

## Review

# Peptidic Modulators of Protein-Protein Interactions: Progress and Challenges in Computational Design

Mor Rubinstein, Masha Y. Niv

The Institute of Biochemistry, Food Science and Nutrition, The Hebrew University of Jerusalem, P.O. Box 12, Rehovot 76100, Israel

Received 23 December 2008; revised 6 February 2009; accepted 6 February 2009

Published online 18 February 2009 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bip.21164

### ABSTRACT:

*With the decline in productivity of drug-development efforts, novel approaches to rational drug design are being introduced and developed. Naturally occurring and synthetic peptides are emerging as novel promising compounds that can specifically and efficiently modulate signaling pathways in vitro and in vivo. We describe sequence-based approaches that use peptides to mimic proteins in order to inhibit the interaction of the mimicked protein with its partners. We then discuss a structure-based approach, in which protein-peptide complex structures are used to rationally design and optimize peptidic inhibitors. We survey flexible peptide docking techniques and discuss current challenges and future directions in the rational design of peptidic inhibitors. © 2009 Wiley Periodicals, Inc. Biopolymers 91: 505–513, 2009.*

**Keywords:** docking; peptides; peptidomimetics; protein motifs; computer-aided drug design

*This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of the preprint by emailing the Biopolymers editorial office at biopolymers@wiley.com*

Correspondence to: Masha Y. Niv; e-mail: niv@agri.huji.ac.il  
Contract grant sponsor: United States-Israel Binational Science Foundation  
Contract grant number: 2007296  
Contract grant sponsor: Niedersachsen-Israeli Research cooperation fund  
© 2009 Wiley Periodicals, Inc.

### INTRODUCTION

Efforts in drug discovery are routinely based on targeting active sites of proteins with small molecules. The characteristics of typical drug-like molecules were summarized in Lipinski's rule of five,<sup>1</sup> namely not more than five hydrogen-bond donors and 10 hydrogen-bond acceptors, a molecular weight under 500 Dalton, and a partition coefficient log P, which is a measure of lipophilicity, of less than five. These were later revised by Veber et al.<sup>2</sup> who proposed a more relaxed set of rules for predicting oral bioavailability: (1) 10 or fewer rotatable bonds and (2) polar surface area equal to or less than 140 Å<sup>2</sup> (or 12 or fewer hydrogen-bond donors and acceptors). However, in recent years, the productivity of drug-discovery efforts has shown a consistent decline.<sup>3</sup> New paradigms have begun to emerge in an attempt to improve the situation. One exciting direction is that of targeting protein-protein interactions as an alternative to targeting active sites of enzymes. Protein-protein interactions between an enzyme and its substrate, between a catalytic domain and its regulatory domains, and between protein motifs and scaffolding domains are involved in most signal-transduction processes in the cell. Stabilization of protein-protein interactions,<sup>4,5</sup> disruption of protein-protein interactions<sup>6</sup> and disruption of intra-protein contacts during folding<sup>7</sup> are all being pursued. However, despite some success stories, discovering small-molecule drugs that modulate protein-protein interactions remains an enormous challenge.<sup>8</sup> The large interfacial areas and the relatively flat topographies of the surfaces of protein-protein interfaces are different from the smaller and deeper cavities of the active sites.<sup>6</sup> Therefore, various types of molecules that are suitable for targeting protein-protein interactions, and that do not necessarily comply with Lipinski or Veber rules, are now actively being sought.

Recent advances in peptide and peptidomimetic synthesis<sup>9,10</sup> are generating renewed enthusiasm in the use of pep-

tides as therapeutics, for example, in vaccine development,<sup>11</sup> angiogenesis-related<sup>12</sup> and auto-immune diseases,<sup>13</sup> diabetes<sup>14</sup> and neuroprotection.<sup>15</sup> In parallel, the role of short linear motifs in mediating protein-protein interactions in signaling complexes<sup>16,17</sup> is being unraveled, and the importance of such motifs (and potentially of their peptidic mimics) in signal channeling has been demonstrated.<sup>17</sup> In general, within many of the highly abundant disordered regions in proteins, lie short (3–10 amino acids long) linear peptide sections that are responsible for thousands of protein-protein interactions. Some examples are the SH3 and WW domains which bind to proline-rich regions, the SH2, PTB and 14-3-3 domains which bind to phosphorylated peptides, the PDZ domain discussed below, and many others.<sup>17</sup> It has become clear that protein-peptide interactions are very abundant in the cell and estimates suggest they represent 15–40% of all interactions,<sup>18</sup> making peptide discovery and development of wide applicability and interest.

Cyclic peptides<sup>19–22</sup> and other modified peptides<sup>23,24</sup> can be highly potent as well as selective, providing a compromise between structural preorganization and sufficient flexibility to mold to a target surface and maximize binding interactions. There is much evidence that despite having molecular masses well above Lipinski's rule constraints, cyclic peptides and other macrocycles can possess drug-like physicochemical and pharmacokinetic properties such as good solubility, lipophilicity, metabolic stability and bioavailability.<sup>25</sup>

Targeting protein-protein interactions by large and flexible molecules requires the adaptation of research methods from several disciplines: suitable compounds need to be synthesized, appropriate biochemical assays that are able to monitor modulation of protein-protein interactions need to be developed, and computational tools that are adequate for analyses of protein interactions with such compounds need to be devised. This review focuses on computational methods for the rational design of peptidic modulators of protein-protein interactions. The fact that the peptides are short proteins allows the use of sequence-based techniques, which are not suitable in the case of small molecules; on the other hand, the large size and high flexibility of peptides present a challenge for the conventional computational docking approaches that rely on the relatively small number of rotatable bonds in the typical "drug-like" molecules. In this review, we highlight the sequence-based approaches to peptide development as well as the adaptation of docking techniques to large and flexible peptides. The approaches described below are schematically represented in Figure 1.

## SEQUENCE-BASED APPROACHES FOR DISCOVERY OF PEPTIDE LEADS

### Mimicking a Binding Partner

One popular strategy used to search for peptides that inhibit a protein-protein interaction is to mimic the sequence of one of the interacting proteins<sup>26</sup> (Figure 1B). Following are some examples of applying this strategy to several important drug targets.

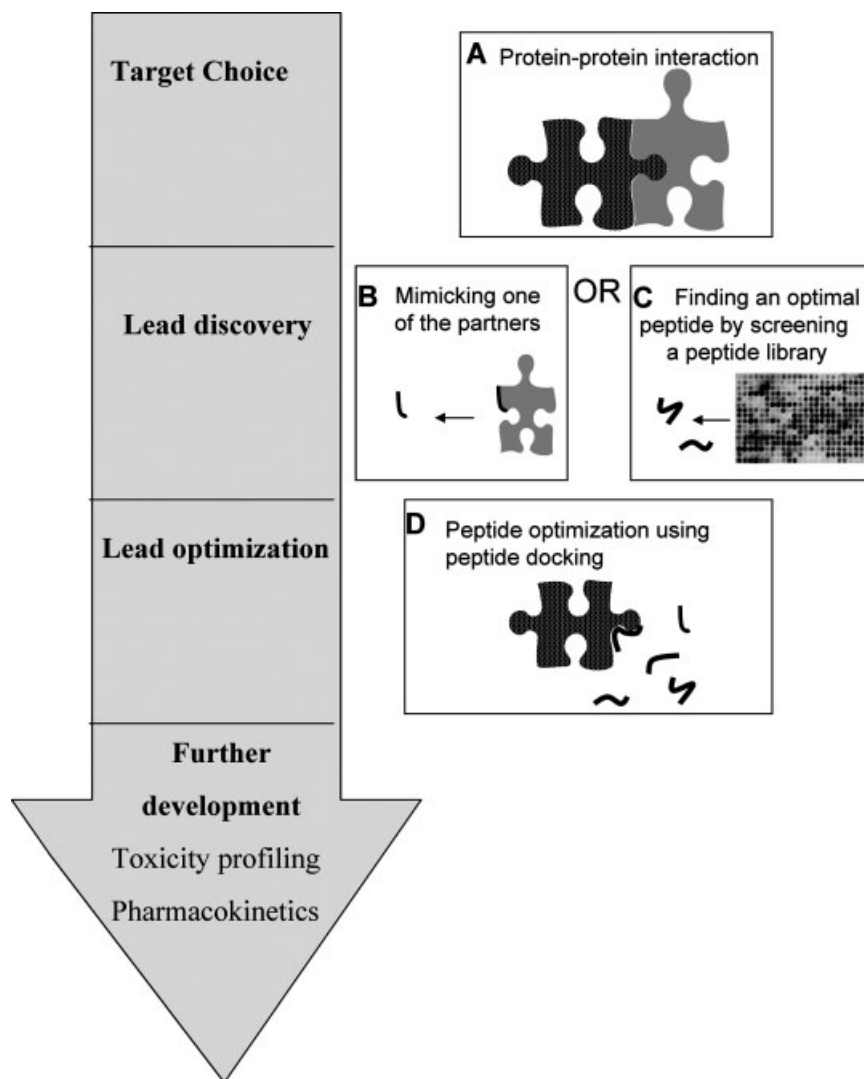
Protein kinases are enzymes that catalyze the phosphorylation of tyrosine or serine/threonine residues. Protein kinases are involved in a myriad of signaling pathways and the regulation of their catalytic activity is crucial for healthy functioning of the cell. Most protein kinase inhibitors in laboratory use, in the clinic and under clinical development are small molecules that interfere with ATP binding of their target kinase.<sup>27,28</sup>

Recently, approaches aimed at targeting the protein-protein interactions of kinases, rather than their ATP-binding site, have emerged. These include inhibiting the interactions of the catalytic domain with its regulatory domains (e.g. Refs. 29–31), inhibiting interactions of the regulatory SH2 domain with its ligands,<sup>32</sup> inhibiting kinase-substrate interactions<sup>22,33–35</sup> etc. The main approaches that use sequence data for selectively targeting interactions between kinases and other proteins are: mimicking a common region of unrelated binders of the same protein, mimicking the substrate, or mimicking the substrate-binding regions in the kinase.

Mochly-Rosen and coworkers derived a peptide based on the only homologous region between annexin I and a member of the 14-3-3 family, two unrelated binders of protein kinase C (PKC). The authors reasoned that the region of homology is what mediates PKC binding to these two proteins. Indeed a 15-mer peptide corresponding to this sequence inhibited PKC binding to the purified or cloned RACK, as well as to annexin I; the peptide inhibited  $\beta_{II}$ PKC function in cells and its injection into a *Xenopus* oocyte inhibited insulin-induced oocyte maturation.<sup>29,36,37</sup>

Inhibiting kinase-substrate interactions by mimicking the sequence of known substrate peptides is an approach that is being actively pursued by several labs.<sup>22,35,38–40</sup> For example, the sequence of the activation loop of insulin-like growth factor-1 receptor, which undergoes autophosphorylation, was used as a pharmacophore template to be mimicked by novel macrocyclic molecules.<sup>22</sup>

A complementary approach is to derive peptides based on the sequence of the enzyme rather than the substrate. We analyzed structures and sequences of protein kinases and identified well-defined substrate-binding regions consisting of accessible and variable patches distributed between conserved and buried anchors. The variable nature of these regions



**FIGURE 1** Schematic representation of the approaches described in this review, in the general context of drug (or modulator of signaling) development. (A) Protein-protein interaction cartoon (B) Mimicking the interface sequence of one of the partners by a peptide. (C) Finding inhibitory peptides by biochemical screening of peptide libraries. (D) Predicting the most probable bound conformation of the peptide in complex with the target protein (peptide docking). This approach can be applied to guide optimization of peptide sequences derived by A or B; or to simultaneously optimize the conformation and the sequence of the peptide.

suggests that they are involved in determining substrate specificity. We used these substrate-binding regions from several different kinases to derive peptides of 7–11 amino acids in length, all myristoylated at the N-terminus and amidated at the C-terminus. These peptides inhibited the activity of kinases from which they were derived in both cell-based and cell-free assays (whereas control peptides failed to do so), and inhibited cancer cell proliferation in the micromolar range.<sup>41</sup> Administration via injection into animal models has confirmed that the kinase-mimicking peptides were nontoxic and had the expected physiological effects.<sup>34,42</sup> Because the

peptides are derived from a well-defined region in the kinase sequence, they can easily be designed for a large number of kinases, including those that are novel or not characterized.

A similar idea was recently adopted for the inhibition of  $\beta$ -APP cleaving enzyme-1 (BACE1),<sup>43</sup> a primary therapeutic target in Alzheimer's disease.<sup>43,44</sup> BACE1 inhibitory peptides of between 9 and 12 residues in length, were derived from the catalytic domains of BACE1 and N-myristoylated for cell permeability. These molecules inhibited the formation of the BACE1 product APP $\beta$ , and significantly reduced the level of the BACE1 product A $\beta$ 40, a neurotoxic peptide whose accu-

mulation induces Alzheimer's disease. Control peptides did not show this inhibitory effect.<sup>43</sup>

gp41, a transmembrane protein which is part of the HIV-1 envelope glycoprotein (Env), mediates HIV-cell fusion. It contains four major functional domains: the fusion peptide, heptad repeat 1 (HR1), HR2, and the transmembrane domain that anchors gp41 to the viral lipid bilayer. The FDA-approved anti-HIV drug Enfuvirtide is a synthetically produced 36-amino-acid peptide derived from a continuous sequence within the HR2 domain of gp41. The fusion of host and viral membranes is a complex process, involving several conformational transitions. In one of those transitions, the HR2 domain "zips" into a hydrophobic groove on the outer face of a pre-existing HR1 helix bundle. Enfuvirtide inhibits this stage by competing with the HR2 domain for HR1 binding.<sup>45</sup>

G-protein coupled receptors (GPCRs) are seven-transmembrane proteins that function as signaling switches throughout the body and regulate virtually every physiological response.<sup>46,47</sup> This is one of the largest gene families targeted for drug discovery.<sup>47</sup> Peptides (usually around 20 amino acids in length) that mimic GPCR-transmembrane domains<sup>48–50</sup> are thought to inhibit GPCR oligomerization. 10–20 amino acids long peptides mimicking the intracellular loops of GPCRs<sup>51–53</sup> were designed to modulate interactions with downstream proteins. Shorter peptides and peptidomimetics mimicking the endogenous peptidic ligands of some GPCRs are under active development.<sup>21,54</sup>

Integrins are a family of heterodimeric cell-surface receptors that regulate cell attachment and response to the extracellular matrix. Many integrins recognize a sequence in the matrix protein that contains the core amino-acid triplet, "Arg-Gly-Asp" (RGD). Cilengitide is a cyclic ligand-mimicking Arg-Gly-Asp peptide, which binds to and inhibits the activities of the  $\alpha v \beta 3$  and  $\alpha v \beta 5$  integrins, thereby inhibiting endothelial cell-cell interactions, endothelial cell-matrix interactions, and angiogenesis.<sup>55</sup> The cyclization of the peptide increases its stability to serum degradation and introduces stereo constraints, that make it selective with respect to other integrins. The compound is currently in clinical trials for the treatment of glioblastoma.

### Finding an Optimal Synthetic Peptide by a Wide-Scale Screen

Recent studies have used wide-scale screens of synthetic or phage-display peptides for binding to a protein of interest (schematically represented in Figure 1C). Combinatorial approaches to peptides as protein-site mimetics have been discussed in a recent review.<sup>56</sup> Such screens of either random<sup>57–61</sup> or biased peptide libraries<sup>22,38,62</sup> of peptides of up

to 20 amino acids in length,<sup>63</sup> may identify potent peptidic inhibitors that are not necessarily similar to native ligands of the protein, but nevertheless efficiently inhibit the protein-protein interaction of interest.

Lead inhibitory peptides, whether based on a known binding partner or obtained from a library screen, often serve as merely the first step in developing an inhibitor. For example, rational changes may be aimed at enhancing the stability of the active conformation, affinity, selectivity and other properties of the peptide. Towards this goal, detailed structural information about the interaction between the peptide and the target protein is most useful. However, such information may be difficult to acquire, both experimentally and computationally. In the following, we describe the challenges and advances in the development of computational tools for structure-based peptide design.

### STRUCTURE-BASED APPROACHES FOR PEPTIDE DESIGN AND OPTIMIZATION

The key challenge for computational prediction of protein-peptide interactions is the flexibility of the systems under consideration: peptides have many more degrees of freedom than typical small molecules. Accounting for this flexibility is crucial for the rational design and optimization of peptides and peptidomimetics. In addition, some disruptors of protein-protein interactions aim to stabilize the inactive form of the protein. Thus, the rational design of such inhibitors must take into consideration the conformational plasticity of the protein and the interplay between different conformations. We focus on the methods used to predict the detailed interaction between the protein and the peptide, namely "peptide docking" (schematically represented in Figure 1D).

An excellent review on ligand docking to proteins can be found in Sousa et al.<sup>64</sup> The ability of docking methods to dock ligands into a known native structure has been evaluated in several recent papers.<sup>65–68</sup> Redocking (docking of the ligand back into its partner which is kept in the "bound" conformation) of ligands with up to 10 and sometimes even 20 rotatable bonds is currently achievable with several docking algorithms. However, docking of highly flexible ligands, such as peptides (which have four rotatable bonds per residue on average) remains nontrivial. There have been only a few comparative studies for such large ligands. In a comparative redocking study performed by Bursulaya et al., two ligands had about 30 rotatable bonds.<sup>67</sup> ICM<sup>69</sup> redocked both within 2 Å; one of these ligands was also successfully docked by Autodock,<sup>70</sup> while the other had a root mean square deviation (RMSD) of 9.10 Å. DOCK,<sup>71</sup> FlexX,<sup>72</sup> and GOLD<sup>73</sup> misdocked these highly flexible ligands (with RMSD >5.9 Å).<sup>67</sup> Another comparative test set<sup>74</sup> included



five ligands with more than 25 rotatable bonds: Surflex<sup>75</sup> redocked three of them with RMSD <2 Å, FLEX<sup>72</sup> and FRED<sup>76</sup> redocked two, GOLD<sup>73</sup> and GLIDE<sup>77</sup> redocked one, and DOCK<sup>71</sup> did not redock any of these flexible ligands within 2 Å.

A few procedures have been developed specifically for peptide docking. A heuristic search of the potential energy surface of peptides in the binding grooves of (rigid) proteins was tested on a dataset of 56 complexes.<sup>78</sup> While all 20 tripeptides were successfully redocked, only 3 of the 36 larger (tetra, penta, hexa, and hepta) peptides could be redocked within 2 Å, suggesting that the sampling strategy needs further adaptation for the longer peptides.

Modeling of the complexes between various nuclear receptors (NRs) and peptide coactivators was performed starting from the bound conformations of both partners and using a knowledge-based coarse-grained potential and replica exchange Monte Carlo sampling.<sup>79</sup> Only 1 out of 10 peptides redocked within 2 Å.

Limiting the conformational space using knowledge-based constraints is often crucial for successful peptide docking. The constraints may come from experiments (such as NOE data<sup>80</sup>) or from previous knowledge of conserved positions for some of the peptide residues.<sup>81–83</sup>

Scheraga and coworkers used a “ligand-growing” procedure. Optimization of the ECEPP/3 potential function<sup>84</sup> supplemented by the experimental NOE constraints was carried out using a collective variable Monte Carlo minimization technique.<sup>85</sup> This approach facilitates structure determination from NMR experiments of weakly binding systems. It was applied to determine the structure of a fibrinogen A  $\alpha$ -like peptide bound to an active-site mutant of thrombin.<sup>80</sup>

Schafroth and Floudas defined the pockets binding each of the peptide residues based on the known structure of the complex of a major histocompatibility complex (MHC) molecule HLA-DRB1\*0101 with a cognate peptide, and used deterministic global optimization to dock MHC peptides. The pockets were predefined and considered to be independent of each other and the protein was considered to be rigid.<sup>86</sup> Another work on MHC peptides made use of the fact that N-terminal and C-terminal parts of the peptides dock to well-conserved sites. The edges were docked as rigid bodies, and the rest of the peptide was placed using a loop-closure technique, achieving C $\alpha$  RMSD <1.5 Å for 9- to 10-amino-acid-long peptides (heavy atom RMSD was not reported).<sup>87</sup>

Well-defined anchoring points of peptides to proteins such as in MHC-peptide binding also occur in other cases: many linear motifs have a well-defined docking region for the conserved residues but more variable docking areas for the residues flanking the motif.<sup>88</sup> Two of the applications described below<sup>81,83</sup>

utilize the anchoring point idea and succeed not only in redocking, but also in the much more challenging task of docking peptides to other structures and to homology models.

Our protocol for flexible peptide docking,<sup>81</sup> implemented in CHARMM,<sup>89</sup> is based on a simulated annealing molecular dynamics (MD) approach. An anchoring point between the peptide and the protein must be known. Then, the peptide-protein complex is heated to an elevated temperature (the peptide does not “evaporate” because of the tethered anchor) and the conformational space in the binding region is explored. Each of the hundreds of conformations obtained from the heated trajectory is cooled, and then optimal side-chain rotamers are chosen using SCWRL3.0.<sup>90</sup> The resulting poses are minimized and scored by the sum of the interaction energies between the protein and the peptide, and the internal energy of the peptide. We tested the ability of this method to predict peptide binding to PDZ domains; the well-conserved position of the C $\alpha$  atom of the C-terminal residue of the peptide was used as an anchoring point. Redocking to native structures yielded excellent results (RMSD <2 Å) for the six tested penta- to octapeptides. In docking to other structures of the same protein (either apo, or originally complexed with a different peptide), and even to homology models, we found that in 9 out of 12 cases, the best-scoring poses had RMSD <2.8 Å for heavy atoms of tetra- to octapeptides.<sup>81</sup>

One of the motivations for studying the structures of PDZ-peptide complexes is the puzzling specificity profile of these interactions: one peptide may bind to dissimilar PDZ domains, while the same PDZ domain may bind dissimilar peptides.<sup>91,92</sup> Recent wide-scale peptide-library studies on PDZ specificity<sup>58,60</sup> have shed more light on the propensities of different PDZ domains towards peptides of particular sequences, but the ability of the same PDZ domain to bind peptides that have residues with different physical properties at a given position is still apparent (see Figure 3 in Tonikian et al.<sup>58</sup>). As suggested previously,<sup>92</sup> this variability probably reflects structural differences in the binding conformations of the peptides and/or their binding-site residues, which enable accommodation of the peptides in the binding groove. The interest in PDZ domains as drug targets, and the attempts to develop peptidomimetic or small-molecule PDZ inhibitors<sup>20,93</sup> require structural data to complement the vast available sequence data. Since determining the crystallographic structure of all peptide-PDZ complexes is unfeasible, docking is an essential procedure.

Bordner and Abagyan<sup>83</sup> tested and optimized peptide-MHC docking simulations using a flexible all-atom model of the complete peptide. A biased-probability Monte Carlo minimization method<sup>94</sup> implemented in the ICM program, combined with grid potentials to represent interactions with

the MHC, allowed a computationally efficient sampling of the peptide's degrees of freedom. Comparison of all available X-ray structures revealed conserved hydrogen bonds between the N and C termini of the peptide and particular MHC residues. Based on this observation, backbone atoms of the peptide's N and C termini were tethered to these conserved positions, by a quadratic energy restraint. This work was not limited to redocking: accurate prediction (0.75 Å backbone RMSD) for cross-docking of a highly flexible decapeptide, as well as docking predictions using homology models with low average backbone RMSDs of less than 1.0 Å were achieved.<sup>83</sup>

Sood and Baker aimed to recover not only the conformations but also the sequences of known binding peptides.<sup>82</sup> Starting from an anchor point, the authors extended the peptide backbone by using a library of about 500,000 peptide fragments of the appropriate length, which was precompiled from a nonredundant subset of the Protein Data Bank (PDB). Extensions that resulted in steric clashes between backbone atoms, or expanded such that there was little chance of favorable interactions with the target protein were eliminated. Side chains were then added using Rosetta-Design, and Monte Carlo search was used to sample all rotamers of all residues at the protein-peptide interface and identify the sequence with the lowest energy. In cases of little sequence variation of the native peptide ligands, the protocol generated models with low  $C_{\alpha}$ -RMSDs from the native structure, and the consensus sequences of the models recapitulated those observed in phage-display or site-directed-mutagenesis experiments. This is a promising approach to simultaneously model backbone flexibility and sequence diversity which may prove useful in protein-peptide interface design and prediction.

## CHALLENGES AND FUTURE DIRECTIONS

Advances in peptide docking and design (for example Refs. 95–98) have raised hope for the tractability and relevance of rational design, including docking of peptides and other large and flexible molecules to proteins. Below we highlight several directions that await further developments.

### Calculation of Binding Affinity and Binder/Nonbinder Classification

In addition to the structural details of interactions that may be obtained by docking or experimentally, it is important to be able to predict the affinity of binding (or complete lack of binding) of peptides to target proteins. The binding affinity is determined by the interplay between energy and entropy. Upon binding, the ligand becomes less mobile, and the

resulting loss in configurational entropy opposes the attractive forces that drive binding.<sup>99</sup> Energy models often account for changes in torsional entropy with a term related to the number of rotatable bonds in a ligand. However, recent results suggest that the physical model underlying these rotamer-counting approaches overestimates the number of stable conformations accessed by the ligand in the free state. Vibrational entropy, on the other hand is often underestimated. The common assumption that different ligands lose the same amount of rotational and translational entropy upon binding also appears to be unsupported.<sup>100</sup> These challenges apply to ligands in general, but are even more significant for the flexible peptides. Nevertheless, a combination of (even nonideal) physics-based methods with machine learning approaches appears to be promising.<sup>83,101</sup> For example, a support vector machine (SVM) was employed by Hou et al. to build classification models based on the molecular interaction energy of peptides with a SH3 domain. The best SVM classifier had prediction accuracies of 78% and 91% for the binding and nonbinding peptides, respectively.<sup>101</sup>

### Incorporation of Explicit Water

Water molecules can modify the shape and flexibility of a ligand-binding site in a protein, improving the complementarity between the two. In drug design, attempts have been made to incorporate water molecules into ligands at their binding sites with the goal of improving binding affinities, as in the design of a class of HIV protease inhibitors.<sup>102</sup> Interestingly, however, the displacement of interfacial water molecules does not always lead to an increase in the ligand's binding affinity. It has been shown that the inclusion of critical water molecules can lead to significant improvement in docking results,<sup>102</sup> and that some, but not all crystallographic water molecules can make significant contributions to the protein-ligand's binding free energies.<sup>103</sup> Explicit solvent MD and calculation of entropic contributions in preformed protein-peptide complexes have been investigated (for example Ref. 104). The simulation of the dynamic process of peptide binding to a protein in explicit solvent is another exciting direction. The metadynamics algorithm was used to investigate the role of two peptides in the folding of HIV-1 protease monomer.<sup>7</sup> This algorithm is aimed at reconstructing the multidimensional free energy of complex systems, and is based on artificial dynamics performed in the space defined by a few collective coordinates, which are assumed to provide a coarse-grained description of the system. The dynamics is driven by the free energy of the system and is biased by a history-dependent potential constructed as a sum of Gaussians centered along the trajectory of the collective varia-

bles.<sup>105</sup> The dynamic process of peptide binding to SH3 binding domain has been simulated using unbiased atomistic MD simulations in explicit water.<sup>106</sup> Thirteen independent, 50–150 ns long simulations were performed starting from peptide-protein distances of 13–20 Å. One of these runs resulted in a near-native conformation of the complex. The simulated binding process was shown to have three characteristic phases: an initial rapid diffusion, a nonspecific-encounter complex, stabilized by salt bridges, and then a specific stable bound complex. The transition from the first to the second phase is driven by long-scale electrostatic force, while use of explicit solvent revealed that the transition to the last phase is accelerated by partial dewetting of the hydrophobic interface.

### Docking Without Prior Knowledge of Interaction Region or Anchoring Point

There are cases in which an extended surface is available for the protein-peptide interactions.<sup>96</sup> In such cases, biochemical (for example, mutational) data typically need to be included to narrow down the possible regions of interaction.<sup>96</sup> Purely *ab initio* prediction of the peptide-protein binding site remains a formidable challenge. Nevertheless, recent work has demonstrated an example of successful *ab initio* docking followed by redesign of a peptide, probably facilitated by the fact that the peptides in that study were cyclic.<sup>98</sup>

A somewhat related issue is that of prediction of peptide aggregation, relevant to the study of Alzheimer's and other protein misfolding diseases.<sup>107</sup> The mechanism by which peptides derived from the C-terminus of the A $\beta$ 42 protein inhibit the formation of neurotoxic oligomers of this protein (considered to be the cause of Alzheimer's disease) was investigated using a coarse-grained protein model and discrete MD approach.<sup>108</sup> Although incapable of providing a high-resolution structure of the complex, such methods enable the rationalization of experimentally observed phenomena.<sup>109</sup>

### Peptide-Like Molecules Including Non-Natural Amino Acids or Building Blocks

Modification of lead peptides often includes introduction of non-natural amino acids or modulation of the backbone. For example, cyclic peptides<sup>19–22,110</sup> and peptides that include non-native residues such as N-methylated<sup>111,112</sup> or D-amino acids,<sup>42</sup> may improve stability, cell permeability and efficacy of natural peptide leads. Sequence-based tools or knowledge-based protein force fields cannot be applied to these compounds, which are no longer merely short proteins. Structure-based approaches are applicable, but require adaptations, such as development of force-field parameters that

would describe the potential energy surface of the modified amino acids.

### Transformation into Small Molecules

A good model of a protein-peptide complex enables the determination of a pharmacophore model (i.e., geometrical arrangements of chemical features such as hydrogen bonding and electrostatic and hydrophobic interactions). This model may then be used to design small molecules that mimic the peptide.<sup>113</sup> New computational tools are being developed to facilitate the search for such peptide-mimicking molecules.<sup>114–116</sup> In a recent study, structures of peptides that correspond to known linear motifs and for which structures in complex with proteins are available, were used to search against FDA approved drugs to identify similar compounds.<sup>117</sup> It will be interesting to see how these approaches contribute to the field of rational design of peptidomimetic and peptide-inspired inhibitors.

## CONCLUSION

The combination of pharmaceutical need and recent advances in peptide synthesis makes it a particularly exciting time in the rational design of peptidic compounds and in the development of adequate computational methods. Indeed, computational tools suitable for the challenge of modeling peptide-protein interactions are being used by several groups and are starting to contribute to the modulation of protein-protein interactions and drug discovery. The development of force-field potentials suitable for peptide-like molecules which include non-natural amino acids is an important and promising direction. The peptidomimetics field is likely to benefit from the synergistic combination of biochemical studies, synthetic efforts and computational (sequence- and structure-based) techniques. We believe that computational effort directed at the many cases in which anchoring information of the peptide to the target protein is available will prove to be most fruitful, and will likely contribute to discovery and development of novel compounds and eventually drug candidates.

## REFERENCES

1. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv Drug Deliv Rev* 1997, 23, 3–25.
2. Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. *J Med Chem* 2002, 45, 2615–2623.
3. Booth, B.; Zimmel, R. *Nat Rev Drug Discov* 2004, 3, 451–456.
4. Block, P.; Weskamp, N.; Wolf, A.; Klebe, G. *Proteins* 2007, 68, 170–186.
5. Hayouka, Z.; Rosenbluh, J.; Levin, A.; Maes, M.; Loyter, A.; Friedler, A. *Biopolymers* 2008, 90, 481–487.

6. Fletcher, S.; Hamilton, A. D. *J R Soc Interface* 2006, 3, 215–233.
7. Bonomi, M.; Gervasio, F. L.; Tiana, G.; Provasi, D.; Broglia, R. A.; Parrinello, M. *Biophys J* 2007, 93, 2813–2821.
8. Wells, J. A.; McClendon, C. L. *Nature* 2007, 450, 1001–1009.
9. Goodman, C. M.; Choi, S.; Shandler, S.; DeGrado, W. F. *Nat Chem Biol* 2007, 3, 252–262.
10. Vagner, J.; Qu, H. C.; Hruby, V. J. *Curr Opin Chem Biol* 2008, 12, 292–296.
11. Purcell, A. W.; McCluskey, J.; Rossjohn, J. *Nat Rev Drug Disc* 2007, 6, 404–414.
12. Sulochana, K. N.; Ge, R. *Curr Pharm Des* 2007, 13, 2074–2086.
13. Yusuf-Makagiansar, H.; Anderson, M. E.; Yakovleva, T. V.; Murray, J. S.; Siahaan, T. J. *Med Res Rev* 2002, 22, 146–167.
14. Jain, R.; Chawrai, S. *Mini-Rev Med Chem* 2005, 5, 469–477.
15. Gozes, I.; Divinski, I. *Curr Alzheimer Res* 2007, 4, 507–509.
16. Reményi, A.; Good, M. C.; Lim, W. A. *Curr Opin Struct Biol* 2006, 16, 676–685.
17. Diella, F.; Haslam, N.; Chica, C.; Budd, A.; Michael, S.; Brown, N. P.; Trave, G.; Gibson, T. J. *Front Biosci* 2008, 13, 6580–6603.
18. Petsalaki, E.; Russell, R. B. *Curr Opin Biotechnology* 2008, 19, 344–350.
19. Kasher, R.; Oren, D. A.; Barda, Y.; Gilon, C. *J Mol Biol* 1999, 292, 421–429.
20. Sharma, S. C.; Rupasinghe, C. N.; Parisien, R. B.; Spaller, M. R. *Biochemistry* 2007, 46, 12709–12720.
21. Hess, S.; Linde, Y.; Ovadia, O.; Safrai, E.; Shalev, D. E.; Swed, A.; Halbfinger, E.; Lapidot, T.; Winkler, I.; Gabinet, Y.; Faier, A.; Yarden, D.; Xiang, Z.; Portillo, F. P.; Haskell-Luevano, C.; Gilon, C.; Hoffman, A. *J Med Chem* 2008, 51, 1026–1034.
22. Qvit, N.; Reuveni, H.; Gazal, S.; Zundevich, A.; Blum, G.; Niv, M. Y.; Feldstein, A.; Meushar, S.; Shalev, D. E.; Friedler, A.; Gilon, C. *J Comb Chem* 2008, 10, 256–266.
23. Chatterjee, J.; Gilon, C.; Hoffman, A.; Kessler, H. *Acc Chem Res* 2008, 41, 1331–1342.
24. Shah, N. H.; Butterfoss, G. L.; Nguyen, K.; Yoo, B.; Bonneau, R.; Rabenstein, D. L.; Kirshenbaum, K. *J Am Chem Soc* 2008, 130, 16622–16632.
25. Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. *Nat Rev Drug Discov* 2008, 7, 608–624.
26. Pecuh, M. W.; Hamilton, A. D. *Chem Rev* 2000, 100, 2479–2493.
27. Sebolt-Leopold, J. S.; English, J. M. *Nature* 2006, 441, 457–462.
28. Bozulic, L.; Morin, P., Jr.; Hunter, T.; Hemmings, B. A. *Sci STKE* 2007, 2007, pe8.
29. Budas, G. R.; Koyanagi, T.; Churchill, E. N.; Mochly-Rosen, D. *Biochem Soc Trans* 2007, 35, 1021–1026.
30. Kheifets, V.; Mochly-Rosen, D. *Pharm Res* 2007, 55, 467–476.
31. Laudet, B. A.; Barette, C.; Dulery, V.; Renaudet, O.; Dumy, P.; Metz, A.; Prudent, R.; Deshiere, A.; Dideberg, O.; Filhol, O.; Cochet, C. *Biochem J* 2007, 408, 363–373.
32. Taylor, J. D.; Ababou, A.; Fawaz, R. R.; Hobbs, C. J.; Williams, M. A.; Ladbury, J. E. *Proteins* 2008, 73, 929–940.
33. Wexler, I. D.; Niv, M. Y.; Reuveni, H. *Biotechniques* 2005, 39, S575–S576.
34. Mack, E.; Ziv, E.; Reuveni, H.; Kalman, R.; Niv, M. Y.; Jorns, A.; Lenzen, S.; Shafir, E. *Diabetes Metab Res Rev* 2008, 24, 577–584.
35. Ye, G.; Tiwari, R.; Parang, K. *Curr Opin Investig Drugs* 2008, 9, 605–613.
36. Souroujon, M. C.; Mochly-Rosen, D. *Nat Biotech* 1998, 16, 919–924.
37. Ron, D.; Mochly-Rosen, D. *J Biol Chem* 1994, 269, 21395–21398.
38. Litman, P.; Ohne, O.; Ben-Yakov, S.; Shemesh-Darvish, L.; Yechezkel, T.; Salitra, Y.; Rubnov, S.; Cohen, I.; Senderowitz, H.; Kidron, D.; Livnah, O.; Levitzki, A.; Livnah, N. *Biochemistry* 2007, 46, 4716–4724.
39. Bogoyevitch, M. A.; Fairlie, D. P. *Drug Discov Today* 2007, 12, 622–633.
40. Harrison, S.; Das, K.; Karim, F.; Maclean, D.; Mendel, D. *Expert Opin Drug Discov* 2008, 3, 761–774.
41. Niv, M. Y.; Rubin, H.; Cohen, J.; Tsurulnikov, L.; Licht, T.; Peretzman-Shemer, A.; Cna'an, E.; Tartakovsky, A.; Stein, I.; Albeck, S.; Weinstein, I.; Goldenberg-Furmanov, M.; Tobi, D.; Cohen, E.; Laster, M.; Ben-Sasson, S. A.; Reuveni, H. *J Biol Chem* 2004, 279, 1242–1255.
42. Goldenberg-Furmanov, M.; Stein, I.; Pikarsky, E.; Rubin, H.; Kasem, S.; Wygoda, M.; Weinstein, I.; Reuveni, H.; Ben-Sasson, S. A. *Cancer Res* 2004, 64, 1058–1066.
43. Yeon, S. W.; Jeon, Y. J.; Hwang, E. M.; Kim, T. Y. *Peptides* 2007, 28, 838–844.
44. Cole, S. L.; Vassar, R. *J Biol Chem* 2008, 283, 29621–29625.
45. Cooper, D. A.; Lange, J. M. A. *Lancet Infect Dis* 2004, 4, 426–436.
46. Conklin, B. R.; Hsiao, E. C.; Claeys, S.; Dumuis, A.; Srinivasan, S.; Forsayeth, J. R.; Guettier, J.-M.; Chang, W. C.; Pei, Y.; McCarthy, K. D.; Nissenson, R. A.; Wess, J.; Bockaert, J.; Roth, B. L. *Nat Methods* 2008, 5, 673–678.
47. Hopkins, A. L.; Groom, C. R. *Nat Rev Drug Discov* 2002, 1, 727–730.
48. Tarasova, N. I.; Rice, W. G.; Michejda, C. J. *J Biol Chem* 1999, 274, 34911–34915.
49. Hebert, T. E.; Moffett, S.; Morello, J.-P.; Loisel, T. P.; Bichet, D. G.; Barret, C.; Bouvier, M. *J Biol Chem* 1996, 271, 16384–16392.
50. George, S. R.; Ng, G. Y.; Lee, S. P.; Fan, T.; Varghese, G.; Wang, C.; Deber, C. M.; Seeman, P.; O'Dowd, B. F. *J Pharmacol Exp Ther* 2003, 307, 481–489.
51. Kubo, S.; Ishiki, T.; Doe, I.; Sekiguchi, F.; Nishikawa, H.; Kawai, K.; Matsui, H.; Kawabata, A. *Ann N Y Acad Sci* 2006, 1091, 445–459.
52. Licht, T.; Tsurulnikov, L.; Reuveni, H.; Yarnitzky, T.; Ben-Sasson, S. A. *Blood* 2003, 102, 2099–2107.
53. Thiagaraj, H. V.; Ortiz, T. C.; Devereaux, M. C., Jr.; Seaver, B.; Hall, B.; Parker, K. K. *Neurochem Int* 2007, 50, 109–118.
54. Linde, Y.; Ovadia, O.; Safrai, E.; Xiang, Z.; Portillo, F. P.; Shalev, D. E.; Haskell-Luevano, C.; Hoffman, A.; Gilon, C. *Peptide Sci* 2008, 90, 671–682.
55. Goodman, S. L.; Holzemann, G.; Sulyok, G. A. G.; Kessler, H. *J Med Chem* 2002, 45, 1045–1051.
56. Eichler, J. *Curr Opin Chem Biol* 2008, 12, 707–713.
57. Hong, H. Y.; Lee, H. Y.; Kwak, W.; Yoo, J.; Na, M. H.; So, I. S.; Kwon, T. H.; Park, H. S.; Huh, S.; Oh, G. T.; Kwon, I. C.; Kim, I. S.; Lee, B. H. *J Cell Mol Med* 2008, 12, 2003–2014.
58. Tonikian, R.; Zhang, Y.; Sazinsky, S. L.; Currell, B.; Yeh, J. H.; Reva, B.; Held, H. A.; Appleton, B. A.; Evangelista, M.; Wu, Y.; Xin, X.; Chan, A. C.; Seshagiri, S.; Lasky, L. A.; Sander, C.; Boone, C.; Bader, G. D.; Sidhu, S. S. *PLoS Biol* 2008, 6, e239.



59. Kim, E.-Y.; Bang, J. Y.; Chang, S.-I.; Kang, I.-C. *Biochem Biophys Res Comm* 2008, 377, 1288–1293.
60. Chen, J. R.; Chang, B. H.; Allen, J. E.; Stiffler, M. A.; MacBeath, G. *Nat Biotech* 2008, 26, 1041–1045.
61. Joo, S. H.; Xiao, Q.; Ling, Y.; Gopishetty, B.; Pei, D. *J Am Chem Soc* 2006, 128, 13000–13009.
62. Edwards, R. J.; Moran, N.; Devocelle, M.; Kiernan, A.; Meade, G.; Signac, W.; Foy, M.; Park, S. D. E.; Dunne, E.; Kenny, D.; Shields, D. C. *Nat Chem Biol* 2007, 3, 108–112.
63. Henderson, G.; Bradley, M. *Curr Opin Biotechnol* 2007, 18, 326–330.
64. Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *Proteins* 2006, 65, 15–26.
65. Kontoyianni, M.; McClellan, L. M.; Sokol, G. S. *J Med Chem* 2004, 47, 558–565.
66. Perola, E.; Walters, W. P.; Charifson, P. S. *Proteins* 2004, 56, 235–249.
67. Bursulaya, B. D.; Totrov, M.; Abagyan, R.; Brooks, C. L. *JCAMD* 2003, 17, 755–763.
68. Dominguez, C.; Boelens, R.; Bonvin, A. M. *J Am Chem Soc* 2003, 125, 1731–1737.
69. Totrov, M.; Abagyan, R. *Proteins* 1997, 29, 215–220.
70. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J Comp Chem* 1998, 19, 1639–1662.
71. Ewing, T. J. A.; Makino, S.; Skillman, A. G.; Kuntz, I. D. *JCAMD* 2001, 15, 411–428.
72. Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J Mol Biol* 1996, 261, 470–489.
73. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J Mol Biol* 1997, 267, 727–748.
74. Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D. *Proteins* 2004, 57, 225–242.
75. Jain, A. N. *J Med Chem* 2003, 46, 499–511.
76. McGann, M. R.; Almond, H. R.; Nicholls, A.; Grant, J. A.; Brown, F. K. *Biopolymers* 2003, 68, 76–90.
77. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J Med Chem* 2004, 47, 1739–1749.
78. Arun Prasad, P.; Gautham, N. *JCAMD* 2008, 22, 815–829.
79. Kurcinski, M.; Kolinski, A. *J Steroid Biochem Mol Biol* 2007, 103, 357–360.
80. Maurer, M. C.; Trosset, J. Y.; Lester, C. C.; DiBella, E. E.; Scheraga, H. A. *Proteins* 1999, 34, 29–48.
81. Niv, M. Y.; Weinstein, H. *J Am Chem Soc* 2005, 127, 14072–14079.
82. Sood, V. D.; Baker, D. *J Mol Biol* 2006, 357, 917–927.
83. Bordner, A. J.; Abagyan, R. *Proteins* 2006, 63, 512–526.
84. Trosset, J.-Y.; Scheraga, H. *J Comp Chem* 1999, 20, 412–427.
85. Noguti, T.; Go, N. *Biopolymers* 1985, 24, 527–546.
86. Schafroth, H. D.; Floudas, C. A. *Proteins* 2004, 54, 534–556.
87. Tong, J. C.; Tan, T. W.; Ranganathan, S. *Protein Sci* 2004, 13, 2523–2532.
88. Stein, A.; Aloy, P. *PLoS ONE* 2008, 3, e2524.
89. Brooks, B. R.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J Comput Chem* 1983, 4, 187–217.
90. Canutescu, A. A.; Shelenkov, A. A.; Dunbrack, R. L., Jr. *Protein Sci* 2003, 12, 2001–2014.
91. Kang, B. S.; Cooper, D. R.; Jelen, F.; Devedjiev, Y.; Derewenda, U.; Dauter, Z.; Otlewski, J.; Derewenda, Z. S. *Structure* 2003, 11, 459–468.
92. Madsen, K. L.; Beuming, T.; Niv, M. Y.; Chang, C. W.; Dev, K. K.; Weinstein, H.; Gether, U. *J Biol Chem* 2005, 280, 20539–20548.
93. Bach, A.; Chi, C. N.; Olsen, T. B.; Pedersen, S. R. W.; Røder, M. U.; Pang, G. F.; Clausen, R. P.; Jemth, P.; Strømgaard, K. *J Med Chem* 2008, 51, 6450–6459.
94. Abagyan, R.; Totrov, M. *J Mol Biol* 1994, 235, 983–1002.
95. Baldauf, C.; Pisabarro, M. T. *J Phys Chem B* 2008, 112, 7581–7591.
96. Niv, M. Y.; Iida, K.; Zheng, R.; Horiguchi, A.; Shen, R.; Nanus, D. M. *Prot Sci*, in press.
97. Caputo, G. A.; Litvinov, R. I.; Li, W.; Bennett, J. S.; DeGrado, W. F.; Yin, H. *Biochemistry* 2008, 47, 8600–8606.
98. Hao, J.; Serohijos, A. W. R.; Newton, G.; Tassone, G.; Wang, Z.; Sgroi, D. C.; Dokholyan, N. V.; Basilion, J. P. *PLoS Comp Biol* 2008, 4, e1000138.
99. Gilson, M. K.; Zhou, H. X. *Ann Rev Biophys Biomol Struct* 2007, 36, 21–42.
100. Chang, C.-E. A.; Chen, W.; Gilson, M. K. *Proc Natl Acad Sci USA* 2007, 104, 1534–1539.
101. Hou, T.; Zhang, W.; Case, D. A.; Wang, W. *J Mol Biol* 2008, 376, 1201–1214.
102. Huang, N.; Shoichet, B. K. *J Med Chem* 2008, 51, 4862–4865.
103. Lu, Y.; Wang, R.; Yang, C. Y.; Wang, S. *J Chem Inf Model* 2007, 47, 668–675.
104. Basdevant, N.; Weinstein, H.; Ceruso, M. *J Am Chem Soc* 2006, 128, 12766–12777.
105. Laio, A.; Rodriguez-Forteza, A.; Gervasio, F. L.; Ceccarelli, M.; Parrinello, M. *J Phys Chem B* 2005, 109, 6714–6721.
106. Ahmad, M.; Gu, W.; Helms, V. *Angewandte Chemie-Int Ed* 2008, 47, 7626–7630.
107. Ma, B.; Nussinov, R. *Curr Opin Chem Biol* 2006, 10, 445–452.
108. Yun, S.; Urbanc, B.; Cruz, L.; Bitan, G.; Teplow, D. B.; Stanley, H. E. *Biophys J* 2007, 92, 4064–4077.
109. Fradinger, E. A.; Monien, B. H.; Urbanc, B.; Lomakin, A.; Tan, M.; Li, H.; Spring, S. M.; Condrón, M. M.; Cruz, L.; Xie, C. W.; Benedek, G. B.; Bitan, G. *Proc Natl Acad Sci USA* 2008, 105, 14175–14180.
110. Hurevich, M.; Barda, Y.; Gilon, C. *Heterocycles* 2007, 73, 617–625.
111. Chatterjee, J.; Ovadia, O.; Zahn, G.; Marinelli, L.; Hoffman, A.; Gilon, C.; Kessler, H. *J Med Chem* 2007, 50, 5878–5881.
112. Chatterjee, J.; Gilon, C.; Hoffman, A.; Kessler, H. *Acc Chem Res* 2008, 41, 1331–1342.
113. Kiran, M. D.; Adikesavan, N. V.; Cirioni, O.; Giacometti, A.; Silvestri, C.; Scalise, G.; Ghiselli, R.; Saba, V.; Orlando, F.; Shoham, M.; Balaban, N. *Mol Pharmacol* 2008, 73, 1578–1586.
114. Eckert, H.; Bajorath, J. *J Chem Inf Model* 2007, 47, 1366–1378.
115. Goede, A.; Michalsky, E.; Schmidt, U.; Preissner, R. *BMC Bioinformatics* 2006, 7, 5.
116. Casey, F. P.; Davey, N. E.; Baran, I.; Varekova, R. S.; Shields, D. C. *J Chem Inf Modell* 2008, 48, 1524–1529.
117. Parthasarathi, L.; Casey, F.; Stein, A.; Aloy, P.; Shields, D. C. *J Chem Inf Modell* 2008, 48, 1943–1948.

Reviewing Editor: Eric Toone