J Comput Aided Mol Des (2006) 20:191–208
DOI 10.1007/s10822-006-9045-3

ORIGINAL PAPER

# Electrostatic evaluation of isosteric analogues

**Roger Sayle · Anthony Nicholls**

Received: 31 October 2005 / Accepted: 28 March 2006 / Published online: 15 July 2006


**Abstract** A method is presented for enumerating a large number of isosteric analogues of a ligand from a known protein–ligand complex structure and then rapidly calculating an estimate of their binding energies. This approach takes full advantage of the observed crystal structure, by reusing the atomic co-ordinates determined experimentally for one ligand, to approximate those of similar compounds that have approximately the same shape. By assuming that compounds with similar shapes adopt similar binding poses, and that entropic and protein flexibility effects are approximately constant across such an isosteric series (''the frozen ligand approximation''), it is possible to order their binding affinities relatively accurately. Additionally, the constraint that the atomic coordinates are invariant allows for a dramatic simplification in the Poisson–Boltzmann method used to calculation the electrostatic component of the binding energy. This algorithmic improvement allows for the calculation of tens of thousands of binding energies per second for drug-like molecules, enabling this technique to be used in screening large virtual libraries of isosteric analogues. Most significantly, this procedure is shown to be able to reproduce SAR effects of subtle medicinal chemistry substitutions. Finally, this paper reports the results of the proposed methodology on seven model systems; dihydrofolate reductase, Lck kinase, ribosome inactivating protein, L-arabinose binding protein, neuraminidase, HIV-1 reverse transcriptase and COX-2.

R. Sayle (✉) · A. Nicholls
OpenEye Scientific Software, Suite 1107, 3600 Cerrillos Road, Santa Fe, New Mexico 87507, USA
e-mail: roger@eyesopen.com

**Keywords** WABE · Docking · Isosteres · Electrostatics · Frozen ligand approximation · De novo structure based drug design · Poisson–Boltzmann equation

## Introduction

Although structure-based drug design has had a number of notable successes, the residual error in current state-of-the-art scoring functions continues to prevent protein–ligand docking from displacing alternative lead discovery methods (both experimental and computational), such as combinatorial chemistry, high-throughput screening (HTS) or compound similarity. The problem is that although the main determinants of molecular recognition are known to involve shape and electrostatics, many of the energetic terms involved in protein–ligand binding are poorly understood or difficult to calculate. Current virtual library screening protocols are able to capture the principal components such as shape complementarity and hydrogen bonding interactions, but are stymied by lesser effects such as entropy, solvent pressure, protein flexibility, conformational strain and potentially as-yet undiscovered influences. The net result of this uncertainty, is that prospective docking is currently just one of a set of approaches available for drug-discovery, often relegated to ''enrichment-style'' compound prioritization, despite being significantly cheaper than experimental methods, such as high-throughput screening. Indeed, the largest benefit of docking in virtual library screening is to capture the dominant term of shape complementarity, and eliminate compounds that could not possibly fit in an active-site.

One approach to analyzing the problems with docking scoring functions is to isolate the tasks of ''posing'' and ''scoring''. Docking faces the dual challenges of posing, orienting a ligand conformation in an active-site, and scoring, estimating the free energy of interaction between the ligand and protein for the selected pose. The requirements of these tasks are sometimes conflicting. To provide accurate scoring of a pose, the interaction energy between a pair of atoms (such as a Van der Waals repulsion) may require a hard potential that changes steeply for small changes in coordinates. For utility in correctly identifying the best scoring pose, a docking program may prefer the same interaction to be represented by a smoother soft potential, that allows a minimum to be found with a larger radius of convergence [6, 14]. This dichotomy results in docking programs that either specialize in reproducing the bound conformation or correctly rank ordering ligands 22.

The experiment investigated in this paper is to evaluate just the scoring aspect of docking potential functions by only considering poses with the same three-dimensional co-ordinates. This restriction, called ''the frozen ligand approximation'', reuses the coordinates of a known protein–ligand complex to approximate and evaluate those of closely related compounds, specifically isosteric analogues. For the purposes of this investigation, an isosteric analogue is a compound with the same heavy atom connectivity, where the hybridization geometry at each non-terminal atom is consistent with the original template ligand. By this definition, the set of isosteric analogues of a compound is a superset of its tautomers and potential protonation states.

The major assumption made by the frozen ligand approximation is that molecules with similar shapes, i.e. the isosteric analogues, adopt similar conformations and binding modes in a protein active site. Further, the sidechain and backbone movements induced in the protein target, the reorganization of solvent around each complex, and even the internal strain of the ligand are likewise assumed to be similar to those in the known reference protein–ligand structure. The validity and limitations of this hypothesis, for a number of model systems, is one aspect evaluated in later sections.

The simplifying assumptions of the frozen ligand approximation provide both scientific and technical benefits. The major scientific benefit is that many of the difficult to account for terms in free energy of binding calculations are held constant (or remain approximately equal). Effects such as entropy, protein movement, co-factor and water displacement, strain energy, and cavity and buried surface area terms are cancelled out when comparing pairs of isosteric analogues. The major technical benefit is the simplification in implementation. Not only does evaluating isosteric analogues eliminate the most time consuming step of posing, but also knowing the coordinates to be evaluated in advance allows much of the computation (such as distances and surface areas) to be pre-computed. This allows evaluation of isosteric analogues to be orders of magnitude faster than conventional docking of unrelated discrete compounds, typically at a rate of thousands or tens of thousands of compounds per second on conventional computer hardware. We take advantage of this speed by analyzing the binding trends of many millions of isosteric analogues.

Inspection of the intermolecular terms used in physic-based scoring functions reveals that once the degrees of freedom associated with atomic movement are eliminated, and only the atom types are allowed to change, the difference between isosteric analogues will be predominantly electrostatic. We do not mean to imply there will not also be differences in Van der Waals interactions from atoms of different radii, differences in total buried area or potentially different rankings in conformer energies, merely that these become second order to the first order changes in electrostatic interaction and desolvation. For typical medicinal chemistry inter-substitutions of carbon, nitrogen and oxygen, the VdW radii are approximately the same, so this assumption is reasonable. For substitution of, say, fluorine to bromine, iodine or sulfur it may not be true, although we note ignoring potential VdW clashes may actually be an advantage. Small shifts in protein atoms to accommodate larger atoms are often observed in crystallography-guided drug optimization and yet are difficult to account for in traditional scoring approaches. A major component of net binding affinity is the hydrophobic effect, typically modeled by a term proportional to the total buried area. Even relatively large radii changes are likely to produce only small differences in this term.

One of the major known limitations of existing docking scoring functions, size bias, is one of the systematic errors conveniently eliminated by this experiment. Most docking scoring functions, whether empirical or molecular mechanics-based, are built upon the summation of pair-wise atomic interaction potentials. Therefore, the more atoms a ligand contains the greater its potential interaction. By restricting consideration to ligands with the same heavy atom count, we eliminate this inherent size bias. In many respects, isosteric analogues provide an ideal set of docking decoys for use in enrichment experiments that attempt to distinguish active compounds from

inactives. As will be shown, the quality of the poses assumed by this technique, in terms of RMS superposition accuracy, is also dramatically better than achievable via current ''ab initio'' posing techniques. This provides the opportunity to investigate docking potentials for a near-ideal world where posing can consistently identify poses within a fraction of an Ångstrom RMS deviation from the experimentally observed configuration.

Conceptually, this approach takes as much advantage of the available X-ray structure information as possible. In conventional docking protocols, using the co-ordinates of a protein from a known complex typically provides far better results than using the co-ordinates of the unbound apo-form of the protein. Indeed the vast majority of existing docking strategies already use some level of ''frozen'' approximation. Typically, the co-ordinates of the entire protein are held rigid whilst the ligands are posed within the active site. Even in docking methods that allow for protein flexibility, the locations of all or most of the backbone atoms or alpha carbons are constrained. It is also preferable to use an experimentally resolved complex structure of the intended target, than attempting to dock into an *apo* structure or a homology model [29]. Even when using molecular dynamics to investigate binding, it is advantageous to start simulations with the ligand correctly posed, and constrain the length of the simulation, rather than to diffuse compounds into relaxed proteins. The frozen ligand approximation is perhaps a natural progression in this ''take advantage of the available information'' philosophy, by asking what value can be made of the co-ordinates of the known complex. Surprisingly, with the exception of retrospective analyses, the experimentally determined ligand co-ordinates are often under-utilized, commonly used only to help define the active site and possibly constrain required interactions. This, at least, is an improvement upon protocols that attempt to locate active sites algorithmically, completely ignoring any biological insight, from previous experiment, such as sequence information or site-directed mutagenesis.

From the medicinal chemistry perspective, the restriction to isosteres is also intuitive, reflecting the incremental changes often performed during lead optimization. Steps such as mutating chlorine into fluorine, carboxylate into amide or amidine, or thioureas into ureas or guanidines, are at the heart of medicinal chemistry research. There is an important distinction between our use of the term isostere, denoting purely the similarity in shape, with the term bio-isostere, which additionally implies similar biological activity. An amidine may have the same shape as a carboxylate (they are isosteres), but the positive formal charge on an amidine may make it a poor mimic of a negatively charged carboxylate if this ionization performs a biological role. Often overlooked in the field of docking are the roles of ionization and tautomerism. The operational definition of an ''isosteric analogue'' used in this work, allows us to consider and evaluate the effects of the protonated and deprotonated forms or amines, carboxylic acids, imidazoles, etc. This is useful not only from the perspective of physiological pH, but also for considering the induced p$K_a$ shifts of both protein on ligand, and ligand on protein.

In retrospect, the concepts underlying the ''frozen ligand approximation'' are applied routinely in X-ray crystallography, though rarely applied to the ligand during docking. The resolution and other limitations of macromolecular X-ray crystallography means that the technique is unable to experimentally distinguish carbon from nitrogen from oxygen, or to observe the protons necessary to identify the appropriate tautomeric form or protonation state. Consequently, a frequent task in structure refinement is assignment of terminal atoms in glutamate (GLU) versus glutamic acid (GLN), and aspargine (ASN) versus aspartic acid (ASP). Likewise, when preparing an active site for docking, assigning protonation states to histidines (as HIE, HID or HIP) and potentially flipping the terminal torsion in gluatamine and asparagine to make improved hydrogen bonding interactions are common place. In these transformation steps, the same underlying assumptions are made, that the heavy atom co-ordinates remain identical, but the atom type (including atomic number, attached hydrogen count and formal charge) assigned to each atomic center may be independently optimized.

## Methods

### Electrostatic evaluation of isosteres

Once the degrees of freedom associated with atomic movement are frozen for a binding energy calculation, the remaining terms are predominantly electrostatic. The electrostatic contributions between different isosteres consist of the energetic differences between the protein and ligand interactions and the ligand solvation when unbound

$$\Delta\Delta E_{\text{binding}} = \Delta\Delta E_{\text{protein-ligand}} - \Delta\Delta E_{\text{ligand}}$$

As a consequence of eliminating all the invariant (and lesser) terms in these free energy calculations, the

interaction energies calculated are only relative energies. The penalty of these simplifications is that what is important is not the absolute energies, or even accurate relative energies, but a correct ordering of affinities.

In recent years it has been shown that continuum solvent models, i.e. the Poisson Equation (PE) and the Poisson–Boltzmann Equation (PBE) for ionic environments, can model this contribution well [20, 39]. In this work, we use a variant of the binding energy calculation used by ZAP [13]. In this scheme, the screened Coulombic interactions are calculated using a finite difference formalization of the Linearized Poisson–Boltzmann equation.

For calculation of molecular binding energies, ZAP has a number of advantages over earlier PB solvers, such as the popular DelPhi program [11, 33, 38]. PB solvers start with a description of a molecular volume and a set of point charges and calculate a potential field. Typically, this is solved by mapping the charges and dielectrics onto points on a regular cubic lattice, and iterating an initial guess of the potential at each point on the lattice until it converges using a discretized form of the PBE. Unfortunately, this discretization process and the method used to describe the water–solute interface can affect the quality of the solution obtained by this approximation process for irregular geometries. The primary source of the error is the discontinuous nature of the dielectric function and the point charges. As the dielectric function has one value inside the molecule and another outside it changes discontinuously at the molecular boundary and this behavior is difficult to capture on a finite resolution lattice. Naturally as grid point separation decreases the accuracy improves, but the number of points, and hence computation time, increases as the third power of the reciprocal grid spacing. ZAP overcomes these issues by using a Gaussian function, instead of the more usual hard molecular surface, to describe the molecular dielectric, providing a smooth, continuous variation of the dielectric between inside and outside of a molecule. In addition, charges are mapped onto the grid using a quadratic function, further removing artifacts from the calculated energies. These improvements allow the electrostatic contribution to binding to be evaluated with a confidence of ~0.5 kcals in about a second, compared to minutes or hours with other PB solvers [13]. Whilst the one-second per molecule performance of ZAP is already impressive, the restriction of electrostatic energy evaluations to isosteric compounds, allows for even greater improvements in performance.

The electrostatic energy of a system consisting of a set of point charges can be written

$$E = \frac{1}{2}\sum_i \phi_i q_i$$

where $\phi_i$ is the potential at charge $q_i$. This is a fundamental equation for electrostatics but is not, strictly, well-defined. Since, in classical theory, a point charge produces an infinite potential at itself, $E$ is by definition infinite. Fortunately, this is not really a problem as we are typically interested in the energy differences between two scenarios. Thus the quantity of interest is $\Delta E$ :

$$\Delta E = \frac{1}{2}\sum_i (\Delta\phi_i) q_i$$

where the infinite self energy cancels.

The second important observation is that for the linearized Poisson–Boltzmann equation (LPBE), the potential produced by any one charge is directly proportional to the size of that charge, and that the net potential of a set of charges is just the sum of the potentials from each individual charge. This implies that $\Delta E$ may always be written as:

$$\Delta E = 0.5 * \sum_i \sum_j k_{ij} q_i q_j$$

where the constants $k_{ij}$ are a property only of geometry. Hence, in linear classical electrostatics, energy is always quadratic function of charge. Note that this formulation contains ''diagonal'' or ''self'' terms that contain the cancellations of infinities. These terms represent the effects of polarization back on the source charge from the dielectric medium.

For the case of a ligand binding to a protein, let the ligand atom charges be specified by $q_i$ and those of the protein by $Q_i$, then the sum over all charges in $\Delta E$ of binding can be split into three terms:

$$\Delta E_{\text{protein}} = \frac{1}{2}\sum_i \sum_j c_{ij} Q_i Q_j$$

$$\Delta E_{\text{ligand}} = \frac{1}{2}\sum_i \sum_j k_{ij} q_i q_j$$

$$\Delta E_{\text{ligand–protein}} = \frac{1}{2}\sum_i \sum_j K_{ij} Q_i q_j$$

These correspond to a calculation of the protein electrostatic energy alone, that of the ligand alone and that of the complex. These can then be substituted into the equation for electrostatic energy given previously:

$$\Delta E = \Delta E_{\text{protein-ligand}} - \Delta E_{\text{ligand}} - \Delta E_{\text{protein}}$$

A key observation underlying this work is that once $\Delta E_{\text{protein}}$, $K_{ij}$ and $k_{ij}$ have been found the binding energy for a different ligand, $q_i$, can be trivially recalculated without the machinery of the PBE solver via the matrix equation:

$$\Delta E = \overline{q^T} \bullet \overline{\overline{M}} \bullet \overline{q} + \overline{I} \bullet \overline{q} + C$$

where $q$ is a vector where element $i$ is the atomic charge on the $i$th atom, $M$ is a square matrix with element $(i,j)$ set equal to $0.5*k_{ij}$, $I$ is a vector with element $(i)$ set equal to $0.5*\Sigma\, K_{ij}Q_i$, and where $C$ is the constant term $(-\Delta E_{\text{protein}})$. For a ligand of $N$ atoms the number of multiplications and additions to calculate $\Delta E$ given $M$, $I$ and $C$ is $N*(2N-1)$, or if the symmetry of $M$ is employed, $(N^2 + 5N - 1)/2$. For typical ligands in medicinal chemistry, this corresponds to about 1000–4000 operations. On a modern workstation this translates to a calculation time of ten to a hundred thousand binding energies per second.

Kangas and Tidor [24–26] have taken this matrix formulation for the binding energy one step further by optimizing for an optimal set of ligand charges. By recognizing that if $Q_i$ are fixed, the equation for the interaction energy is a quadratic function of the ligand charges, which can be solved directly to find a minimal value. These 'perfect' charges then represent the best possible electrostatic design for a given set of atom centers and a given pose in an active site. Unfortunately, the optimal partial charges calculated by this method are not guaranteed to be physically meaningful. Small eigenvalue of $M$, when inverted, give charges to large to be realized. This is not improved by limiting the domain of possible charges or the net charge. It is also possible that numerical imprecision causes the quadratic form to have negative eigenvalues, indicating that the optimal interaction can be found by placing an infinite charge on one or more atom centers. These issues have limited the application of the 'perfect' charge method. The approach described here does not suffer from these drawbacks as it merely uses the formalism to allow the electrostatic calculation to keep pace with the enumeration of isosteric molecules.

The calculation of $M$, $I$ and $C$ involves a sequence of $2N + 2$ ZAP calculations, where $N$ is the number of ligand atoms. The first $N$ are performed on the ligand alone, wherein each atom is given a charge of 1.0 and the potential at each atom in the ligand from this charge is calculated. The second $N$ calculations are identical except that the ligand is placed in the context of the protein active site, but with all charges on the protein atoms set to zero. The difference in the atom potentials found from these two sets of calculation construct the matrix $M$. Next a calculation is performed on just the protein, with all atoms carrying their normal charges $Q_i$. Finally, this calculation is repeated with the ligand molecule in the active site but with all of the ligand charges set to zero. The differences between the potentials at each protein atom allow $C$ to be calculated, and the interaction vector $I$ is just the potential at each ligand atom from the last calculation. Typical preprocessing time for an average protein and ligand is between one and two minutes.

## Generation of isosteres

The other component used in this evaluation is an isosteric analogue generator. In the above matrix formulation, each trial compound to be evaluated is represented by a vector of atomic charges. The approach proposed to construct these vectors is to enumerate a large number of isosteric analogues, and then for each one use a standard partial charging method, such as MMFF94 charges. The isostere enumeration algorithm used in this evaluation is implemented in the program WABE that takes as input a starting compound and a training database of allowable chemistry. The first step of the algorithm is to reduce the specified molecule to an abstract representation of its shape that encodes only its heavy atom connectivity and non-terminal hybridization. The hybridization states can be assigned either from the 2D valence and local connectivity, or from the 3D geometry. At this stage, the atomic number, formal charge, and implicit hydrogen count of the original template are suppressed. A graphical representation of this generic reduced representation is shown in Fig. 1.

The next stage of the WABE algorithm is to repopulate this template. Abstractly, the method is to assign atom types to each vertex. Suitable atom types are discovered from analysis of a training database of molecules. An atom type is potentially substitutable in the template if it has the same hybridization and heavy atom degree (the number of bonds to non-hydrogen

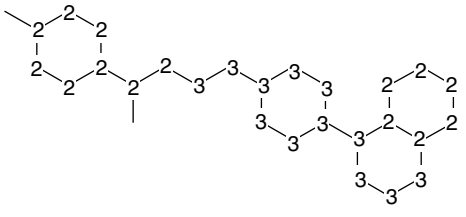


**Fig. 1** Sudoku for chemists: the hybridization labeled skeleton used to define a set of isosteres

atoms). However, if each atom position in the scaffold were assigned independently, the virtual molecules generated would be unrealistic, containing bizarre chemistries based on unusual juxtapositions of elements. To overcome this, and restrict the enumeration to reasonable looking molecules, WABE adds an additional constraint to each assignment. A substitution is only allowed if the environment, formed by an atom type and its immediate neighbors, is seen in the training database. The training database is analyzed, a molecule at a time, and reduced to sets of atom types and atom environments. An atom type encodes the atomic number, formal charge, implicit hydrogen count, hybridization and heavy degree. An atom environment encodes a central atom type and a list of its neighboring atom types, including the bond order to each. Once each of the atom types and atom environments has been identified, the algorithm determines the substitution classes. An atom class is the set of atom types that have the same hybridization and heavy degree. Likewise, an environment class is the set of atom environments whose atom types belong to the same atom class.

The task of enumerating isosteres is analogous to solving a jigsaw puzzle. Each piece of the puzzle is an atom type, and the shape of each piece is its environment class. The edges of each piece correspond to the neighboring atom constraints. The actual enumeration itself is implemented using a simple backtracking algorithm that continually attempts to assign atom types to each vertex, checking that each assignment is consistent with any previously assigned neighbors (Fig. 2).

By modifying the contents of the training database, it possible to control the types of isosteres generated. Giving a larger training database typically results in enumerating more isosteres. If the database contains drug-like molecules, the enumerated isosteres should be composed of drug-like functionality. Using a company's registry as a training set restricts isosteres to the types of chemistries used within an organization. Using a database of available vendor compounds produces isosteres composed of fragments that may be purchased. Finally, using a database of known actives generates isosteres containing functionality known to be of relevant to a given target.

Refinements to the above algorithm include (i) duplicate elimination, to avoid alternate Kekulé forms or trivial automorphisms, (ii) ability to tolerate and preserve atom environments in the original molecule that are not represented in the training set, and (iii) the ability to specify which parts of the input molecule are mutable and which parts should be held fixed.

This approach to the generation of isosteres is similar in concept to the MARABOU program developed by the University of Leeds for use with their SPROUT *de novo* structure-based drug design program [10, 28]. In their approach, the pose generation stage places "skeletons" with no information concerning atom types in the active site. As above, these skeletons contain just hybridization and bonding information. Once posed, the interactions between the atoms of the skeleton and the protein active site are analyzed, and positions on the skeleton classified as donor, acceptor or donor/acceptor in order to confer the appropriate binding character. Finally, the program MARABOU (Modify Atoms in Results to Achieve Binding Organic Units) addresses the problem of atom type assignment to match the desired constraints using an expert system with a knowledge base of 46 functional group substitutions.

## Hydrogen placement

One subtlety arising from the synergy of the two methods described above concerns the handling of hydrogen atoms. After failed attempts to use united atom representations for the electrostatic evaluations, it was deemed necessary to sprout explicit hydrogen atoms at standard geometries. For each isostere, prior to electrostatic evaluation, the previously implicit hydrogen atoms are placed around the template using standard bond lengths and ideal bond angles [34].

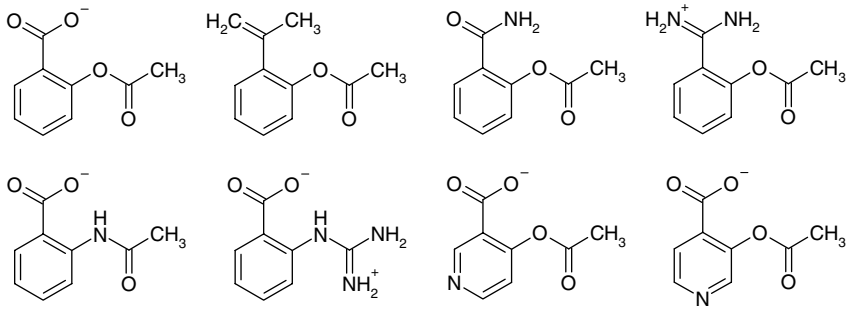


**Fig. 2** Examples of isosteric analogues generated for aspirin

During this procedure, no reference is made to the protein structure to avoid bumps or select preferred torsions for rotatable hydrogens on terminal $sp^3$ hybridized atoms. The adverse impact of placing this freely rotatable hydrogen arbitrarily is discussed in the results section, and will be addressed in future work.

Because the hydrogens are deterministically placed around the scaffold based on the heavy atom co-ordinates of the template, the number of possible hydrogen atom locations is bounded. For the matrix approximation, the charge vector represent all of the heavy atom locations, plus the set of all generated hydrogens. The dielectric volume from protons is small but when combined from many 'virtual' positions on a single atom, it becomes significant. Calculations showed that the off-diagonal terms (atom-atom potentials) and heavy atom self-energies (diagonal) were better represented by removing the volume of such protons (effectively burying them in the bonding atom) but that the self-energy terms for each such proton was more reliably calculated by assigning a normal proton radius. For most of this work we have used a proton radius of 1.0Å. Other radii were: C = 1.7 Å, N = 1.55 Å, O = 1.6 Å, S = 1.9 Å, F = 1.2 Å, I = 2.1 Å. The internal dielectric was set to 1.0. This is not physically correct for either protein or ligand but reflects the choice of charging model, MMFF, where charges have been developed to reproduce vacuum potentials. Results not presented suggest this leads to a relatively constant shift in energies compared to a charge description suitable for, and in conjunction with, more physical dielectrics.

## Results and discussion

### PB approximation accuracy

The first validation of our approach to evaluating isosteric analogues is to compare the binding energies calculated with our implementation of the ''matrix'' approximation (called GIMBLE) with those calculated by performing a full Poisson–Boltzmann calculation in ZAP.

The plot below shows a comparison of the ZAP calculated and GIMBLE calculated energies for 1320 isosteric analogues of methotrexate bound to dihydrofolate reductase (using the co-ordinates of PDB file 3dfr). The details of the protein preparation are given in the section on dihydrofolate reductase below (Fig. 3).

This graph shows good agreement between the two sets of calculations, especially when considering the

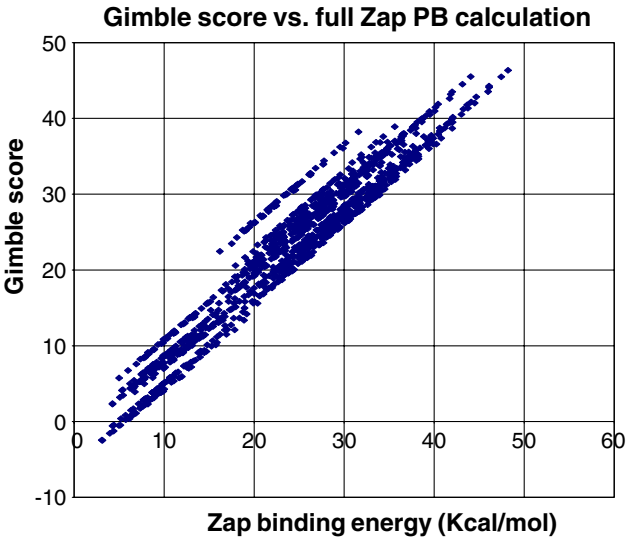



**Fig. 3** Comparison of binding energies as determined by GIMBLE's matrix approximation to full ZAP electrostatic binding calculation for 1320 analogues of methotrexate to DHFR

required CPU times. The time taken to perform the full ZAP calculations for these 1320 compounds was 54 min and 55 s (on a 2.8 GHz Pentium-4 PC), averaging about 2.5 s per complex. The time taken using the matrix approximation in GIMBLE (on identical hardware) was only one minute 34 s (or less than a tenth of a second per molecule). Even this result is skewed by the large set-up time and small number of isosteric analogues. Repeating the experiment with 33,250 isosteric analogues of methotrexate increases the GIMBLE run-time to 2 min 41 s.

A noteworthy feature of the above scatter plot is the marked striation, where clearly the analogues are grouped in bands, each with a small constant error in their estimated binding energies. The cause of this error is due to our neglect of the effects of changing atomic radius when mutating atoms from one element to another. Because the energies calculated from the PBE are sensitive to the dielectric boundary, and therefore the radius of each atom, the assumption that the same values of $M$, $I$ and $C$ can be used for all isosteres is flawed and introduces these small errors. The primary variation occurs in the diagonal terms of $M$, which describe the change in the self-energy of the charge on each atom, the ''desolvation'' term. For example, all of the compounds in the top band of the scatter plot above contain a bromine atom, attached to the central (pteridine) ring system.

For this work, this residual error was considered tolerable (for example, the best scoring compound of one method is almost always the best scoring of the

other), but this issue will be addressed in future work. The three mitigation strategies under consideration are: (i) to restrict the set of isosteric analogues to elements with approximately the same VdW radius. (ii) use GIMBLE as a fast pre-screen and then use ZAP to accurately rescore the best few thousand results. (iii) categorize the atomic substitutions into groups based upon significant changes in radii, and then perform multiple GIMBLE runs, one for each subset of compounds. This last strategy takes advantage of the GIMBLE's impressive speed, where the overhead of ten (or so) approximate calculations is still dramatically faster than a single full PB calculation.

## Isostere enumeration

One potential concern of a method that restricts itself to isosteres of a given lead compound is whether there are a sufficient number of analogues to be useful. Moreover, a follow-up question is what sort of chemical diversity can be found in such a restricted subset of chemical space or virtual library. Interestingly, it transpires that common drug-like molecules can often have many millions of isosteric analogues. Indeed, the vastness of virtual chemical space means that the number of potential compounds with a high 2D similarity to any starting point is huge.

An analysis of the 250,251 compounds in the National Cancer Institute's NCI00 database reveals it contains 173,784 unique shape templates, of which 143,285 (82%) are singletons. The most common shape occurs 258 times.

The particularly high fraction of singletons found in discrete compound databases suggests that isosteric analogues are perhaps an under-explored aspect of virtual chemical space.

By definition, some physical chemical properties of a set of isosteres are constrained; all enumerated compounds have the same heavy atom count and number of rotatable bonds. However, for other properties of interest, isosteric analogues show a wide range of values, indicating a reasonable chemical diversity. The graphs below show the calculated molecular weight, XlogP [45] and hydrogen bond donor and acceptor counts for a set of 3874 isosteric analogues of celecoxib (Fig. 4).

As shown by the above histograms, although (as expected) the variation in molecular weight is not large across a virtual library of isosteres, there is a reasonable distribution of partition coefficients, hydrogen bond donors and hydrogen bond acceptors. This diversity should enable isosteric analog exploration to be used to optimize ADME properties or to trade-off potency against bioavailability using techniques such as those proposed in this work.

## Dihydrofolate reductase

The first test case used to evaluate the methods described above is the optimization of methotrexate isosteres in the active site of dihydrofolate reductase (DHFR), a classic benchmark in the field of docking. A schematic representation of the interactions between methotrexate and the active site of DHFR is shown in Fig. 5.

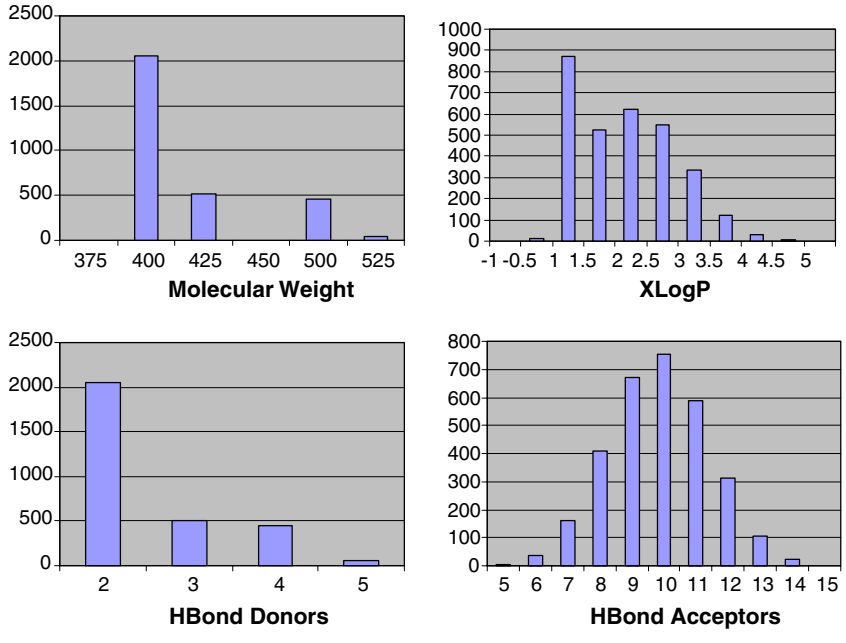



**Fig. 4** Distributions of physical properties for 3874 isosteric analogues of celecoxib (vioxx)

A critical binding interaction in this model system is the protonated nitrogen at the 1 position on the pteridine ring, which forms a hydrogen bonding interaction with Asp27 (as shown in the top right corner). The difficulty of this system for many docking protocols is that the pKa of this N1 nitrogen is greater than 7, implying that this position is normally unprotonated at physiological pH. We therefore assumed that this would be an ideal system to evaluate our isosteric evaluation, expecting to preferentially select isosteres that were suitably protonated at this position.

Using the coordinates from the PDB file 3dfr, the active site was prepared by removing all of the water atoms from the structure, placing the ligand residue ''MTX'' in one file, and the remaining protein atoms in another. The protonation states of ionizable residues were automatically assigned physiological defaults, and hydrogen atoms (on both the protein and ligands) sprouted at standard geometries [34]. Finally, MMFF94 partial charges were assigned to all atoms.

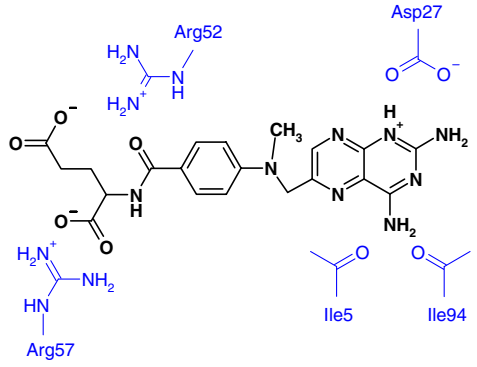



**Fig. 5** Interactions of methotrexate to the active site of DHFR

Our very first trial used a relatively small training database that resulted in enumerating 1320 analogues of methotrexate. Examples of the types of substitutions to the core ring system are shown in Fig. 6.

As described in the above section on PB approximation accuracy, the GIMBLE calculation completed in 94 s and agreed with a full ZAP binding calculation on the best and the worst scoring isosteres. The best and worst of the 1320 are shown in Figs. 7 and 8.

As we had hoped, the proposed method correctly identified the protonation of N1 as beneficial to binding, in that it was present in all the top-scoring compounds. However, the isostere enumeration and evaluation also identified several other substitutions that it predicted to improve binding. A literature search revealed all of the individual modifications suggested to methotrexate had been observed in experimental SAR studies [12, 17, 43]. Specifically, the substitution of the terminal carboxylate to an amidine group is reasonable as this functional group points out into solvent. Its role in the biological activity of methotrexate is to assist in transport. The non-terminal carboxylate makes a critical interaction with DHFR and is correctly identified and preserved. The nitrogen that is replaced by a carbon in the linker region has been shown experimentally not to be important in binding. Finally, the substitutions of the aromatic

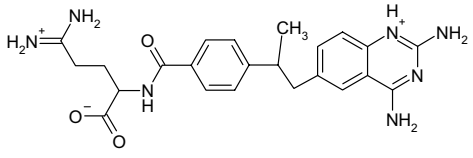


**Fig. 7** Best scoring methotrexate analogue (of 1320) in DHFR

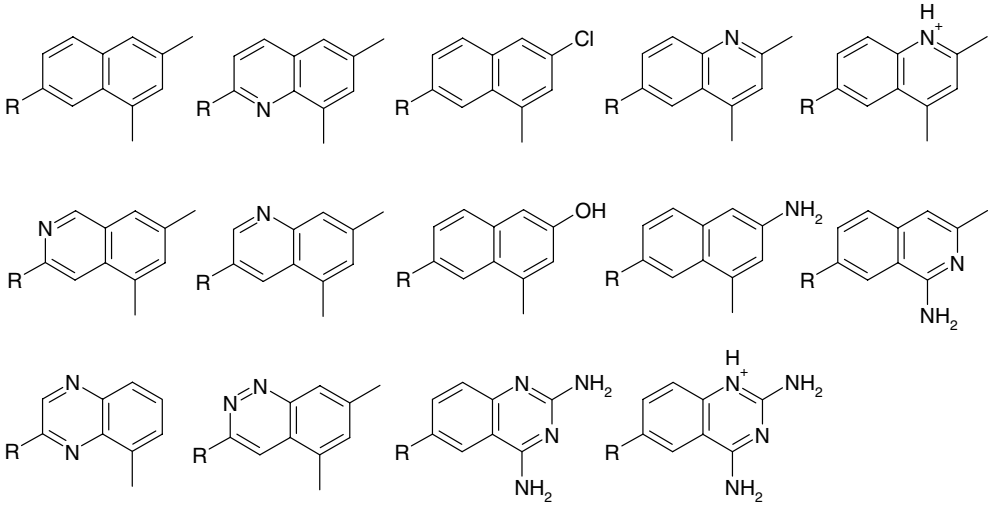


**Fig. 6** Example isosteric substitutions made to methotrexate's core ring system. The R represents the rest of methotrexate analogues that has been omitted for brevity

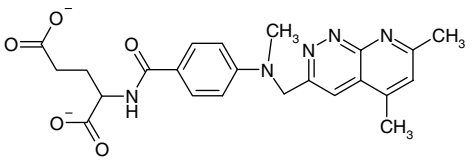


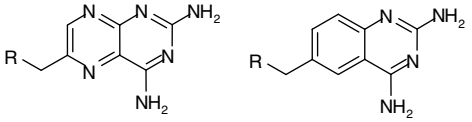
**Fig. 8** Worst scoring methotrexate analogue (of 1320) in DHFR

nitrogens in the pteridine ring system to form a quinazoline have been shown experimentally to improve binding (Fig. 9).

Comparing the best scoring analogue to the worst scoring analogue generates an interesting observation. Although the size of the training database was small, such that some of the atom positions in methotrexate had no substitutions, the non-terminal alpha carboxylate (which was freely substitutable) appears preserved in both the best and worst scoring compounds. The explanation is that the PB electrostatic binding energy calculation is a whole molecule property, and not just the sum of independent contributions. The reason for GIMBLE's selection of the terminal amidinium was to balance the net charge on the system, and minimize the desolvation penalty. In the worst scoring analogue, the desolvation penalty from the net charge of negative two is greater than the positive contribution made by the carboxylate.

The input structure of methotrexate (unprotonated at N1) was ranked 221st of 1320. The explanation for this poor showing is the combinatorial nature of the multiple improvements discovered by WABE/GIMBLE. Most of the higher scoring compounds were protonated at N1, or contained different combinations of one, two or three of the nitrogen substitutions.

After the success of the methodology on this model system, we repeated the experiment with a larger training database to investigate whether this approach scaled well. The second trial, which generated 33,250 analogues of methotrexate, identified exactly the same compound as the best scoring. The worst scoring compound, however, demonstrated that sampling had indeed been increased (Fig. 10).



Exptl $IC_{50}$ 3.1μM Exptl $IC_{50}$ 0.125μM

**Fig. 9** Experimental binding affinities for pteridines versus quinazolines in methotrexate-like compounds bound to DHFR [12]. For brevity, the R represents the remaining identical part of the measured compounds

p56 Lck kinase

In order to evaluate whether the proposed methodology was transferable to other systems, we were invited by AstraZeneca to repeat the experiment on a system of pharmaceutical interest. The system suggested was inhibition of p56 Lck kinase, where a significant amount of proprietary assay data was available for closely related (isosteric) series. The figure below shows the structure of AZ compound M532121 for which the structure of the protein–ligand complex had been solved experimentally by X-ray crystallography (J. A. Grant et al. 2005, submitted) (Fig. 11).

The crystal structure for the complex was resolved to a high resolution (1.5 Å resolution with an $R$ of 19.7% and a $R_{free}$ of 21.7%).

The $IC_{50}$ inhibition data available for this system prompted us to slightly tweak our protocol. Instead of using M53212 as the template, we elected to cleave the 2-hydroxyethyl group from the guanidine, and evaluate the isosteric analogues of this slightly smaller compound. The motivation for this modification was the availability of experimental binding information for similar compounds containing a urea or thiourea at this position in the structure (Fig. 12 and Table 1).

The inhibition data made available was from multiple runs of a high-throughput screening assay, with compounds often measured multiple times. Despite the

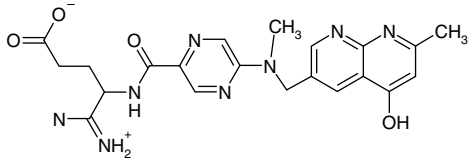


**Fig. 10** Worst scoring methotrexate analogue (of 33,250) in DHFR

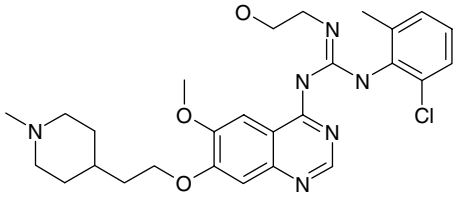


**Fig. 11** Structure of Astra Zeneca M53212

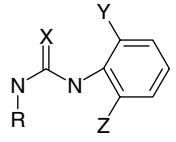


**Fig. 12** The fragment of interest of the Lck kinase inhibitors. R represents the rest of the molecule, held constant for this study

**Table 1** Experimental $IC_{50}$ (nM) inhibitions for Lck kinase

| X | Y | Z | IC50 (nM) |
|---|---|---|---|
| =O | Cl | Cl | 1–3 |
| =O | F | F | 0.8–6.9 |
| =O | CH3 | CH3 | 1.4–2.8 |
| =O | CH3 | Cl | 1–5.3 |
| =S | Cl | Cl | 3–8.7 |
| =S | F | F | 4.5–23.7 |
| =S | CH3 | CH3 | 4.8–755 |
| =S | CH3 | Cl | 2.4–6.2 |
| =N | Cl | Cl | 14–22 |
| =N | F | F | 41–46 |
| =N | CH3 | CH3 | 4.8–6.8 |
| =N | CH3 | Cl | 3–7.8 |

experimental uncertainty, the data shows a preference for ureas and thioureas over guanidines (with ureas slightly better than thioureas), i.e. = O ≤ = S < = N, and ring substitutions patterns, CH3/Cl < CH3/CH3 ≤ Cl/Cl ≤ F/F.

The binding scores predicted by GIMBLE (lower is better) are given in the tables below. In these calculations, the piperidine nitrogen was assumed protonated, but identical trends are observed when the ring nitrogen is kept neutral. (Tables 2 and 3).

In the first table, it can be seen that GIMBLE is able to predict that the ureas and thioureas bind better than the guanidines, though the method mispredicts that thioureas will score slightly better than ureas. This error is due to the significantly larger radius of sulfur than oxygen, as discussed previously. Rescoring these compounds with a full ZAP calculation correctly reorders ureas as better than thioureas. One advantage of theoretical methods such as GIMBLE is that it allows the binding energy prediction of both the neutral and protonated forms of guanidine. In this case, the protonated guanidinium, likely at physiological conditions, results in worse binding than the neutral form.

**Table 2** GIMBLE predicted Lck binding scores #1

| X | Y | Z | Gimble |
|---|---|---|---|
| =O | CH3 | Cl | 4.193 |
| =S | CH3 | Cl | 4.133 |
| =NH | CH3 | Cl | 4.375 |
| =NH2+ | CH3 | Cl | 6.512 |

**Table 3** GIMBLE predicted Lck binding scores #2

| X | Y | Z | Gimble |
|---|---|---|---|
| =NH2+ | CH3 | F | 6.472 |
| =NH2+ | CH3 | Cl | 6.513 |
| =NH2+ | Cl | Cl | 6.876 |
| =NH2+ | F | F | 6.883 |
| =NH2+ | CH3 | CH3 | 8.331 |
| =NH2+ | Cl | CH3 | 8.468 |
| =NH2+ | F | CH3 | 8.498 |

The second table also shows that the gross ordering of substitutions on the phenyl ring system are correctly predicted by GIMBLE; $CH_3/Cl < CH_3/CH_3 \leq Cl/Cl \leq F/F$. Again, an advantage of theoretical methods is their ability to distinguish the preferred binding orientation of asymmetric di-ortho-substituted isosteres. The table above gives two binding energy values for $CH_3/Cl$, evaluating two binding conformations of the same ligand.

In addition to confirming that the protocol was able to reasonably reproduce some of the available SAR for Lck kinase, the isostere enumerator was run in prospective mode with a large training database. This enumeration resulted in 1,792,000 isosteres. The best scoring compound is shown below (Fig. 13).

Although, this molecule retains a large number of features recognizable as important for kinase inhibition, features such as the selection of an olefin are uncommon in medicinal chemistry. However, an important aspect of the SAR, previously known to AstraZeneca and identified by GIMBLE was the importance of protonation of the piperidine ring. Poisson–Boltzmann methods had previously been used to rationalize why substitutions at this position were important to binding affinity, although this group is deep in solvent and makes no close contacts with the protein (J. A. Grant et al. 2005, submitted). Instead of the more usual single short-range electrostatic interaction (such as a hydrogen bond or salt bridge), this group is important for its cumulative long-range interactions with the many negatively charged groups on this side of the protein target.

Studying the rank ordered list of the nearly 1.8 million isosteres, shows that many of the top scoring compounds contain recognizable groups. In second place is a fluoro/amino substituted phenyl group, and third place is the chloro/methyl of the original input structure. The methoxy group at the original 7-locant appears in the fifth placed compound, a 1,5-naphthyridine by 13th place, a thiourea at 15th, a quinazo-

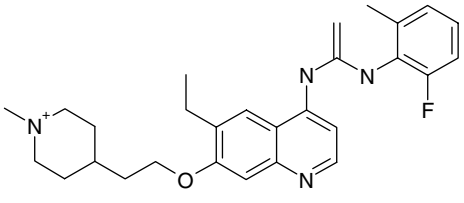



**Fig. 13** Best scoring Lck inhibitor analogue (of 1.8 million)

line by the 71st position, and a urea by the 82nd position. Indeed, after screening 1,792,000 compounds the experimentally assayed Astra Zeneca thiourea appeared at position 360, and the Astra Zeneca urea at position 535.

## Critical evaluation

To determine the limits of the proposed isostere evaluation protocol, we selected five ''amenable'' model systems at random from the recent medicinal chemistry literature. Although, in-house crystallography efforts at pharmaceutical companies typically have a wealth of protein–ligand structures for closely related compounds, there are relatively few examples in the publicly available protein databanks. The five case studies presented below, ribosome inactivating protein, L-arabinose binding protein, neuraminidase, HIV-1 reverse transcriptase and COX-2, are ordered by increasing difficulty.

### Ribosome inactivating protein

The first of the six systems used as a validation benchmark, was ribosome inactivating protein (RIP) selected from Miller et al., ''Structure-Based Design and Characterization of Novel Platforms for Ricin and Shiga Toxin Inhibition'' [30]. Pharmacologically, this protein class is the point of interaction of the toxin Ricin, used in the famous ''umbrella tip'' assassination of Georgy Markov. Of interest here are protein–ligand complex structures of two isosteric compounds to the same target deposited in the PDB. The structure of 9-deazaguanidine (9DG) is available as PDB code 1il4, and that of 7-deazaguanidine (7DG) as 1il3. The structures of these two compounds are given below, together with that of xanthine, whose binding affinity is also reported to the same target (Fig. 14).

To confirm that the ''frozen ligand hypothesis'' (that isosteric compounds often adopt similar binding poses) applies in this case, the two protein structures were superposed using the protein alpha carbon atoms. The image below shows that the binding modes of 9DG and 4DG in RIP are nearly identical, and that the atomic coordinates of one ligand can serve as an accurate approximation of the other (Fig. 15).

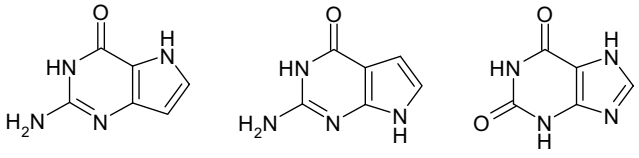



**Fig. 14** Structures of 9-deazaguanidine (9DG) from 1il4, 7-deazaguanidine (7DG) from 1il3, and xanthine

Table 4 shows that the GIMBLE predicted binding scores, using 1il3 as a template, correlate well with the experimentally determined binding affinities, i.e. that 9DG binds better the 7DG which in turn binds better than xanthine. The two predicted values for xanthine reported above, XAN1 and XAN2, describe the scores given for the two different tautomers of xanthine in the active site. Identical results, 9DG < 7DG < XAN, are also obtained for this series using the structure from 1il4 as the template.

In this example, we introduce the principle that the best scoring tautomer of xanthine is probably the bound form (though no account is taken of the relative energetic penalties of different tautomeric forms). This is similar to the assumption made previously with DHFR that the energetic cost associated with the protonation of methotrexate's N1 is minor compared to the beneficial interaction it gains. One topic of future study is to investigate the contribution of this penalty in GIMBLE's binding energy calculation.

### L-Arabinose binding protein

The second validation system is L-arabinose binding protein (ABP) taken from Shen and Quiocho [39]. In

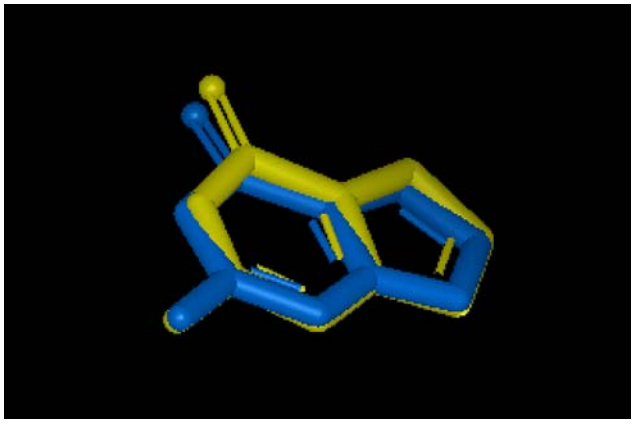


**Fig. 15** Overlay of the ligand binding poses from 1il3 (blue) and 1il4 (yellow)

**Table 4** GIMBLE energy score vs. experimental binding affinity

| | Gimble | IC50 (mM) |
|---|---|---|
| 9DG | −6.437 | 1.4 |
| 7DG | −5.081 | 2.8 |
| XAN1 | 3.346 | 3.6 |
| XAN2 | 7.907 | |

that work, the binding of a number of closely related (but not quite isosteric) sugars was studied, and the protein–ligand structures of several deposited in the PDB. For this case study, the concept of the frozen ligand approximation was stretched, in that we used the coordinates for a subset of the ligand atoms in Lck-kinase. Although, each compound does not have an identical ''shape'', all are substructures of one of the ligands, which can be used as a template (Fig. 16).

Repeating the protocol above, the crystal structures of the three ABP complexes were superposed and the resulting overlay shown in the figure below. Once again, for this case study, the ligand structures overlay almost identically, with the template providing a reasonable approximation of the other ligands' poses (Fig. 17).

Using the structure of α-D-galactose (GLA) from 5abp as a template, GIMBLE was used to predict the electrostatic binding energies of α-D-fucose (FCA), α-L-arabinose (ARA), 6-F-deoxy-galactose (6FA), 2-deoxy-galactose (2DA) and 1-deoxy-galactose (1DA). The GIMBLE scores together with experimental ΔΔG (in kcal/mol) relative to GLA are presented in Table 5.

With the exception of the template GLA, the predicted affinities correlate well with the experimentally

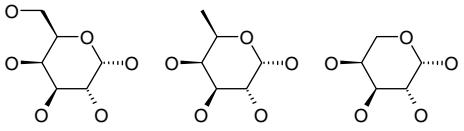


**Fig. 16** The structures of α-D-galactose (PDB code 5abp), α-D-Fucose (1abf) and α-L-arabinose (1abe)

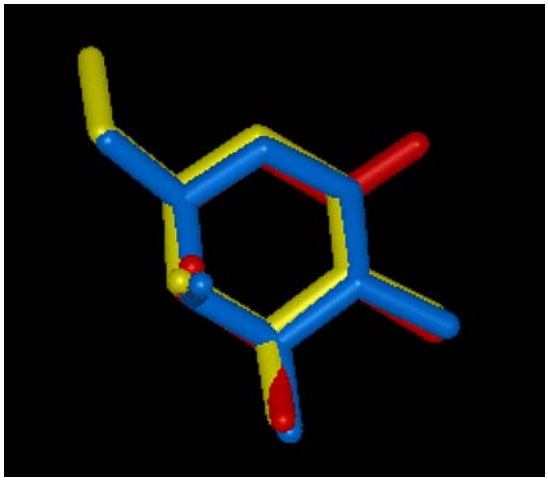


**Fig. 17** Overlay of the ligand binding poses from 5abp (yellow), 1abf (blue) and 1abe (red)

**Table 5** GIMBLE energy score versus experimental binding affinity

| | Gimble | Exptl ddG |
|---|---|---|
| GLA | 14.380 | 0.0 |
| FCA | 10.086 | 1.7 |
| ARA | 9.739 | –0.5 |
| 6FA | 13.079 | 2.6 |
| 2DA | 15.137 | 3.6 |
| 1DA | 10.053 | 1.8 |

determined values, ARA < FCA ~ 1DA < 6-FA < 2DA. Indeed, excluding the GLA result, the correlation coefficient $R^2$ between the experimental and predicted values is 0.685.

The GIMBLE misprediction of GLA's binding affinity can be traced back to the simplistic protocol used to prepare isosteres for this evaluation. The only difference between the poorly predicted GLA and the more accurately predicted FCA is the addition of a single hydroxyl group. As described previously, the hydrogen on this freely rotating position is positioned arbitrarily with respect to the protein. In the case, of D-galactose binding in ABP, the problematic hydroxyl forms a hydrogen-bonding interaction to ASP89 of the active site. If the hydrogen is placed optimally toward the carboxylate of ASP89, the electrostatics match well, but if the hydrogen is placed in the opposite direction, the interaction between the oxygens score poorly. Clearly, in order to correctly evaluate electrostatic binding interactions the effects of hydroxyl rotors needs to be taken into account. Fortunately, the dummy hydrogen position mechanism used by GIMBLE (as described in the methods section) provides a mechanism for addressing this issue, which will be a factor in future work. It is interesting to observe that hydrogen mobility did not appear to be a factor in any of the previous results.

## Neuraminidase (Sialidase)

The model system of 2,3-didehydro-2-deoxy-*N*-acetyl-neuraminic acid (DANA) bound to influenza virus neuraminidase was identified in a paper by Fornabaio et al. [7]. This case study makes use of four neuraminidase crystal structures from the PDB containing analogues of DANA, as illustrated by the structure below (Fig. 18).

The four isosteric compounds are DANA itself, where $R_1 = OH$ and $R_2 = OH$ (PDB code 1f8b), 4-amino-DANA, where $R_1 = OH$ and $R_2 = NH_2$ (1f8c), 9-amino-DANA, where $R_1 = NH_2$ and $R_2 = OH$ (1f8d) and finally 4,9-diamino-DANA, where

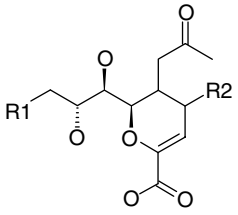



**Fig. 18** Generic structure of DANA analogues considered in this study

$R_1 = NH_2$ and $R_2 = NH_2$ (1f8e). As shown below, the binding modes of these isosteric compounds are nearly identical (Fig. 19).

The table below shows the agreement between the GIMBLE scores and the experimental ΔG (in kcal/mol) as reported in Fornabaio et al. [7]. For these GIMBLE calculations, the amino substitutions were all treated as charged, which would be their expected protonation state at physiological pH (Table 6).

Once again the agreement to experiment and its ability to correctly rank order the four isosteric analogues is impressive, with 4A < DANA < < 49A < 9A. The $R^2$ correlation for the GIMBLE energies is 0.89, and for full Poisson–Boltzmann calculations using ZAP, 0.95.

There is, however, one caveat to the impressive results obtained above for this case study, which is why it was presented third in order of increasing difficulty. The experimental $K_i$ values used to calculate the above ΔGs were observed in an assay performed at pH 5.5. Indeed, the topic of the Fornabaio et al. paper is on pH effects to the free energy of binding. Following the hypothesis proposed in their discussion section, the active site residue GLU276 was forced neutral, deviating from the default protonation state rule used in preparing the other active sites used in this study. Indeed, the exceptional agreement produced by Poisson–Boltzmann methods would appear to strongly support the hypothesis of the Kellogg group that this residue is neutral under the conditions of the experimental assay. For the purposes of influenza drug-design, running GIMBLE under default conditions should be sufficient, and the requirement to model protein $pK_a$ shifts only necessary to reproduce the non-physiological assay results.

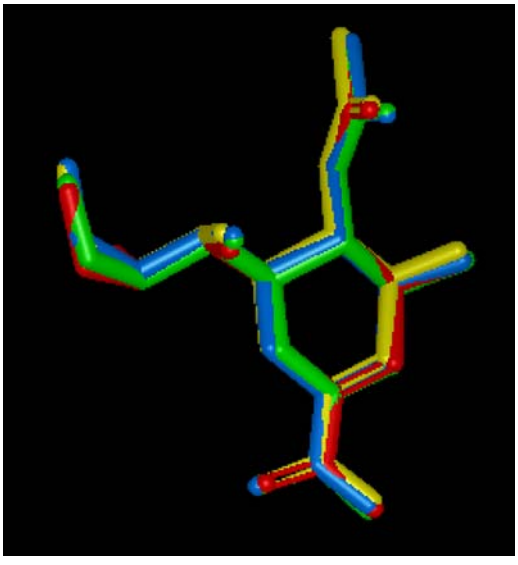


**Fig. 19** Overlay of the binding poses of DANA (1f8b), 4-amino-DANA (1f8c), 9-amino-DANA (1f8d) and 4,9-diamino-DANA (1f8e)

**Table 6** GIMBLE energy score versus experimental binding affinity

| | Gimble | Zap | delta *G* |
|---|---|---|---|
| DANA | −18.359 | −18.359 | −7.4 |
| 4A | −28.349 | −29.777 | −10.1 |
| 9A | 6.904 | 2.63 | −4.6 |
| 49A | −0.047 | −6.321 | −6.6 |

## HIV-1 reverse transcriptase (RT)

The fourth model system, non-nucleoside inhibitors of HIV-1 reverse transcriptase, is at the very limits of the proposed protocol's ability to accurately predict electrostatic binding energies. This case study selected from a paper by Rizzo et al. 36 investigated the binding of a set of near isosteric analogues of HEPT. The generic structure of the considered HEPT analogues is shown in the diagram below (Fig. 20).

This validation set consisted of a set of eight compounds taken from the relevant paper. The crystal structures of complexes involving HEPT, where X = S, $R_1 = CH_3$ and $R_2 = CH_2OCH_2CH_2OH$ (H01), and MKC-442, where X = C, $R_1 = CH(CH_3)_2$ and $R_2 = CH_2OCH_2CH_3$ (H11) are available in PDB as 1rti and 1rt1, respectively. Once again the measured binding data is for similar but not isosteric compounds,

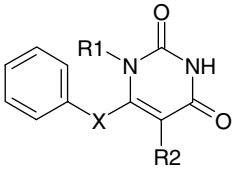



**Fig. 20** Generic structure of HEPT analogues considered in this study

so again the common substructure relaxation of the frozen ligand approximation is applied.

The binding mode superposition of HEPT (from 1rti) and MKC-442 (from 1rt1) shows that although the binding modes are relatively similar, the overlay is poorer than those previously presented (Fig. 21).

The results for the GIMBLE electrostatic binding energy calculation for the HEPT isosteres, together with their identities and experimental affinities, are presented below (Table 7).

Unfortunately, this time the correlation for these compounds is relatively poor. Whilst, there is some correlation for the first four results, there appears to be no signal in the last four. There are a number of possible causes for these limitations of the proposed method. One possible explanation is the resolution of 1rti (3.0 Å) and 1rt1 (2.55 Å) is lower than those presented earlier. This reduced quality of these structures may account for the poor quality of the binding mode overlay in this test case. Another hindrance is the nature of the binding site. The non-nucleoside binding site of reverse transcriptase forms only a single hydrogen bond to HEPT, and this interaction was to the ring nitrogen that was held invariant over the analyzed isosteres. All the interactions evaluated were principally hydrophobic or desolvation in nature. There were also two classes of changes studied in this series; the structural changes to $R_2$ and substitutions of the linker atom X. The $R_2$ changes involve significant shortening of a sidechain. In a hydrophobic pocket, the varying Van der Waals contributions and the possibility of alternate alkyl conformations and their entropic effects stretch the assumptions underlying the investigated frozen ligand approximation. Likewise, the substitution of the sulfanyl linker atom of HEPT with either a methylene or oxy linker potentially results in a significant conformational/shape change. In addition to the VdW radius changes mentioned earlier, the conformational effects of changing bond lengths and bond angles, by substituting sulfur for carbon or oxygen, is amplified in the hinge region of articulated molecules.

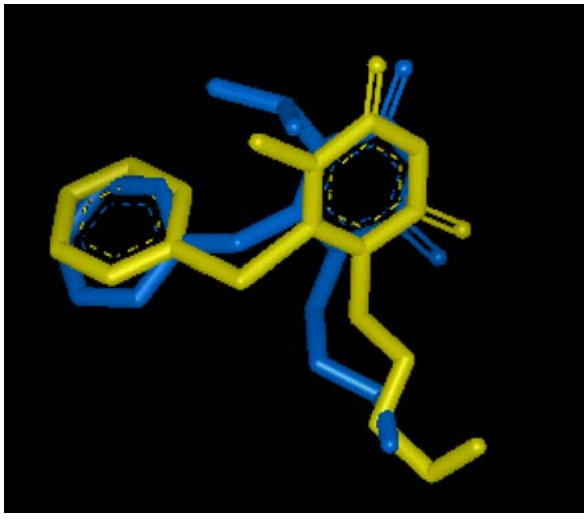


**Fig. 21** Overlay of the ligand binding poses from 1rti (yellow) with 1rt1 (blue)

**Table 7** GIMBLE energy score versus experimental binding affinity

| | $R_1$ | $R_2$ | X | Gimble | delta *G* |
|---|---|---|---|---|---|
| H01 | C | COCCO | S | −21.151 | −7.32 |
| H02 | C | COCCC | S | −22.513 | −7.73 |
| H20 | C | CCCC | S | −23.743 | −8.4 |
| H07 | C | CC | S | −22.095 | −8.03 |
| H13 | C | COCCO | C | −21.003 | −6.52 |
| H14 | C | COCCO | O | −21.052 | −5.78 |
| H03 | C | COCC | S | −21.37 | −9.2 |
| H04 | C | COC | S | −20.643 | −8.06 |

## Cyclooxygenase 2 (COX-2)

Perhaps the greatest disappointment of the validation exercise was the inability of GIMBLE to accurately evaluate the binding interactions of celecoxib (vioxx) isosteres to its binding site of cyclooxygenase 2 (COX-2). Although only a single protein–ligand complex has public X-ray structures (PDB codes 6cox and 1cx2), the quantity of experimental binding data makes this system attractive for validation. COX-2 is a classical example of isosterism in drug-design with celecoxib, rofecoxib and valecoxib sharing an obvious shape similarity (often called the ''mickey mouse'' motif). An impressive total of 39 suitable isosteric (and near isosteric) analogs were identified in the supplementary material of the recent paper by Mozziconacci et al. [31] (Fig. 22).

For the evaluation, the coordinates of the ligand SC-558 in PDB code 1cx2 were used as a template structure. In addition to the usual substitutions, the sulfonamide and methylsulfonyl functional groups

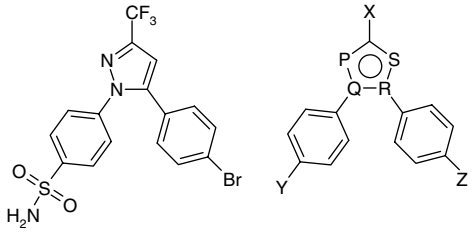



**Fig. 22** Structure of SC-558 and generic structure of analogues considered in this study

were investigated as terminal rotors. Hence for every isostere, three configurations were evaluated corresponding to the different rotations (or permutations) of the oxygen and nitrogen atoms using the experimental coordinates.

The resulting GIMBLE scores are presented in Table 8.

The ability of the proposed isostere evaluation methodology to perform well on this particular case study would have gone a long way in ensuring its utility. Unfortunately, there seems to no correlation with estimated binding scores and the experimental affinities. Again, there could be several possible causes of this failure. As before, the nature of this active site makes it a difficult target, with no strong electrostatic interactions and the majority of the significant mutations appearing in hydrophobic pockets. A significant concern is the quality of the structure of 1cx2 as a template. The structure itself is a particularly low resolution for structure based drug design, only resolved to 3.0 Å. The strained puckering of a phenyl ring that almost certainly should be planar is noticeable. The long bond length of the bromo substitution in SC-558 is probably a poor template for the chloro, fluoro, methyl, amino and hydrogen substitutions evaluated at that position. Even the analysis of the sulfonamide rotor indicated that the atomic assignments made in the original PDB file are probably incorrect. This agrees with the previous studies of COX-2s X-ray structures [35, 41]. Another noteworthy observation [19] is that the structure of 1cx2 was particularly highly restrained, almost homology like, in that molecular mechanics minimization of the protein structure results in almost no movement. Another plausible explanation of GIMBLE's poor results is the suggestion that much of the experimental data on the COX-2 system is performed in a non-aqueous phase assay [2].

**Table 8** GIMBLE energy score versus experimental binding affinity

| Ring # | X | Y | Z | Gimble | IC50 | pIC50 |
|---|---|---|---|---|---|---|
| 1 | H | SO2N | F | −18.177 | 0.014 | 7.85 |
| 1 | H | SO2C | F | −18.836 | 0.513 | 6.29 |
| 2 | H | SO2N | F | −18.21 | 10.233 | 4.99 |
| 3 | CF3 | SO2C | H | −19.473 | 0.120 | 6.92 |
| 3 | CF3 | SO2C | C | −20.309 | 0.158 | 6.8 |
| 3 | CF3 | SO2N | Cl | −19.059 | 0.010 | 8 |
| 3 | CF3 | SO2N | F | −18.877 | 0.010 | 8 |
| 3 | CF3 | SO2N | H | −18.83 | 0.040 | 7.4 |
| 3 | CF3 | SO2N | C | −19.684 | 0.040 | 7.4 |
| 3 | C | SO2C | Cl | −19.807 | 0.240 | 6.62 |
| 3 | CF3 | SO2C | Cl | −19.765 | 0.110 | 6.96 |
| 3 | CF2 | SO2C | Cl | −20.219 | 0.617 | 6.21 |
| 3 | CF | SO2C | Cl | −19.963 | 0.407 | 6.39 |
| 3 | CF3 | SO2C | F | −19.455 | 0.100 | 7 |
| 4 | CF3 | SO2C | F | −18.946 | 5.888 | 5.23 |
| 4 | CO | SO2C | Cl | −19.842 | 8.318 | 5.08 |
| 5 | CF3 | SO2C | H | −19.104 | 1.698 | 5.77 |
| 5 | CF3 | SO2C | C | −19.97 | 53.703 | 4.27 |
| 5 | CF3 | SO2N | H | −18.444 | 0.437 | 6.36 |
| 5 | CF3 | SO2N | C | −19.278 | 51.286 | 4.29 |
| 5 | C | SO2C | H | −19.166 | 79.433 | 4.1 |
| 6 | CF3 | SO2C | H | −19.213 | 1.514 | 5.82 |
| 6 | CF3 | SO2C | C | −20.041 | 1.288 | 5.89 |
| 6 | CF3 | SO2N | C | −19.351 | 0.724 | 6.14 |
| 7 | CO | SO2N | Cl | −19.148 | 0.832 | 6.08 |
| 7 | CF | SO2N | H | −18.844 | 0.200 | 6.7 |
| 7 | CF3 | SO2N | H | −18.23 | 0.032 | 7.49 |
| 7 | CF3 | SO2N | F | −18.226 | 0.041 | 7.39 |
| 7 | CF3 | SO2N | Cl | −18.415 | 0.010 | 8 |
| 7 | CF3 | SO2N | N | −18.708 | 0.339 | 6.47 |
| 7 | CF2 | SO2N | Cl | −18.848 | 0.010 | 8 |
| 7 | CF2 | SO2N | C | −19.515 | 0.013 | 7.89 |
| 7 | CF2 | SO2N | H | −18.688 | 0.129 | 6.89 |
| 7 | CF3 | Cl | Cl | −21.05 | 4.786 | 5.32 |
| 7 | C | SO2N | H | −18.208 | 63.096 | 4.2 |
| 7 | CF3 | SO2C | F | −18.88 | 0.100 | 7 |
| 7 | CF3 | SO2N | C | −19.093 | 0.040 | 7.4 |
| 8 | CF3 | SO2N | H | −17.975 | 45.709 | 4.34 |
| 9 | CF3 | SO2N | | −17.481 | 64.565 | 4.19 |

Hopefully, the future availability of a high-resolution protein–ligand structure and/or improved experimental data, will allow this system to be re-analyzed. As a consolation the docking scoring protocol proposed by Mozziconacci fares just as poorly on this difficult subset of their data set. Tantalizingly, the ordering of chloro and fluoro being better than methyl, being better than hydrogen, being better than amino is preserved in GIMBLE's scores, only in exactly reverse order, perhaps indicating a more systematic imbalance of electrostatic terms.

## Conclusions

We have presented a method for enumerating thousands to millions of isosteric analogues of the ligand from the crystal structure of a protein–ligand complex, and then efficiently calculating their electrostatic binding energies. The enumeration of molecules with the same shape (including tautomers and protonation states) allows for a dramatic simplification in the computation of the Poisson–Boltzmann binding energy. The performance resulting from this optimization allows the evaluation of many thousands of drug-like compounds per second without sacrificing significant accuracy.

In the initial experiments investigating methotrexate analogues in the active site of dihydrofolate reductase, this approach showed a surprising ability to reproduce known structure activity relationships. While this

method is not a magic bullet, and encounters problems on classically difficult binding sites, the validation reported here reveals that this technique works well on many docking systems. One of the major merits of this approach is that by minimizing many of the unknown variables in binding energy calculations, it is much easier to analyze, interpret and correct for the failures. As a result, many of the incorrect predictions can be understood, and enhancements to tackle the known deficiencies of VdW radius differences, proton placement and protein p$K_a$ can be tackled in follow-up investigations.

To assess the applicability of the proposed method, the ''frozen ligand hypothesis'' was tested on a number of model systems where the bound crystal structures of isosteric and near isosteric compounds were available. To summarize, the method performs well for truly isosteric series with significant electrostatic interactions (such as hydrogen bonding) to constrain the pose/conformation in a rigid active site. The procedure performs less well for flexible ligands of different shapes, in low resolution or flexible active sites with few constraining interactions, such as a poorly fitting hydrophobic pocket. This makes this procedure particularly well suited for ligand optimization once the natural substrate or a tightly binding lead compound has been co-crystallized in an enzymatic active site.

Two obvious questions that this work poses are whether Van der Waals interactions should be integrated into the binding energy evaluation so as to improve the predictive ability of the method, and is it possible to relax each isostere in the active site, perhaps using a constrained molecular mechanics force field, without losing the electrostatic signal?

Although the methods and results presented here describe evaluation of binding energies using the ZAP method for calculating electrostatic energies [13] and Halgren's MMFF94 bond charge increment method for assigning partial charges [15, 16], the same procedure can be used to benchmark and compare other scoring functions, such as molecular mechanics force fields [3, 4, 23] or empirical potentials [5, 32, 44], although potentially without the rapidity of the matrix formulation to electrostatics described here. In addition, other partial charging algorithms [1, 8, 9, 21, 46] or alternate electrostatics approaches [11, 18, 27, 33, 38, 40, 42] could be assessed against available experimental data.

**Acknowledgements** We thank Andrew Grant at AstraZeneca who kindly provided the crystal structure and $IC_{50}$ inhibition data for the Lck kinase case study, and Bob Tolbert for his work integrating WABE and GIMBLE into a single package. We also thank the anonymous reviewers, Geoff Skillman, Paul Hawkins and Christopher Bayly for their insightful comments. Finally, we would like to thank Juan Alvarez for a description in his recent book to an early draft of this work as ''an inspired experiment'' [37].

## References

1. Bayly CI, Cieplak P, Cornell WD, Kollman PA (1993) J Phys Chem 97:10269
2. Chris Bayly, Merck, Personal Communication, September 2005
3. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Roux B, Schlenkrick M, Smith JC, Stote R, Straub J, Wiorkiewicz-Kuczera J, Karplus M (1992) FASEB J 6:A143
4. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) J Am Chem Soc 117:5179
5. Eldridge D, Murray CW, Auton TR, Paolini GV, Mee RP (1997) J Comp Aided Mol Design 11(5):425
6. Ferrari AM, Wei BQ, Constantino L, Shoichet BK (2004) J Med Chem 47:5076
7. Fornabio M, Cozzini P, Mozzarelli A, Abraham DJ, Kellogg GE (2003) J Med Chem 46(21):4487
8. Gasteiger J, Marsili M (1978) A new model for calculating atomic charges in molecules. Tetrahed Lett 3181
9. Gasteiger J, Marsili M, (1980) Tetrahedron 36:3219
10. Gillet VJ, Johnson AP, Mata P, Sike S, Williams P (1993) J Comp Aided Mol Design 7:127
11. Gilson M, Sharp KA, Honig B (1987) Calculating the electrostatic potential of molecules in solution: method and error assessment. J Comput Chem 9327
12. Graffner-Nordberg M, Marelius J, Ohlsson S, Persson A, Swedberg G, Anderssin P, Andersson SE, Aqvist J, Hallberg A (2000) J Med Chem 43(21):3853
13. Grant JA, Pickup BT, Nicholls A (2001) J Comput Chem 22:608
14. Graves AP, Brenk R, Shoichet BK (2005) J Med Chem (JMC) 48(11):3714
15. Halgren TA (1996) J Comput Chem 17(5–6):490
16. Halgren TA (1996) J Comput Chem 17(5):553
17. Hansch C, Fukunaga JY, Jow PYC, Hynes JB (1977) J Med Chem 20:1
18. Hawkins GD, Cramer CJ, Truhlar DG (1996) J Phys Chem 100:19824
19. Hawkins, P (2005) OpenEye Scientific Software, Personal Communication, September 2005
20. Honig B, Nicholls A (1995) Science 268:1144
21. Jakalian A, Jack DB, Bayly CI (2002) J Comput Chem 23:1623
22. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) J Mol Biol (JMB) 207:727
23. Jorgensen WL, Maxwell D, Tirado-Rives J (1996) J Am Chem Soc 118:11225
24. Kangas E, Tidor B (1999) Phys Rev E 59(5):5958
25. Kangas E, Tidor B (2000) J Chem Phys 112(20):9120
26. Kangas E, Tidor B (2000) Electrostatic optimization in ligand complementarity and design. In: Floudas CA, Pardalos PM (eds) Optimization in computational chemistry and molecular biology. Kluwer Academic Publishers, pp 231–242
27. Madura JD, Briggs JM, Wade RC, Davis ME, Luty BA, Ilin A, Antosiewicz J, Gilson MK, Bagheri B, Scott LR, McCammon JA (1995) Comp Phys Commun 91:57

28. Mata P, Gillet VJ, Johnson AP, Lampreia J, Myatt GJ, Sike S, Stebbings A (1995) J Chem Inform Comp Sci 35:479
29. McGovern SL, Shoichet BK (2003) J Med Chem 46:2895
30. Miller DJ, Ravikumar K, Shen H, Suh J-K, Kerwin SM, Robertus JD (2002) J Med Chem 45(1):90–98
31. Mozziconacci J-C, Arnoult E, Bernard P, Do QT, Marot C, Morin Allory L (2005) J Med Chem 48(4):1055
32. Muegge I, Martin YC (1999) J Med Chem 45(5):791
33. Nicholls A, Honig B (1991) J Comput Chem 12:435
34. OEChem Theory Manual, Version 1.4, OpenEye Scientific Software, Santa Fe, New Mexico. http://www.eyesopen.com/
35. Plount Price ML, Jorgensen WL (2000) J Am Chem Soc (JACS) 122:9455
36. Rizzo RC, Udier-Blagovic M, Wang D-P, Watkins EK, Kroeger Smith MB, Smith RH, Tirado-Rives J, Jorgensen WL (2002) J Med Chem 45(14):2970
37. Rush Thomas III S, Manas Eric C, Tawa Gregory J, Alvarez Juan C (2005) Solvation-based scoring for high throughput docking, Chapter 10. In: Alvarez J, Shoicet B (eds) Virtual screening in drug discovery. Taylor & Franics Group, CRC Press, pp 249–277
38. Sharp K, Honig B (1990) Ann Rev Biophys Biophys Chem 19:301
39. Shen J, Quiocho FA (1995) J Comp Chem 16(4):445
40. Sitkoff D, Sharp KA, Honig B (1994) J Phys Chem 98:1978
41. Soliva R, Almansa C, Kalko SG, Luque FJ, Orozco M (2003) J Med Chem 46(8):1372
42. Still WC, Tempczyk A, Hawley RC, Hendrickson T (1990) J Am Chem Soc 112:6127
43. Vaidya CM, Wright JE, Rosowsky A (2002) J Med Chem 45(8):1690
44. Verkivker GM, Bouzida D, Gehlaar DK, Rejto PA, Arthurs S, Colson AB, Freer ST, Larson V, Luty BA, Marrone T, Rose PW (2000) J Comput Aided Mol Design 14:731
45. Wang Renxiao, Fu Y, Lai L (1997) J Chem Inform Comp Sci 37:615
46. Li J, Zhu T, Cramer CJ, Truhlar DG (1998) J phys chem 102:1820