

Computer-Aided Drug Design 2003–2005

BY BERNARD COUPEZ, HENRIK MÖBITZ AND RICHARD A. LEWIS

Novartis Institutes for Biomedical Research, Basel CH-4002, Switzerland

1 Introduction

The themes for this review again have been driven strongly by the need of the Pharmaceutical industry to make the discovery process quicker and more reliable. Virtual screening in all its forms is at the heart of most research, from bioavailability filters through to rigorous estimations of the free energy of binding. Two areas of relative heat have been docking/scoring, and ADME/Tox. On the other hand, 3D-QSAR and pharmacophores have become quiet. Part of the reason for this may arise from the successes in high-throughput crystallography, delivering more targets and complexes, the relative failure of HTS, and the increase in the amount of high quality data coming from late-phase research/early-phase development concerning the fate of clinical candidates. These trends look set to continue in the future, and the next two years should yield many new breakthroughs.

2 ADME/Tox and Druggability

There has been a fresh impetus to the modelling of ADME, Toxicity and druggability phenomena, partly driven by a desire to understand why such complex phenomena can, apparently, be described so simply, and partly to see if better models can be built, to improve the attrition rate in medicinal chemistry still further.

2.1 Druggability and Bioavailability. – In the continuing debate over what physicochemical properties are required for bioavailability, Vieth *et al.*¹ have surveyed 1729 marketed drugs with respect to their route of administration, h-bonding capability, lipophilicity and flexibility. One conclusion they draw is that these properties have not varied substantially over time, implying that oral bioavailability is independent of target or molecular complexity. Compounds with lower molecular weight, balanced lipophilicity and less flexibility tend to be favoured. Leeson and Davis² claim that molecular weight, flexibility, the number of O and N atoms and hydrogen-bond acceptors have risen, by up to

29%. This may be partly due to the choice of 1983 as the reference year, or the advent of more complex targets with greater selectivity needs (*e.g.* kinases). In the same vein, a study³ re-examined the correlation of flexibility and polar surface area (PSA) with bioavailability proposed by Veber *et al.*⁴ One conclusion is that there are significant differences in the ways of defining flexibility and PSA, and the correlations depend markedly on the method used (this is not surprising, as neither quantity is precisely definable). A second conclusion was that the limits defined (Number of rotatable bond < 10, PSA < 140 Å²) excluded a significant number of compounds with acceptable rat bioavailability. In the authors' words, "This observation underscores the potential danger of attempting to generalise a very complicated endpoint and of using that generalisation in a prospective selection application". Despite this, another bioavailability score⁵ has been devised, to predict the probability that a compound has >10% bioavailability in the rat. Compounds are grouped by ionisation class (anions, cations, neutral). It was found that the standard rule-of-5 does well for cations and neutrals (88% of the compounds predicted to have low bioavailability are observed as such). Anionic compounds were better described by PSA limits. Some simple rules are given to compute the bioavailability score. In Abbott laboratories, this score is now routinely computed for all compounds and is used for hit-list triaging. It will be interesting to see if the results can be repeated on other data sets; the paper has certainly sparked much interest in the modelling community. Wegner⁶ provides support for the idea that human intestinal absorption correlates with PSA, by generating a classification model. The justification is that the error in the experimental data is 25%, and 80% of the observations occur in the top and bottom quartiles, that is, the data is more binary than evenly spread. In addition to PSA, other descriptors that reflect the electronic character of atoms and their environment also came to the fore.

2.2 Metabolism, Inhibitors and Substrates. – The field of cytochrome modelling is becoming more mature as we begin to understand the limitations of the experimental data and the subtleties of the mechanisms (the whole field of cytochrome P450 modelling, including homology, pharmacophore and 3D-QSAR models has been reviewed in detail recently⁷). Empirical models are still preferred, especially for rapid evaluation of large libraries. In one case, use of a jury system improved prediction accuracy to over 90%.⁸ Chohan *et al.*⁹ have developed 4 models for Cytochrome P450 (Cyp) 1A2 inhibition, and identified the expected descriptors as being important to the QSAR (lipophilicity, aromaticity, HOMO/LUMO energies). Perhaps a more interesting result in this paper was the use of the *k* index to assess predictive powers of the models using test data.

$$k = \frac{\text{observed agreement-chance agreement}}{\text{total observed-chance agreement}}$$

This index should prove useful for data sets that are diverse and noisy. The validity of QSAR model predictions has also been studied by Guha and Jurs.¹⁰

The protocol is quite straightforward. The initial QSAR models were built, and the residuals of the compounds in the training set were used to classify the training set predictions into good and bad. The threshold for the classification is arbitrary. Test compounds were predicted, and the predictions were grouped by substructural similarity to the nearest neighbour in the training set. It was seen that test compounds that had neighbours with low/good residuals were themselves well-predicted, with the reverse being the case for neighbours with high residuals. The success rate for classifying the strength of the prediction was 73% to 94%. The Merck group¹¹ performed a retrospective study of in-house data sets, and concluded that the distance to the nearest neighbour, and the number of nearest neighbours (local density) were the two most useful measures for predicting prediction quality. They also concluded that distance does not have to be measured in the same descriptor space as was used to build the QSAR model. Topological descriptors combined with a Dice coefficient worked equally well.

A number of groups have been active in the prediction of the most likely sites of metabolism of molecules that are substrates for cytochromes. Singh *et al.*¹² developed a semi-quantitative method based on the energy barrier to the creation of hydrogen radicals as calculated by AM1. Using a set of 50 substrates for Cyp 3A4, they were able to show that only hydrogens with a solvent-accessible surface area over 8 \AA^2 are susceptible to attack. The expensive quantum mechanic calculations could be approximated by local neighbourhood descriptors which could be well correlated to the energies ($R^2 = 0.98$), offering a fast and practical method for screening large libraries. An extension of this concept is embodied in the MetaSite program,¹³ which uses propensity to react, accessibility and GRID molecular interaction fields as descriptors. The methodology is more general, and can be applied to any cytochrome structure: in validation experiments, an accuracy of 80% is claimed. It is also important to be able to predict which compounds will be inhibitors as well as substrates, to avoid drug-drug interactions. A classifier based on a support vector machine (SVM)¹⁴ has been created that correctly predicts compounds into high, medium and low affinity at 70% accuracy, even with simple 2D descriptors. The improved accuracy was obtained through a systematic variation and optimisation of the SVM parameters.

Considering the success of surprisingly simple, semiempirical methods in ADME modelling, it is interesting to see whether more advanced methods could bring further improvements. A recent paper of Beck¹⁵ provides a link to the rich literature of DFT studies of hemes and cytochromes. The author uses Fukui functions to gauge the site of highest nucleophilicity of a number of known drugs. The predictions give mixed results and demonstrate that the implicit assumption of Fukui functions, *i.e.* an isotropic electrophilic attack, is flawed, not to mention that their MO-like shape does not allow a ranking of single atoms. In conclusion, the study suggests that it is more important to have an accurate description of the cytochrome-ligand complex than to invest in a high-level description of the chemical reactivity. De Visser *et al.*¹⁶ have used DFT on 10 C–H barriers with reference to bacterial cytochromes, and claim an

excellent correlation between bond energy and observed activation energy barriers, so there is still some mileage in this approach.

2.3 Toxicity. – Unacceptable toxicity is still a key source of compound failure in clinical trials. Several groups have developed tools and programs for predicting toxicity for use in early phase, but the question arises about the accuracy of these models, and the levels of false positives and negatives that are acceptable. In research, an overly strict model with no false negatives may cause the discarding of a perfectly reasonable lead series. In development, missing a toxic alert which shows up in a later phase is unacceptable. Similarly, any program that is used by regulatory authorities to screen compounds must be very unforgiving of any flaw. In a recent study by the FDA¹⁷ on maximum human therapeutic dose, rules-based programs managed 64% accuracy, not much over random, giving an indication of the pitfalls in this field; Helma has given an overview of this area.¹⁸ Clearly, the domain of the models is critical and this has been addressed explicitly for QSARs that make toxicity predictions.¹⁹ Another route to predicting ADME properties is to use screening results, as exemplified by the Bioprint approach.²⁰ 1198 drugs have been assayed against 130 screens, to give an activity fingerprint. QSAR models are then derived using pharmacophore descriptors. New compounds can be run through the models to predict binding affinity in all the screens, compared to the nearest neighbours in the database and finally fingerprinted themselves for confirmation. Using the affinity fingerprint alone, one can again identify similar molecules (sometimes with surprising results) and extrapolate to the potential side-effect profiles. This is very useful when selecting one from several lead series for optimisation.

3 Docking and Scoring

3.1 Ligand Database Preparation. – The ligand database is the basis for virtual screening (VS). Special care must be taken at this stage; accurate and physically relevant tautomeric and protonation states need to be assigned. Often compounds are registered in a database as a tautomer that is not necessarily the most probable state of the molecule and it is difficult to assign the correct state, so all relevant states should be generated. Similarly, as the stereochemistry of chiral centres is often not known, one must generate all stereoisomers. A recent article reveals the impact of pre-processing a database containing both known actives and inactives, where multiple protonated, tautomeric, stereochemical, and conformational states have been enumerated.²¹ The authors show that the interplay between 2D representations, stereochemical information, protonation states, and ligand conformation ensembles has a profound effect on the success rates of VS and conclude that the enrichment is highly dependent on the initial treatments used in database construction. In a paper that is bound to become a citation classic for the service that it has provided to the academic modelling community, Irwin and Shoichet describe the creation of the ZINC database of

commercially available compounds, available via the web.²² The resource can be used in virtual screening studies, as the authors have taken care to provide compounds in multiple protonation and tautomer states, even multiple conformations. The paper provides a useful recipe for creating such a database for general use.

3.2 Target Preparation. – Thanks to high throughput crystallography and structural genomics, we have the X-ray structures of many targets of therapeutic interest, with the obvious exception of membrane proteins. When no experimental structure is available, it is possible to generate a 3D structure based on a template protein of similar sequence and a known structure, for example the model of CDK10,²³ based on the CDK2 crystal structure, that was successfully used for a docking study.

If several structures of the target are available, which structures should be used: the apo form, a holo complex, or a homology model? This issue was examined by McGovern *et al.*²⁴ They docked a large number of small molecules against 10 targets using the apo-, holo-, and modelled forms of the binding site. Using enrichment rates, they found that the holo form gave the best results (70% enrichment) followed by the apo (20%) and then the modelled form (10%). However, the holo form can be over influenced by the ligand in holo complex, if the active site has “collapsed” around the ligand. Then one would get a lower retrieval rate of similar but larger ligands, due to the increased steric constraints; the apo form of an active site can be markedly different from holo form.²⁵ The conclusion is that VS using any form of the target will do better than random, but the holo form will give a best enrichment. This was also confirmed by Erickson *et al.*²⁶ which show that the docking accuracy decreases dramatically if one uses an average or apo structure. Another approach is to use softened repulsive terms in the Lennard-Jones potential, to allow a closer approach of ligand and protein atoms that could be later resolved by minimisation.²⁷ The T4 lysozyme system was used, with the ACD database as the source of ligands. The soft function was worse than the hard function, if multiple protein conformations were used, and vice versa for a single model. It was concluded that soft potential favour the decoys as much as the true ligands, so needs to be used with care.

Like the ligand preparation, the preparation of the target also requires great care. Incorrect protonation states or tautomers of histidines can lead to serious docking errors. For example Polgar *et al.*²⁸ demonstrate the importance of protonation states in virtual screening for β -secretase (BACE1) inhibitors. They observed improvement of enrichment rates when they assigned different protonation states to catalytic Asp32 and Asp228 residues. Some docking methods require the addition of hydrogens. It is recommended that after the addition of hydrogen atoms to the protein, the positions of the hydrogens are relaxed by energy minimization to avoid any steric clashes. The positioning of hydrogen atoms on hydroxyl groups in the active site should also be checked and changed if necessary. In some instances, hydrogen bonds to crystallographic waters might need to be maintained for the docking.

However increasing the degrees of flexibility also increases the computational complexity and cost. Different methods have been described in the literature to tackle this critical issue (for a review see ref. 29). Often these methods model the flexibility in the binding site exclusively, by sampling the protein conformational space using molecular dynamics or Monte Carlo calculations or rotamer libraries. Another way of treating protein flexibility is to use an ensemble of protein conformations, rather than a single one. In a recent paper, Barril and Morley³⁰ use all the X-ray structures of cyclin dependant kinase 2 (CDK2) and heatshock protein 90 (HSP90) to assess the performance of flexible receptor docking. They observe that flexible receptor docking performs much better in binding-mode prediction than rigid receptor docking. However, they also noticed that for library screening, ensembles of cavities often result in worse hit rates than rigid docking. This trend can be reversed by selecting those ligands that bind consistently well to many cavities in the ensemble.

3.3 Water Molecules. – Another challenge in protein-ligand docking is the modelling of the water molecules in protein ligand recognition. Water can form hydrogen bonds between the protein and the ligand or can be displaced by the ligand.³¹ Recently a new approach that allows this was implemented by Verdonk *et al.* in GOLD.³² The method allows water molecules to switch on and off and to spin. The explicit inclusion of water molecules in a docking program improves the binding mode when a ligand interacts with a water molecule. A distinction can also be made between the compounds that can displace a water molecules and the compounds that cannot. They claim that their algorithm correctly predicts water mediation/displacement in 93% of their tests and they observe some slight improvements in binding mode quality for water-mediated complexes. Similar results were reported by De Graf *et al.*³³ for cytochrome P450s. The waters were either removed, or the crystallographic waters retained, or waters in GRID minima were used. Surprisingly, the last scenario gave the best results by up to 20% in the number of correct poses that scored highest.

3.4 Comparison of Docking Methods. – The flow of papers performing comparative evaluations of docking/scoring programs continues,^{26,34} although there is an increasing feeling that these studies offer only limited insight.³⁵ Another potential pitfall in studies evaluating docking and scoring functions has been highlighted.³⁶ Enrichment rates can be artificially boosted by not matching the ID properties of the decoy set to the true ligands (for example, if the ligand is much larger than the decoys, it will be favoured). It was also observed that incorporating even small amounts of chemical knowledge, in the form of pharmacophoric constraints, could improve the quality of binding modes, and hence the enrichment. The work of Warren *et al.*³⁷ deserves mention as their protocol did not rely on evaluations performed with default parameter settings, but rather let expert users set up the runs, which is a more realistic scenario. Their conclusion was that no one approach was clearly better than any other for all targets; all failed to predict binding affinity with any

confidence. Against that, the study was performed with quite old versions of the software, so whether the same conclusions are valid today is moot.

3.5 Scoring. – Success of VS depends more strongly on the quality of the scoring function, than the method for generating dockings. An imperfect scoring function can mislead by predicting incorrect ligand geometries or by selecting nonbinding molecules over true ligands. Graves *et al.*³⁸ consider these false-positive hits as decoys and have used them to improve their protein-protein docking algorithms. A new version of the knowledge-based scoring function DrugScore has been published,³⁹ based on better quality small molecule X-ray data. This has the advantage of being higher resolution and better populated than protein x-ray data sources. For common interactions (for example, C.sp3 – C.sp3), the shapes of the potentials are the same, with more definition from the new potentials, reflecting the higher resolution of the underlying data. When this scoring function was used to dock and score 100 complexes with decoys, the crystallographic pose was ranked in the top 3 for 90% of the complexes, a 57% improvement over the previous version. The rank order coefficient for the prediction of binding energy was improved slightly (0.62), but we are still not doing significantly better than the correlation with molecular weight (0.56). This phenomenon is also observed in high-throughput screening, leading to measures for ligand efficiency^{40,41} that correct for molecular weight. A simple method for computing thermodynamic energies of binding,⁴² allowing flexibility in the protein side chains via Monte Carlo sampling, and a very simple model for van der Waals and electrostatic interactions nonetheless proved to be quite effective in predicting the selectivity of 6 kinase inhibitors when tested against a panel of 20 receptors. The authors identified a strong dependence on a good initial binding pose, and saw that minimisation with the function did not improve the results. Some of the success may have come from working in a target family, when one can assume that many of the errors are consistent, so that the relative energies can be trusted. While free energy methods remain the gold standard for ligand affinity prediction, the associated computational cost prohibits their routine use in the pharmaceutical industry. In a series of papers, Oostenbrink and van Gunsteren^{43–45} have tackled this problem and extended the one-step perturbation method with the aim of a fast, accurate prediction of structurally diverse compounds. In principle, through the use of a well-chosen reference state, the computational cost is reduced to a single full simulation. In a study of the estrogen binding receptor, a series of biphenyl compounds were predicted with an error of $<1 \text{ kcal} \cdot \text{mol}^{-1}$, whereas the predictions for a more diverse set of compounds hint that the method needs further improvement before it can be generally applied. One such improvement is the reparameterization of the underlying GROMOS force-field for the prediction of thermodynamic properties of hydration and solvation.⁴⁶

As no scoring function is perfect and each scoring function has its own strengths and weaknesses, we can combine different scoring functions to balance errors of one single scoring function and improve the probability of

identifying ‘true’ ligands by reducing the false positive rates. This approach is called consensus scoring. However, the potential value of consensus scoring might be limited, if terms in different scoring functions are significantly correlated, which could amplify calculation errors, rather than balance them. The success of the consensus scoring approach was analysed by Yang *et al.*⁴⁷ Using data from five scoring systems with two evolutionary docking algorithms on four targets, thymidine kinase, human dihydrofolate reductase, and estrogen receptors in antagonist and agonist conformations, the authors demonstrated that combining multiple scoring functions improves the enrichment of true positives only if each of the individual scoring functions has relatively high performance and if the individual scoring functions are very distinct in their philosophy. Recently an alternative way of combining various scores was proposed by Vigers and Rizzi;⁴⁸ this approach called “multiple active site correction” can correct library ranking using scores calculated for several active sites. The corrected score is now high only if compounds are found to score well with the target of interest and not with others.

3.6 New Methods. – New docking methods have been developed during these last two years. Glide,⁴⁹ developed by Schrödinger, is one of the most popular. Firstly the properties of the active site are mapped on a grid. Then a set of low energy conformations of the ligand is generated using a Monte Carlo approach. These poses are used as input and the ligand is minimized in the binding site and three to six low energy poses are selected and a Monte Carlo simulation is performed on these. The AFMoC protocol has been further developed,⁵⁰ to adapt a scoring function with local knowledge provided by known complexes and measured affinities. The key advances have been to use filtering of grid point variables by Shannon entropy, and to use sensible defaults for potentials that became repulsive under the AFMoC protocol. Using a challenging test set of 66 highly flexible HIV-1 protease inhibitors, they were able to identify a correct binding pose with the top binding score in 75% of cases, an improvement of 14% over native scoring functions. Another twist for knowledge-based scoring functions is to optimise the ligand positions before fitting the scores to experiment.⁵¹ This removes the bias of the x-ray refinement protocol. An accuracy of 2 kcal·mol⁻¹ in binding energy prediction is claimed, but the results were not compared to the correlation of score to molecular weight.

A current trend in the field is to focus on the inclusion of various solvation and rotational entropy contributions. However the terms currently used to approximate entropy or desolvation energy provide only incomplete descriptions of these effects on protein–ligand binding. For example, Krammer *et al.*⁵² present developed two new empirical scoring functions that possess good predictive accuracy in determining the ligand-receptor binding affinities over a wide range of protein classes. A recently introduced new methodology based on ultrashort (50–100 ps) molecular dynamics simulations with a quantum-refined force-field (QRFF-MD)⁵³ was evaluated by Ferrara *et al.* using CDK2 kinase.⁵⁴ The QRFF-MD method achieves a correlation of 0.55, which is significantly better than that obtained by a number of traditional approaches in

virtual screening but only slightly better than that obtained by consensus scoring (0.50). The authors also introduced a new scoring function that combines a QRFF-MD based scoring function with consensus scoring, which resulted in substantial improvement on the enrichment profile.

With the increase of the computational power it is now possible to use more rigorous theoretical and more CPU-intensive approaches. Kuhn *et al.*⁵⁵ reported the usefulness of the MM-PBSA approach for VS: they showed that applying the MM-PBSA energy function to a single, relaxed complex structure is an adequate and sometimes more accurate approach than the standard free energy averaging. MM-PBSA can also be used as a post-docking filter for enriching virtual screening results, and for distinguishing between good and weak binders for which $\Delta pIC_{50} \geq 2-3$. Huang *et al.*⁵⁶ developed a two-stage virtual screening protocol: first a rapid, grid-based scoring function is used to dock large compound databases to a receptor. In the second step the OPLS all-atom force field and a generalized Born implicit solvent model is used to minimize the ligand in the cavity and to rescore the poses for the top 25% of the ligands from the docking phase.

One well-known strategy for improving throughput and accuracy of docking for hit-finding is to apply some extra screens to reduce the size of the database to be screened. Then one can use more expensive but hopefully more accurate protocols for the docking and scoring. Maiorov and Sheridan⁵⁷ started by using a fast docking protocol FLOG, then fed the best 1000 scoring results into ICM-Dock for redocking. They showed a 5-fold improvement of the enrichment obtained by FLOG alone. Use of this two step method meant that the entire MDDR database could be screening in under a day.

The role of fluorine in hydrogen-bonding has been difficult to quantify. In some cases, addition of a fluorine can bring great improvements in binding affinity, in many other cases, it seems to be neutral, even in cases where a positive interaction should take place. The GRID program⁵⁸ now includes a potential function for fluorine, based on the new survey of protein-ligand x-ray complexes. Aliphatic fluorines make straighter and shorter hydrogen bonds than aromatic fluorines. Bifurcated bonds are not observed at all. When the new term was added to GRID and the GRID field used as a scoring function for docking, an improvement of about 20% in pose generation and ranking was observed.

3.7 Application of Virtual Screening. – Forino *et al.*⁵⁹ present an interesting case study of a difficult target, the protein kinase PKB/Akt, notorious for yielding a mere 2 hits in a HTS campaign. In several schemes that rely solely on docking or consensus scores, the authors report near random hit rates. However, when the final selection was based on a visual inspection of consensus hits for the potential to form hydrogen bonds similar to ATP, the hit rate increased significantly, leading to the identification of 3 μM competitive inhibitors. Although anecdotal, this story may offer comfort to modellers that brains can easily be as productive as brute force methods. Similarly, Huang *et al.*⁶⁰ were able to find hits for β -secretase after disappointing results from HTS.

Mozziconacci *et al.*⁶¹ used Cox-2 as their test case. The first part of the paper looks at the selection of the optimal parameters for the docking protocol (here using DOCK), followed by a consensus scoring approach. Having optimised the protocol with known ligands, a large (13,711) virtual library was screened. Of the 12 compounds selected and available for assay, 4 had IC₅₀'s < 1 μM.

4 *De Novo*, Inverse QSAR and Automated Iterative Design

There is continuing activity in the area of *de novo* design, driven partly by the increased interest in fragment-based screening (FBS). FBS is an experimental method for identifying small (<250 Da) molecules that bind to pockets of an active site, rather than to the site as a whole. Then the fragments should be joined into larger composite structures, with hopefully a large gain in affinity. This is the traditional territory of *de novo* design. Schnieder has reviewed the whole area since its inception.⁶² SPROUT has been used to design NK2 antagonists⁶³ based on a GPCR model. The best structure had an affinity of 2 μM as the racemate. In a less ambitious use of *de novo* design, some D3 agonists were designed based on a CoMFA model.⁶⁴

As new compounds in corporate pipelines gravitate towards higher molecular weight and ClogP, it is interesting to see what small building blocks may have been missed in the vastness of chemical space. Fink *et al.*⁶⁵ report their findings from virtual database of 14 million compounds weighing less than 160 Da. The exhaustive enumeration of all possible molecules containing C, N, O, H and F was achieved by a mathematical graph representation of the saturated hydrocarbons, followed by permutation of each core. Connectivity criteria were used to obtain a comparable composition and number of basic cores as present in the 36,000 known compounds in public data bases. Not surprisingly, the authors report a denser coverage of the property space of drug-likeness descriptors. In a virtual screening of three representative targets, a mere 10% of the virtual hits are outside the property space covered by existing compounds. Although some of the example structures do not seem desirable from a medicinal chemistry perspective, there are surprising examples of drug-like small molecules not known in any data bases.

The bridge between inverse QSAR and *de novo* design is neatly illustrated by the work of two groups. The CoG program by Brown *et al.*⁶⁶ uses a genetic algorithm to evolve similar molecules to starting structures, with fingerprints as the internal definition, and some QSAR models as the external definition of similarity.⁶⁷ The molecules are evolved by simple graph operations within a genetic algorithm framework, to change element types, valency and bond orders. The fitness of the new structures is calculated using Tanimoto similarity to a reference set of molecules. The ranking of the molecules is performed by a Pareto score based on the similarities to all the reference structures, to avoid the generation of highly localized islands around the reference set. In the example given, menthol and camphor were the reference molecules, and the method was able to produce a large number of sensible structures that were intermediate between the two. The experiment was repeated with aqueous solubility as the

target, and again the program could evolve molecules with the desired characteristics. In a related approach, Lewis⁶⁸ developed a full inverse QSAR protocol, using the mutation of structures to drive towards compounds with better fitness. The key findings of the research were that the palette of reactions used had a strong influence on the quality and improvements found, and that one had to take great care that the fitness function could be applied to the molecules that were generated. Restricting the reactions to functional groups found within the set of molecules used to generate the QSAR model helped greatly, as did imposing a core structure as a constraint. To prevent extrapolation, molecules were kept with the QSAR space using distance to the nearest neighbour in the training set as a strategy. The task of computing extrapolation has been taken up by others, as is discussed elsewhere in this review.¹⁰ In two studies with real QSAR sets, Lewis was able to propose molecules with 1–2 fold improvement in predicted activity, and that were similar to the original series.

5 3D-QSAR

The long awaited validation study for the XED method for molecular similarity has been published.⁶⁹ Molecules are described using the maxima and minima in the electrostatic and steric fields around the molecule. These points form a pharmacophore, and so can be used to search databases for alternative chemotypes. As the representation is sparse, several conformations per molecule can be considered. A new chemotype with nanomolar potency for CCK2 and improved excretion properties was found. This is a good concrete example of the power of the XED representation. Other than that, this area has been comparatively quiet, awaiting more developments in alignment methods, as discussed in the next section (Table 1).

6 Pharmacophores

Questions around the quality of our methods for generating pharmacophores are being raised, especially as there have been no major advances since GASP and DISCOtech. Three groups are revisiting some of the fundamental issues

Table 1 *Some representative high-quality 3D-QSAR models*

<i>Target</i>	<i>Method</i>	<i>Alignment</i>	<i>Q²</i>
Sigma-1 ⁷⁰	CoMFA	DISCOtech	0.7
COX-2 ⁷¹	COMFA/CoMSIA	Docking	0.74
EGFR ⁷²	CoMFA	Docking	0.7
Choline acetyltransferase ⁷³	CoMFA	Reference ligand	0.76
Oxytocin ⁷⁴	CoMFA	Docking	0.85
Catechol-O-methyltransferase ⁷⁵	CoMFA/GOLPE	Docking	0.6
Androgen receptor ⁷⁶	CoMSIA	Docking	0.66
δ , μ , κ -Opioids ⁷⁷	CoMFA	Reference ligand	0.67
NMDA ⁷⁸	CoMFA	Reference ligand	0.5
PPAR α /PPAR γ ⁷⁹	CoMFA	Reference ligand	0.7
PEPT1 ⁸⁰	CoMSIA	Reference ligand	0.82

Table 2 Some high quality pharmacophore models published during the review period

Target	Method	Features	Tolerances/Ang
HIV-1 Integrase ^{83–86}	MD/LigBuilder	4	1
General kinases ⁸⁷	MOE	5	n/a
Biogenic amine GPCRs ⁸⁸	Catalyst	3–5	n/a
GABA-A ⁸⁹	Catalyst	5 + 3 exclusion volumes	n/a

around sampling of conformational space, the generation of ensembles of solutions, and the scoring of those solutions. The Sheffield group⁸¹ have developed a multi-objective genetic algorithm (MOGA), based on their experiences with GASP. The conflicting objectives are conformational energy and the degree of overlap/similarity of the structures when overlaid according to the pharmacophore hypothesis. One inherent difficulty is that the ‘correct’ answer is often not known. Most methods can produce several plausible solutions but this may reflect the difficulty the programs have in sampling the search space. The MOGA does find a wider range of solutions than other stochastic approaches. Another advantage is that conformational space is sampled on the fly, rather than relying on a precomputed set of conformers, which will bias the search space. The disadvantage is that the MOGA does not allow for partial matches, so the pharmacophore needs to be built from compounds that all have (similar) high affinity.

Kristam *et al.*⁸² looked at the approaches used in the literature to generate Catalyst pharmacophore models. In many cases, they found it hard to reproduce the results, leading to them to propose a template for describing Catalyst models which should aid in reproducibility. In addition, they found that the less-expensive rules-based methods for generating conformational ensembles did just as well as much more expensive methods, in line with the earlier findings of Bostrom (Table 2).⁸³

7 Library Design

Multi-objective optimisation has again been linked to combinatorial chemistry. In this work,⁹⁰ the chemistry space is described by reaction transforms that contain information both about the reaction and the functional groups that the reaction will work for. The fitness function is a combination of basic bioavailability together with some empirical SAR. These reaction schemes are used to operate on a family of starting structures via a Genetic Algorithm. The number of reactions steps is limited to a small number (3). In the case considered, kinase inhibitors, basic kinase scaffolds could be rapidly decorated and their predicted affinity improved according to this protocol. Validation of the approach was performed by comparison to known kinase inhibitors rather than experimentally. In a similar vein, Brauer *et al.*⁹¹ used evolutionary chemistry to find novel inhibitors of glucose-6-phosphate translocase. They start with a selection of

low-affinity compounds from screening. After each round of optimisation, compounds are synthesised and assayed, and the authors claim the discovery of a new class of imidazoles from their work.

8 Cheminformatics and Data Mining

It is a pleasure to cite one review⁹² on the whole area of similarity and its application to structure retrieval. Prof. Willett has been a leading figure in the area, and the review marks the award of the 2005 ACS award for computers in Chemical and Pharmaceutical research. A nice application of the similarity principal based on some of his work is provided by Krumrine, Maynard and Lerman⁹³ for the analysis of virtual screening data. In any type of screening, there is a balance between the rigour/throughput of the assay and the number of acceptable false positives and negatives. Most scoring/assay systems can provide some degree of discrimination between hits and inactives, but the overlap between the two classes can be quite significant. In this approach, each hit/inactive is characterised by the behaviour of neighbouring compounds. These compounds should be similar and there exhibit similar behaviour. A false positive is therefore a hit surrounded by inactives, and vice versa for a false negative. The neighbourhood is defined by a threshold of 70% similarity, as measured by Daylight fingerprints and a Tanimoto coefficient. The probability of a compound being truly active/inactive, given its neighbourhood, can then be computed. This enables more effective selection of compounds for further assay. As a sidebar, the authors note that early application of common filters for bioavailability and ADMET would have resulted in the loss of a number of perfectly reasonable lead series. Caveat emptor!

8.1 Scaffold Hopping. – Scaffold hoping is an attractive road to novel chemotypes, yet there are few routine methods that reliably deliver. An interesting case study is presented by Rush *et al.*⁹⁴ which uses the overlap of shape Gaussians to search for chemically diverse analogues of a μM HTS hit in the conformationally expanded data base. To weed out steric clashes, a MM-PBSA scoring function was applied. Among the 30 candidates, 3 hits with comparable binding affinities and better overall-properties were identified. As the 2D-similarity of the hits was rather low, the authors claim that shape alone can be a more powerful descriptor for scaffold-hopping. Another explanation why the method works well in this case is that the molecules are flat and conformationally restricted. It would be interesting to see how this method fares across various target classes. The XED approach for scaffold hopping is discussed above.⁶⁹ Rarey *et al.*⁹⁵ have extended their feature tree algorithm by devising a scheme to combine single trees into a multiple tree. In a sense, this is the same as deducing a pharmacophore model using just 2D topology. They showed that the multiple trees outperformed single trees and Catalyst models in terms of enrichment during virtual screening, and that alternative scaffolds

could also be retrieved. The study however only considered two cases, ACE inhibitors and $\alpha 1$ antagonists.

8.2 Descriptors and Atom Typing. – Bender *et al.*⁹⁶ have proposed a new descriptor. Molecules are described using atom environments up to an optimal radius (found to be 2). The most information-rich features are then selected by trying to optimise the separation between active and inactive structures, and these features are used to drive a Bayesian classification algorithm. In the tests described, these targeted features seem to do better than analogous descriptors (*e.g.* MACCS keys or Daylight fingerprints) that are unfiltered, and have equivalent performance to other directed feature algorithms like feature trees.⁹⁷ The enrichment factors are between 6.5 and 11 fold.

Labute⁹⁸ presents a new algorithm for the automated prediction of atom types from structural data, that is based on a maximum weighted matching rather than knowledge-based rules. From the coordinates, a covariance-matrix is used to determine candidate bonding partners. Based on these connectivities and geometric thresholds, a preliminary hybridization is assigned. The final assignment uses the likelihood rates for higher bond orders estimated from a set of 200k commercial compounds. Over a range of test sets from the PDB, the success rates are reported to be in the low 90 percentiles. Poor perception was encountered in the case of strained, unusual conformations, underlining that visual inspection by chemist remains the gold standard. In the advent of high-throughput crystallography, where the interpretation of results might eventually become the scarcer resource, such a method might prove useful.

A fundamental question for most medicinal chemists, is “what makes a compound attractive”. This is particularly important when lists of hits from high-throughput screening are being assessed and the series for optimisation is being chosen. Lajiness *et al.*⁹⁹ set out to study how subjective the “chemist’s eye” really was by asking a set of chemists to independently score the same set of 250 compounds that had already been rejected by a senior and experienced medicinal chemists. The result was that the chemists rejected compounds on a very inconsistent basis, which has serious implications for hit follow-up studies as an entire chemotype can be dropped for no good reason. Even when the same chemist looked at the same compounds for a second time, the pattern of rejections varied greatly. There was also no relation between experience and consistency of opinion; assessment of attractiveness cannot be learned. Overall the consistency of opinion between two chemists was 24%.

The use of biological fingerprints is part of the emerging area of chemogenomics. A seminal paper in the field is on kinomics.¹⁰⁰ Rather than classifying kinases according to sequence alignment or other bioinformatics methods, the authors propose a scheme based entirely on small molecule selectivity data. This enables one to perform structure-based drug design on kinases that are otherwise seeming unrelated. Naturally, at high levels of homology, the results of the two approaches are the same. When faced with a new kinase target or selectivity profile, one could deduce similar targets and hence screening sets.

Conversely, given a new chemotype, one could also predict which kinases it is likely to hit.

8.3 Tools. – The emergence of several tools focussed on the medicinal chemist has been welcome. In most cases, these tools have been built inside Pharma companies to deal with a pressing need. The selection of bioisoteric heterocycles¹⁰¹ is a good example. There are many possible heterocycles, and the ones favoured by this program are derived from molecules in Phase II clinical trials or later. This should filter out systems associated with poor toxicity, metabolism and the like. The heterocycles are oriented to a common frame of reference, and relevant descriptors, such as charge, shape size *etc* are computed. The most similar systems to a query ring system can be found and presented quickly to the chemist to seed further ideas for lead optimisation. Although it is not really within the scope of this review, the discovery of ferrocene as an effective bioisostere for benzene in GPCR ligands¹⁰² is included as an example of how far we still have to go in understanding similarity.

9 Structure-Based Drug Design

9.1 Analysis of Active Sites and Target Tractability. – Hajduk *et al.*¹⁰³ have looked at a large set of NMR-based screening data to assess the important factors that make a protein pocket druggable. Among a diverse group of 23 targets, the authors identify 57 potential binding sites of which only 28 are targeted by small molecules. The highly consistent hit rate of >90% for known binding sites underscores the general wisdom of targeting substrate binding sites. The single most important factor was found to be the apolar surface area, followed by the size and roughness of the pocket. The authors also present an eight-parameter correlation function which may serve to prioritize among several potential targets. The ability to predict the druggability of a potential protein target will certainly be much appreciated when the genomic era of drug design comes to full fruition.

The hardy problem of protein flexibility has been re-examined using B-factors.¹⁰⁴ Despite the known dangers in equating high B-factors with intrinsic mobility, the authors conclude that, in a study of 800 high-resolution x-ray structures, 71% of atoms in the binding sites become less mobile on binding, but that 29% become more so. Explanation of this observation is not easy, and may reflect loss of water, or induced fit, freeing up side chains. The prediction of binding site flexibility on a fairly extensive test set of both NMR and x-ray structures has been studied by Mancera *et al.*¹⁰⁵ Their algorithm DYNASITE uses an iterative procedure of rotamer generation based on a rotamer library and minimization of cluster representatives to identify alternative conformations of binding site side chains. This method obtains up to 20 clusters of alternative conformations and in all but two cases the experimental alternative conformations were among those. Exceptions were ligands which induced a high-energy conformation of the active site. While the authors have

demonstrated that a rotamer-approach reliably generates proximal conformations, the prediction of which conformation prevails for a given ligand in an unknown complex places a heavy burden on the scoring function.

9.2 Kinase Modelling. – Selectivity in kinase inhibition is a topic that is bound to gain importance as more and more kinase targets are discovered. Sheinerman *et al.*¹⁰⁶ describe a 3-step approach to distinguish and predict selectivity profiles. After a structure-guided alignment of representative kinases from all families, they use continuum electrostatics to identify the active site residues that contribute most to binding. The spatial arrangement of the important residues then constitutes what the authors term the “binding site signature”. Substitutions are classified as non-conservative or conservative based on polarity and volume of the residue. In some cases, this crude method works surprisingly well, *e.g.*, kinases of the ABL-family that are inhibited by Imatinib in the sub- μM range are distinguished from non-targeted families ($\text{IC}_{50} > 10 \mu\text{M}$) based on only one non-conservative substitution. In some cases, kinases with conserved signature are not inhibited, prompting the authors to concede that conservation of the signature is a necessary but not sufficient criterion. While one non-conservative substitution is not sufficient in all cases to predict whether a drug is active on a given kinase, the correlation for two non-conservative substitutions is very good. By comparison, sequence similarity appears to be a very poor criterion. In this new field of kinomics,¹⁰⁰ approaches like this which combine a simple methodology with high predictive power are very welcome to digest the steadily growing amount of selectivity data.

Comparison of bindings sites, to examine questions of selectivity, has been tried with alignment-free approaches, for example the work of Vulpetti *et al.*¹⁰⁷ on kinases using GRID/CPCA.¹⁰⁸ By analysing the molecular interaction fields, they were able to identify key residues at the back of the ATP-binding pocket. These differences were exploited to design compounds that were both potent and selective. The advantage of this approach is that it is mostly automatic, and only needs unaligned models of the proteins. In case where there structural data is not available, Bayesian models¹⁰⁹ have been built from experimental inhibition data. The models do show some enrichment, and their speed means that they could be used for screening vendor libraries; however, it is still too early to expect the same power as the structure-based methods.

9.3 GPCR Modelling. – Despite the lack of further structures of GPCRs, modelling of this protein family is still an active area. Rognan *et al.*¹¹⁰ have published an automated protocol for building models of the seven transmembrane domains. They have created a database of 277 human GPCRs from the secretin, rhodopsin and mGlutamate classes. The most interesting part of the approach is the use of docking to identify good and bad models. For a set of known ligands and decoys, good models can pick out the known ligands from the decoys: bad models can also be identified and refined further. The authors also note that the models can be used as a screen to identify possible receptors for ligands of unknown selectivity. In a study of the power of different

techniques for identifying leads for GPCRs, Evers *et al.*¹¹¹ showed that pharmacophore, 2D-QSARs and Feature Trees did better than classical docking experiments, when this extra information was available. However, they conclude that docking to GPCRs still gave good levels of enrichment, so that this is still a valid method for hit identification for the target class.

A general QSAR model for GPCRs has also been developed.¹¹² The activity data was taken from 1939 compounds tested over 40 GPCRs, and the QSAR was built using pharmacophore-based descriptors. Experimental validation of 360 compounds against 21 targets showed that the model gave > 5-fold enrichment even at low threshold for the definition of activity. Enrichment was also observed for targets not in the original assay panel, so that ligands could be designed for orphan receptors based on the data from related GPCRs. Bock and Gough also describe a GPCR QSAR model based on very simple descriptors and the protein sequence.¹¹³ Kratochwil *et al.*¹¹⁴ have devised a method for abstracting the active site residues from GPCRs into a simple pharmacophore-like model. This representation is useful for comparison of active sites, the interpretation of mutagenesis data and the early assessment of selectivity issues. Lamb *et al.*¹¹⁵ claim a pharmacophore-based model for GPCR screening, with a hit-rate of 2.6% against the μ -opioid receptor ($IC_{50} < 10 \mu M$). The model was derived from activity classes in the MDDR database, and 3- and 4-centre pharmacophores were used. Screening of virtual libraries is then performed by looking for subsets that overall cover the same pharmacophoric signature as the set of GPCR ligands.

Another study into classification of GPCR ligands used support vector machines and pharmacophore fingerprints.¹¹⁶ The model was very good at identifying inactives (<99%) and could recall actives at 75%. To provide a stringent test (the space of GPCR ligands is relatively crowded, and there will often be a similar ligand left in the training set for leave-N-out validations), an entire chemotype class was removed. By refinding members of this class, for which the nearest neighbour in the training set had a similarity of 0.38 by Daylight fingerprints, the model showed that it could also be used for chemotype hopping.

Bridging the fields of structure-based drug design and chemogenomics is protein function and similarity. Given a query binding site, can one find proteins with similar function, or given a protein, can one identify the active site(s) by analogy with similar proteins. In addition to the standard CavBase,¹¹⁷ which can be slow to run, Nussinov *et al.*¹¹⁸ have developed a much faster method, based in essence on pharmacophore triplets. As always, speed is at the cost of resolution, and the pharmacophore features are crude, and not influenced by any local electrostatic perturbations. Also the representation is a snapshot of a single conformation. In tests, adenine and estradiol binding sites were accurately recognised, as could analogues of fatty-acid binding protein. A sequence-based tool has been developed by Lichtarge *et al.*,¹¹⁹ based on their evolutionary trace algorithm. A set of analogues to the query protein are retrieved using BLAST, then aligned using ClustalW. The residues in the sequence are given a score depending on the degree of conservation: the

number of branches the alignment must be split into before the residue becomes invariant within the branch. Clusters of residues of similar rank indicate some evolutionary focussing *i.e.* an active site. In their hands, the best-ranked residues cluster around active sites, and can be shown to occur at highly significant probabilities over chance. Using this measure they were able to identify 70–90% of active sites from sequence alone. This provides two good tools to complement biological fingerprinting.

10 Conclusions

In the last review period, we highlighted the advances in the field of ADME/Tox and docking. This trend has been sustained. One can speculate that collaborations driven by confidential in-house data have provided the impetus for the fresh attempts on some very hardy problems. Virtual screening in all its forms is at the heart of most research, from bioavailability filters through to rigorous estimations of the free energy of binding. It is pleasing that the field of *de novo* design is undergoing a small renaissance, as we understand how to control the changes made to molecules more carefully. We can also see the development of target-specific approaches becoming more popular over generalised solutions, so that local knowledge can be properly used, rather than washed out in the noise. Finally, we predict that chemogenomics will become a more common-place and accessible tool, allowing all modellers to mine our wealth of experimental data more fully.

References

1. M. Vieth, M.G. Siegel, R.E. Higgs, I.A. Watson, D.H. Robertson, K.A. Savin, G.L. Durst and P.A. Hipskind, *J. Med. Chem.*, 2004, **47**, 224.
2. P.D. Leeson and A.M. Davis, *J. Med. Chem.*, 2004, **47**, 6338.
3. J.J. Lu, K. Crimin, J.T. Goodwin, P. Crivori, C. Orrenius, L. Xing, P.J. Tandler, T.J. Vidmar, B.M. Amore, A.G.E. Wilson, P.F.W. Stouten and P.S. Burton, *J. Med. Chem.*, 2004, **47**, 6104.
4. D.F. Veber, S.R. Johnson, H.Y. Cheng, B.R. Smith, K.W. Ward and K.D. Kopple, *J. Med. Chem.*, 2002, **45**, 2615.
5. Y.C. Martin, *J. Med. Chem.*, 2005, **48**, 3164.
6. J.K. Wegner, H. Frohlich and A. Zell, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 931.
7. C. de Graaf, N.P.E. Vermeulen and K.A. Feenstra, *J. Med. Chem.*, 2005, **48**, 2725.
8. S.E. O'Brien and M.J. de Groot, *J. Med. Chem.*, 2005, **48**, 1287.
9. K.K. Chohan, S.W. Paine, J. Mistry, P. Barton and A.M. Davis, *J. Med. Chem.*, 2005, **48**, 5154.
10. R. Guha and P.C. Jurs, *J. Chem. Inf. Model.*, 2005, **45**, 65.
11. R.P. Sheridan, B.P. Feuston, V.N. Maiorov and S.K. Kearsley, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 1912.
12. S.B. Singh, L.Q. Shen, M.J. Walker and R.P. Sheridan, *J. Med. Chem.*, 2003, **46**, 1330.
13. G. Cruciani, E. Carosati, B. DeBoeck, K. Ethirajulu, C. Mackie, T. Howe and R. Vianello, *J. Med. Chem.*, 2005, **48**, 6970.

14. J.M. Kriegl, T. Arnhold, B. Beck and T. Fox, *J. Comput. -Aided Mol. Des.*, 2005, **19**, 189.
15. M.E. Beck, *J. Chem. Inf. Model.*, 2005, **45**, 273.
16. S.P. deVisser, D. Kumar, S. Cohen, R. Shacham and S. Shaik, *J. Am. Chem. Soc.*, 2004, **126**, 8362.
17. J.F. Contrera, E.J. Matthews, N.L. Kruhlik and R.D. Benz, *Reg. Tox. Pharm.*, 2004, **40**, 185.
18. C. Helma, *Curr. Op. Drug Disc. Dev.*, 2005, **8**, 27.
19. S. Dimitrov, G. Dimitrova, T. Pavlov, N. Dimitrova, G. Patlewicz, J. Niemela and O. Mekenyan, *J. Chem. Inf. Model.*, 2005, **45**, 839.
20. C.M. Krejsa, D. Horvath, S.L. Rogalski, J.E. Penzotti, B. Mao, F. Barbosa and J.C. Migeon, *Curr. Op. Drug Disc. Dev.*, 2003, **6**, 470.
21. A.J.S. Knox, M.J. Meegan, G. Carta and D.G. Lloyd, *J. Chem. Inf. Model.*, 2005, **45**, 1908.
22. J.J. Irwin and B.K. Shoichet, *J. Chem. Inf. Model.*, 2005, **45**, 177.
23. M. Sun, Z.S. Li, Y. Zhang, Q.C. Zheng and C.C. Sun, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 2851.
24. S.L. McGovern and B.K. Shoichet, *J. Med. Chem.*, 2003, **46**, 2895.
25. C. Bissantz, P. Bernard, M. Hibert and D. Rognan, *Proteins*, 2003, **50**, 5.
26. J.A. Erickson, M. Jalaie, D.H. Robertson, R.A. Lewis and M. Vieth, *J. Med. Chem.*, 2004, **47**, 45.
27. A.M. Ferrari, B.Q.Q. Wei, L. Costantino and B.K. Shoichet, *J. Med. Chem.*, 2004, **47**, 5076.
28. T. Polgar and G.M. Keseru, *J. Med. Chem.*, 2005, **48**, 3749.
29. H.A. Carlson, *Current Opinion in Chemical Biology*, 2003, **6**, 01.
30. X. Barril and S.D. Morley, *J. Med. Chem.*, 2005, **48**, 4432.
31. J.E. Ladbury, *Chemistry & Biology*, 1996, **3**, 973.
32. M.L. Verdonk, G. Chessari, J.C. Cole, M.J. Hartshorn, C.W. Murray, J.W.M. Nissink, R.D. Taylor and R. Taylor, *J. Med. Chem.*, 2005, **48**, 6504.
33. C. de Graaf, P. Pospisil, W. Pos, G. Folkers and N.P.E. Vermeulen, *J. Med. Chem.*, 2005, **48**, 2308.
34. M.D. Cummings, R.L. DesJarlais, A.C. Gibbs, V. Mohan and E.P. Jaeger, *J. Med. Chem.*, 2005, **48**, 962.
35. J.C. Cole, C.W. Murray, J.W.M. Nissink, R.D. Taylor and R. Taylor, *Proteins*, 2005, **60**, 325.
36. M.L. Verdonk, V. Berdini, M.J. Hartshorn, W.T.M. Mooij, C.W. Murray, R.D. Taylor and P. Watson, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 793.
37. G.L. Warren, C.W. Andrews, A.M. Capelli, B. Clarke, J. LaLonde, M.H. Lambert, M. Lindvall, N. Nevins, S.F. Semus, S. Senger, G. Tedesco, I.D. Wall, J.M. Woolven, C.E. Peishoff and M.S. Head, *J. Med. Chem.*, 2005, dx.doi.org/10.1021/jm050362n.
38. A.P. Graves, R. Brenk and B.K. Shoichet, *J. Med. Chem.*, 2005, **48**, 3714.
39. H.F.G. Velec, H. Gohlke and G. Klebe, *J. Med. Chem.*, 2005, **48**, 6296.
40. C. bad-Zapatero and J.T. Metz, *Drug Discovery Today*, 2005, **10**, 464.
41. Y. Pan, N. Huang, S. Cho and A.D. MacKerell Jr., *J. Chem. Inf. Comp. Sci.*, 2003, **43**, 267.
42. W.M. Rockey and A.H. Elcock, *J. Med. Chem.*, 2005, **48**, 4138.
43. C. Oostenbrink and W.F. van Gunsteren, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 6750.
44. C. Oostenbrink and W.F. van Gunsteren, *Proteins*, 2004, **54**, 237.

45. C. Oostenbrink and W.F. van Gunsteren, *J. Comput. Chem.*, 2003, **24**, 1730.
46. C. Oostenbrink, A. Villa, A.E. Mark and W.F. van Gunsteren, *J. Comput. Chem.*, 2004, **25**, 1656.
47. J.M. Yang, Y.F. Chen, T.W. Shen, B.S. Kristal and D.F. Hsu, *J. Chem. Inf. Model.*, 2005, **45**, 1134.
48. G.P.A. Vigers and J.P. Rizzi, *J. Med. Chem.*, 2004, **47**, 80.
49. R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis and P.S. Shenkin, *J. Med. Chem.*, 2004, **47**, 1739.
50. S. Radestock, M. Bohm and H. Gohlke, *J. Med. Chem.*, 2005, **48**, 5466.
51. A.E. Muryshev, D.N. Tarasov, A.V. Butygin, O.Y. Butygina, A.B. Aleksandrov and S.M. Nikitin, *J. Comput. -Aided Mol. Des.*, 2003, **17**, 597.
52. A. Krammer, P.D. Kirchhoff, X. Jiang, C.M. Venkatachalam and M. Waldman, *J. Mol. Graph. Mod.*, 2005, **23**, 395.
53. A. Curioni, T. Mordasini and W. Andreoni, *J. Comput. -Aided Mol. Des.*, 2004, **18**, 773.
54. P. Ferrara, A. Curioni, E. Vangrevelinghe, T. Meyer, T. Mordasini, W. Andreoni, P. Acklin and E. Jacoby, *J. Chem. Inf. Model.*, 2006, **46**, 254.
55. B. Kuhn, P. Gerber, T. Schulz-Gasch and M. Stahl, *J. Med. Chem.*, 2005, **48**, 4040.
56. N. Huang, C. Kalyanaraman, J.J. Irwin and M.P. Jacobson, *J. Chem. Inf. Model.*, 2006, **46**, 243.
57. V. Maiorov and R.P. Sheridan, *J. Chem. Inf. Model.*, 2005, **45**, 1017.
58. E. Carosati, S. Sciabola and G. Cruciani, *J. Med. Chem.*, 2004, **47**, 5114.
59. M. Forino, D. Jung, J.B. Easton, P.J. Houghton and M. Pellicchia, *J. Med. Chem.*, 2005, **48**, 2278.
60. D.Z. Huang, U. Luthi, P. Kolb, K. Edler, M. Cecchini, S. Audetat, A. Barberis and A. Caffisch, *J. Med. Chem.*, 2005, **48**, 5108.
61. J.C. Mozziconacci, E. Arnoult, P. Bernard, Q.T. Do, C. Marot and L. Morin-Allory, *J. Med. Chem.*, 2005, **48**, 1055.
62. G. Schneider and U. Fechner, *Nat. Rev. Drug Disc.*, 2005, **4**, 649.
63. M.A. Ali, N. Bhogal, J.B.C. Findlay and C.W.G. Fishwick, *J. Med. Chem.*, 2005, **48**, 5655.
64. J. Elsner, F. Boeckler, F.W. Heinemann, H. Hubner and P. Gmeiner, *J. Med. Chem.*, 2005, **48**, 5771.
65. T. Fink, H. Bruggesser and J.L. Reymond, *Ang. Chem. Int. Ed*, 2005, **44**, 1504.
66. N. Brown, B. McKay and J. Gasteiger, *J. Comput. -Aided Mol. Des.*, 2004, **18**, 761.
67. N. Brown, B. McKay, F. Gilardoni and J. Gasteiger, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 1079.
68. R.A. Lewis, *J. Med. Chem.*, 2005, **48**, 1638.
69. C.M.R. Low, I.M. Buck, T. Cooke, J.R. Cushnir, S.B. Kalindjian, A. Kotecha, M.J. Pether, N.P. Shankley, J.G. Vinter and L. Wright, *J. Med. Chem.*, 2005, **48**, 6790.
70. D. Jung, J. Floyd and T.M. Gund, *J. Comp. Chem.*, 2004, **25**, 1385.
71. P.A. Datar and E.C. Coutinho, *J. Mol. Graph. Mod.*, 2004, **23**, 239.
72. C.L. Kuo, H. Assefa, S. Kamath, Z. Brzozowski, J. Slawinski, F. Saczewski, J.K. Buolamwini and N. Neamati, *J. Med. Chem.*, 2004, **47**, 385.
73. V. Chandrasekaran, G.B. McGaughey, C.J. Cavallito and J.P. Bowen, *J. Mol. Graph. Model.*, 2004, **23**, 69.
74. B. Jojart, T.A. Martinek and A. Marki, *J. Comput. -Aided Mol. Des.*, 2005, **19**, 341.
75. A.J. Tervo, T.H. Nyronen, T. Ronkko and A. Poso, *J. Comput. -Aided Mol. Des.*, 2003, **17**, 797.

76. A.A. Soderholm, P.T. Lehtovuori and T.H. Nyronen, *J. Med. Chem.*, 2005, **48**, 917.
77. Y.Y. Peng, S.M. Keenan, Q. Zhang, V. Kholodovych and W.J. Welsh, *J. Med. Chem.*, 2005, **48**, 1620.
78. B.M. Baron, R.J. Cregge, R.A. Farr, D. Friedrich, R.S. Gross, B.L. Harrison, D.A. Janowick, D. Matthews, T.C. McCloskey, S. Meikrantz, P.L. Nyce, R. Vaz and W.A. Metz, *J. Med. Chem.*, 2005, **48**, 995.
79. S. Khanna, M.E. Sobhia and P.V. Bharatam, *J. Med. Chem.*, 2005, **48**, 3015.
80. A. Biegel, S. Gebauer, B. Hartrodt, M. Brandsch, K. Neubert and I. Thondorf, *J. Med. Chem.*, 2005, **48**, 4410.
81. S.J. Cottrell, V.J. Gillet, R. Taylor and D.J. Wilton, *J. Comput. -Aided Mol. Des.*, 2004, **18**, 665.
82. R. Kristam, V.J. Gillet, R.A. Lewis and D. Thorner, *J. Chem. Inf. Model.*, 2005, **45**, 461.
83. J. Bostrom, J.R. Greenwood and J. Gottfries, *J. Mol. Graph. Model.*, 2003, **21**, 449.
84. G.L. Mustata, A. Brigo and J.M. Briggs, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 1447.
85. J.X. Deng, K.W. Lee, T. Sanchez, M. Cui, N. Neamati and J.M. Briggs, *J. Med. Chem.*, 2005, **48**, 1496.
86. A. Brigo, K.W. Lee, F. Fogolari, G.L. Mustata and J.M. Briggs, *Proteins*, 2005, **59**, 723.
87. A.M. Aronov and M.A. Murcko, *J. Med. Chem.*, 2004, **47**, 5616.
88. T. Klabunde and A. Evers, *Chembiochem.*, 2005, **6**, 876.
89. P. Kahnberg, M.H. Howard, T. Liljefors, M. Nielsen, E.O. Nielsen, O. Sterner and I. Pettersson, *J. Mol. Graph. Model.*, 2004, **23**, 253.
90. S.C. Schurer, P. Tyagi and S.A. Muskal, *J. Chem. Inf. Model.*, 2005, **45**, 239.
91. S. Brauer, M. Almstetter, W. Antuch, D. Behnke, R. Taube, P. Furer and S. Hess, *J. Comb. Chem.*, 2005, **7**, 218.
92. P. Willett, *J. Med. Chem.*, 2005, **48**, 4183.
93. J.R. Krumrine, A.T. Maynard and C.L. Lerman, *J. Med. Chem.*, 2005, **48**, 7477.
94. T.S. Rush, J.A. Grant, L. Mosyak and A. Nicholls, *J. Med. Chem.*, 2005, **48**, 1489.
95. G. Hessler, M. Zimmermann, H. Matter, A. Evers, T. Naumann, T. Lengauer and M. Rarey, *J. Med. Chem.*, 2005, **48**, 6575.
96. A. Bender, H.Y. Mussa, R.C. Glen and S. Reiling, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 170.
97. M. Rarey and J.S. Dixon, *J. Comput. -Aided Mol. Des.*, 1998, **12**, 471.
98. P. Labute, *J. Chem. Inf. Model.*, 2005, **45**, 215.
99. M.S. Lajiness, G.M. Maggiora and V. Shanmugasundaram, *J. Med. Chem.*, 2004, **47**, 4891.
100. M. Vieth, R.E. Higgs, D.H. Robertson, M. Shapiro, E.A. Gragg and H. Hemmerle, *Bio. Biophys. Acta.*, 2004, **1697**, 243.
101. H.B. Broughton and I.A. Watson, *J. Mol. Graph. Model.*, 2004, **23**, 51.
102. K. Schlotter, F. Boeckler, H. Hubner and P. Gmeiner, *J. Med. Chem.*, 2005, **48**, 3696.
103. P.J. Hajduk, J.R. Huth and S.W. Fesik, *J. Med. Chem.*, 2005, **48**, 2518.
104. C.Y. Yang, R.X. Wang and S.M. Wang, *J. Med. Chem.*, 2005, **48**, 5648.
105. A.Y.C. Yang, P. Kallblad and R.L. Mancera, *J. Comput. -Aided Mol. Des.*, 2004, **18**, 235.
106. F.B. Sheinerman, E. Giraud and A. Laoui, *J. Mol. Biol.*, 2005, **352**, 1134.
107. A. Vulpetti, P. Crivori, A. Cameron, J. Bertrand, M.G. Brasca, R. D'Alessio and P. Pevarello, *J. Chem. Inf. Model.*, 2005, **45**, 1282.

108. M.A. Kastenholz, M. Pastor, G. Cruciani, E.E. Haaksmá and T. Fox, *J. Med. Chem.*, 2000, **43**, 3033.
109. X.Y. Xia, E.G. Maliski, P. Gallant and D. Rogers, *J. Med. Chem.*, 2004, **47**, 4463.
110. C. Bissantz, A. Logean and D. Rognan, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 1162.
111. A. Evers, G. Hessler, H. Matter and T. Klabunde, *J. Med. Chem.*, 2005, **48**, 5448.
112. C. Rolland, R. Gozalbes, E. Nicolai, M.F. Paugam, L. Coussy, F. Barbosa, D. Horvath and F. Revah, *J. Med. Chem.*, 2005, **48**, 6563.
113. J.R. Bock and D.A. Gough, *J. Chem. Inf. Model.*, 2005, **45**, 1402.
114. N.A. Kratochwil, P. Malherbe, L. Lindemann, M. Ebeling, M.C. Hoener, A. Muhlemann, R.H.P. Porter, M. Stahl and P.R. Gerber, *J. Chem. Inf. Model.*, 2005, **45**, 1324.
115. M.L. Lamb, E.K. Bradley, G. Beaton, S.S. Bondy, A.J. Castellino, P.A. Gibbons, M.J. Suto and P.D. Grootenhuis, *J. Mol. Graph. Model.*, 2004, **23**, 15.
116. J.C. Saeh, P.D. Lyne, B.K. Takasaki and D.A. Cosgrove, *J. Chem. Inf. Model.*, 2005, **45**, 1122.
117. M. Hendlich, A. Bergner, J. Gunther and G. Klebe, *J. Mol. Biol.*, 2003, **326**, 607.
118. A. Shulman-Peleg, R. Nussinov and H.J. Wolfson, *J. Mol. Biol.*, 2004, **339**, 607.
119. H. Yao, D.M. Kristensen, I. Mihalek, M.E. Sowa, C. Shaw, M. Kimmel, L. Kavrakı and O. Lichtarge, *J. Mol. Biol.*, 2003, **326**, 255.