996). Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **274**, 1531–1534.

Smith, L. J., Redfield, C., Smith, R. A., Dobson, C. M., Clore, G. M., Gronenborn, A. M., Walter, M. R., Naganbushan, T. L. & Wlodawer, A. (1994). Comparison of four independently determined structures of human recombinant interleukin-4. *Nature Structural Biology* **1**, 301–310.

Smith, R. D., Dunbar, J. B., Jr., Ung, P. M. U., Esposito, E. X., Yang, C.-Y., Wang, S. & Carlson, H. A. (2011). CSAR benchmark exercise of 2010: combined evaluation across all submitted scoring functions. *Journal of Chemical Information and Modeling* **51**(9), 2036–2046.

Sotriffer, C. A. (2011). Accounting for induced-fit effects in docking: what is possible and what is not? *Current Topics in Medicinal Chemistry* **11**, 179–191.

Sousa, S. F., Fernandes, P. A. & Ramos, M. J. (2006). Protein–ligand docking: current status and future challenges. *Proteins* **65**, 15–26.

Sperandio, O., Mouawad, L., Pinto, E., Villoutreix, B. O., Perahia, D. & Miteva, M. A. (2010). How to choose relevant multiple receptor conformations for virtual screening: a test case of Cdk2 and normal mode analysis. *European Biophysics Journal* **39**, 1365–1372.

Stultz, C. M. & Karplus, M. (1999). MCSS functionality maps for a flexible protein. *Proteins* **37**, 512–529.

Subramanian, J., Sharma, S. & B-Rao, C. (2006). A novel computational analysis of ligand-induced conformational changes in the ATP binding sites of cyclin dependent kinases. *Journal of Medicinal Chemistry* **49**, 5434–5441.

Subramanian, J., Sharma, S. & B-Rao, C. (2008). Modeling and selection of flexible proteins for structure-based drug design: backbone and side chain movements in p38 MAPK. *ChemMedChem* **3**, 336–344.

Sullivan, S. M. & Holyoak, T. (2008). Enzymes with lid-gated active sites must operate by an induced fit mechanism instead of conformational selection. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 13829–13834.

Sweeney, Z. K., Harris, S. F., Arora, N., Javanbakht, H., Li, Y., Fretland, J., Davidson, J. P., Billedeau, J. R., Gleason, S. K., Hirschfeld, D., Kennedy-Smith, J. J., Mirzadegan, T., Roetz, R., Smith, M., Sperry, S., Suh, J. M., Wu, J., Tsing, S., Villaseñor, A. G., Paul, A., Su, G., Heilek, G., Hang, J. Q., Zhou, A. S., Jernelius, J. A., Zhang, F-J. & Klumpp, K. (2008). Design of annulated pyrazoles as inhibitors of HIV-1 reverse transcriptase. *Journal of Medicinal Chemistry* **51**, 7449–7458.

Taft, C. A., Da Silva, V. B. & Da Silva, C. H. (2008). Current topics in computer-aided drug design. *Journal of Pharmaceutical Sciences* **97**, 1089–1098.

Talele, T. T., Khedkar, S. A. & Rigby, A. C. (2010). Successful applications of computer aided drug discovery: moving drugs from concept to the clinic. *Current Topics in Medicinal Chemistry* **10**, 127–141.

Taylor, R. D., Jewsbury, P. J. & Essex, J. W. (2003). FDS: flexible ligand and receptor docking with a continuum solvent model and soft-core energy function. *Journal of Computational Chemistry* **24**, 1637–1656.

Teague, S. J. (2003). Implications of protein flexibility for drug discovery. *Nature Reviews Drug Discovery* **2**, 527–541.

Teodoro, M. L. & Kavraki, L. E. (2003). Conformational flexibility models for the receptor in structure based drug design. *Current Pharmaceutical Design* **9**, 1635–1648.

Totrov, M. & Abagyan, R. (2008). Flexible ligand docking to multiple receptor conformations: a practical alternative. *Current Opinions in Structural Biology* **18**, 178–184.

Tsai, C. J., Kumar, S., Ma, B. & Nussinov, R. (1999a). Folding funnels, binding funnels, and protein function. *Protein Science* **8**, 1181–1190.

Tsai, C. J., Ma, B. & Nussinov, R. (1999b). Folding and binding cascades: shifts in energy landscapes. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 9970–9972.

Tsai, C. J., Ma, B., Sham, Y. Y., Kumar, S. & Nussinov, R. (2001). Structured disorder and conformational selection. *Proteins* **44**, 418–427.

van Westen, G. J. P., Wegner, J. K., Bender, A., Ijzerman, A. P. & van Vlijmen, H. W. T. (2010). Mining protein dynamics from sets of crystal structures using "consensus structures". *Protein Science* **19**, 742–752.

Velec, H. F., Gohlke, H. & Klebe, G. (2005). DrugScore(CSD)-knowledge-based scoring function derived from small molecule crystal data with superior recognition rate of near-native ligand poses and better affinity prediction. *Journal of Medicinal Chemistry* **48**, 6296–6303.

Verdonk, M. L., Mortenson, P. N., Hall, R. J., Hartshorn, M. J. & Murray, C. W. (2008). Protein-ligand docking against non-native protein conformers. *Journal of Chemical Information and Modeling* **48**, 2214–2225.

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press

Wang, Z., Zhu, G., Huang, Q., Qian, M., Shao, M., Jia, Y. & Tang, Y. (1998). X-ray studies on cross-linked lysozyme crystals in acetonitrile–water mixture. *Biochimica and Biophysica Acta* **1384**, 335–344.

Warren, G. L., Andrews, C. W., Capelli, A. M., Clarke, B., LaLonde, J., Lambert, M. H., Lindvall, M., Nevins, N., Semus, S. F., Senger, S., Tedesco, G., Wall, I. D., Woolven, J. M., Peishoff, C. E. & Head, M. S. (2006). A critical assessment of docking programs and scoring functions. *Journal of Medicinal Chemistry* **49**, 5912–5931.

Wei, B. Q., Weaver, L. H., Ferrari, A. M., Matthews, B. W. & Shoichet, B. K. (2004). Testing a flexible-receptor docking algorithm in a model binding site. *Journal of Molecular Biology* **337**, 1161–1182.

Weiki, T. R. & von Deuster, C. (2009). Selected-fit versus induced-fit protein binding: kinetic differences and mutational analysis. *Proteins* **75**, 104–110.

Wiesmann, C., Barr, K. J., Kung, J., Zhu, J., Erlanson, D. A., Shen, W., Fahr, B. J., Zhong, M., Taylor, L., Randal, M., McDowell, R. S. & Hansen, S. K. (2004). Allosteric inhibition of protein tyrosine phosphatase 1B. *Nature Structural and Molecular Biology* **11**, 730–737.

Wong, S. & Jacobson, M. P. (2008). Conformational selection *in silico*: loop latching motions and ligand binding in enzymes. *Proteins* **71**, 153–164.

Yang, C. & Wang, S. (2010). Computational analysis of protein hotspots. *ACS Medicinal Chemistry Letters* **1**, 125–129.

Yang, L. W., Eyal, E., Chennubhotla, C., Jee, J., Gronenborn, A. M. & Bahar, I. (2007). Insights into equilibrium dynamics of proteins from comparison of NMR and X-ray data with computational predictions. *Structure* **15**, 741–749.

Zacharias, M. (2004). Rapid protein-ligand docking using soft modes from molecular dynamics simulations to account for protein deformability: binding of FK506 to FKBP. *Proteins: Structure, Function, and Bioinformatics* **54**, 759–767.

Zacharias, M. (2008). Combining elastic network analysis and molecular dynamics simulations by hamiltonian replica exchange. *Journal of Chemical Theory and Computation* **4**, 477–487.

Zacharias, M. & Sklenar, H. (1999). Harmonic modes as variables to approximately account for receptor flexibility in ligand-receptor docking simulations: application to DNA minor groove ligand complex. *Journal of Computational Chemistry* **20**, 287–300.

Zavodszky, M. I. & Kuhn, L. A. (2005). Side-chain flexibility in protein–ligand binding: the minimal rotation hypothesis. *Protein Science* **14**, 1104–1114.

Zavodszky, M. I., Lei, M., Thorpe, M. F., Day, A. R. & Kuhn, L. A. (2004). Modeling correlated main-chain motions in proteins for flexible molecular recognition. *Proteins: Structure, Function, and Bioinformatics* **57**, 243–261.

Zavodszky, M. I., Sanschagrin, P. C., Korde, R. S. & Kuhn, L. A. (2002). Distilling the essential features of a protein surface for improving protein–ligand docking, scoring, and virtual screening. *Journal of Computer-Aided Molecular Design* **16**, 883–902.

Zentgraf, M., Fokkens, J. & Sotriffer, C. A. (2006). Addressing protein flexibility and ligand selectivity by "in situ cross-docking". *ChemMedChem* **1**, 1355–1359.

Zentgraf, M., Steuber, H., Koch, C., La Motta, C., Sartini, S., Sotriffer, C. A. & Klebe, G. (2007). How reliable are current docking approaches for structure-based drug design? Lessons from aldose reductase. *Angewandte Chemie* **46**, 3575–3578.

Zhao, Y. & Sanner, M. F. (2007). FLIPDock: docking flexible ligands into flexible receptors. *Proteins: Structure, Function, and Bioinformatics* **68**, 726–737.

Zhao, Y. & Sanner, M. F. (2008). Protein-ligand docking with multiple flexible side chains. *Journal of Computer-Aided Molecular Design* **22**, 673–679.

Zhu, J., Fan, H., Liu, H. & Shi, Y. (2001). Structure-based ligand design for flexible proteins: application of new F-DycoBlock. *Journal of Computer-Aided Molecular Design* **15**, 979–996.

https://doi.org/10.1017/S0033583512000066 Published online by Cambridge University Press