Perspectives in Drug Discovery and Design, **20:** 145–169, 2000.
KLUWER/ESCOM

Hydrophobicity maps and docking of molecular fragments with solvation

NICOLAS MAJEUX, MARCO SCARSI, CATHERINE TENETTE-SOUAILLE and AMEDEO CAFLISCH*
Department of Biochemistry, University of Zürich, Winterthurerstraße 190, CH-8057 Zürich, Switzerland

Summary. Two methods for structure-based computational ligand design are reviewed. Hydrophobicity maps allow to quantitatively estimate and graphically display the propensity of nonpolar groups to bind at the surface of a protein target [Scarsi et al., Proteins Struct. Funct. Genet., 37 (1999) 565]. The program SEED (Solvation Energy for Exhaustive Docking) finds optimal positions and orientations of nonpolar fragments using the hydrophobicity maps, while polar fragments are docked with at least one hydrogen bond with the protein [Majeux et al., Proteins Struct. Funct. Genet., 37 (1999) 88]. An efficient evaluation of the binding energy, including continuum electrostatic solvation, allows to dock a library of 100 fragments into a 25-residue binding site in about five hours on a personal computer. Applications to thrombin, a key enzyme in the blood coagulation cascade, and the p38 mitogen-activated protein kinase, which is a target for the treatment of inflammatory and neurodegenerative diseases, are presented. The role of the hydrophobicity maps and structure-based docking of a fragment library in exploiting genomes to design drugs is addressed.

Key words: continuum electrostatics, docking of fragment library, generalized Born approximation, hydrophobicity map, molecular surface, p38 MAP kinase, SEED, solvation, thrombin

Introduction

Hydrophobicity is an important factor in molecular recognition [1–4] and the accurate prediction of the binding modes of nonpolar molecules to proteins in aqueous solvent is useful for ligand docking and drug design [5]. We have recently developed an approach to calculate and visualize the hydrophobicity at the surface of a protein (hydrophobicity maps) [6]. It is based on the evaluation of the nonpolar energy and electrostatic desolvation of the receptor with a continuum model. These energy contributions determine the binding modes of a nonpolar compound to the hydrophobic surface regions of a re-

* To whom correspondence should be addressed. E-mail: caflisch@bioc.unizh.ch

ceptor. Electrostatic interactions between ligand and receptor do not play a significant role in the case of a nonpolar ligand.

Computer programs for structure-based ligand design are useful tools for *de novo* design and lead modification [7–10]. The combinatorial strategy chosen for structure-based ligand design consists of three parts: the docking of molecular fragments, the connection of the docked fragments by combinatorial principles to generate candidate ligands, and the estimation of (relative) binding affinities [11]. The docking approach implemented in the program SEED determines optimal positions and orientations of small to medium-size molecular fragments in the binding site of an enzyme or receptor [12]. SEED docks polar fragments so that at least one hydrogen bond with optimal distance to a protein polar group is made. The hydrophobicity maps are used for the docking of apolar fragments. Our numerical continuum electrostatic methodology [13,14] and *ad hoc* look-up tables are employed to efficiently evaluate the protein and fragment desolvation upon binding and the screened electrostatic interaction. The fragments are then connected in a combinatorial way by the program CCLD [15]. For the final evaluation of candidate ligands, which is not discussed in this article, one could use a multi-layer scoring system that utilizes more than one binding affinity estimation method [16].

A recent review on computational approaches for drug design contains a detailed discussion of QSAR methods, structure-based docking and design programs, as well as a complete recapitulation of the available techniques to estimate (relative) binding affinities, i.e., from knowledge-based scoring functions to molecular dynamics-based free energy techniques [17]. The reader, interested in implicit solvation models is referred to two recent and comprehensive reviews [18,19].

Hydrophobicity maps

Methods

Hydrophobicity maps are a graphical representation of the binding energy of a nonpolar probe sphere rolling over the surface of the receptor. The binding energy includes both the electrostatic and nonpolar contributions to the association of a hydrophobic compound at the surface of a receptor. A continuum approach is used for the electrostatic component, whereas the van der Waals interaction describes the nonpolar contribution. The binding energy is displayed by color-rendering on the surface of the receptor. This yields a precise visualization of the surface hydrophobicity as well as a clear distinction between hydrophobic and hydrophilic zones in close proximity.

Binding energy of a nonpolar probe sphere at the surface of a receptor
The solvent accessible surface (SAS) is spanned by the center of a probe sphere rolling over the van der Waals surface of a molecule [20]. A number of points are distributed uniformly on the SAS of the receptor to describe in a discrete manner the different positions of the center of the probe sphere. On each of these points the binding energy of the nonpolar probe sphere (ΔE) is approximated, as explained in the next subsection, by the sum of van der Waals interaction energy (E^{vdW}) and electrostatic desolvation of the receptor ($\Delta E^{\mathrm{desolv}}$):

$$\Delta E = E^{\mathrm{vdW}} + \Delta E^{\mathrm{desolv}} \tag{1}$$

Parameters for the van der Waals energy and partial charges from standard force fields can be used. In the applications presented in this work, the all-hydrogen MSI CHARMm22 parameter set [21,22] was used.

The evaluation of ΔE for about 55 000 positions of the probe sphere on the thrombin surface requires about 35 s on a 195 MHz R10000 processor.

Van der Waals interaction energy The nonelectrostatic contributions to binding consist of the solute–solute van der Waals energy (favorable to binding), the loss of solute–solvent van der Waals energy (unfavorable), and the disruption of water structure which is a favorable entropic effect at room temperature [23]. A number of approaches have been proposed to evaluate these contributions [23–28]. Here it is assumed that solute–solvent van der Waals interactions and disruption of water structure compensate each other (see Reference 28 and Figure 6 of Reference 29), and that the solute–solute van der Waals energy can account for the nonelectrostatic component of the binding energy. Therefore, the van der Waals energy between the probe sphere and the receptor atoms (E^{vdW}) is assumed to account for all the nonelectrostatic contributions to the association of the probe sphere to the receptor. It is calculated as:

$$E^{\mathrm{vdW}} = \sum_{i \in \mathrm{receptor}} \sqrt{\varepsilon_i \varepsilon_{\mathrm{probe}}} \left\{ \left(\frac{R_i + R_{\mathrm{probe}}}{r_i} \right)^{12} - 2 \left(\frac{R_i + R_{\mathrm{probe}}}{r_i} \right)^{6} \right\} \tag{2}$$

where r_i is the distance between the receptor atom i and the probe sphere. ε_i and R_i are the van der Waals energy minimum and radius of atom i. The probe sphere van der Waals radius (R_{probe}) and energy minimum ($\varepsilon_{\mathrm{probe}}$) are input values. Since the probe sphere is rolled over the receptor van der Waals surface, it does not clash with it and is always at optimal distance from at least one receptor atom.

Electrostatic desolvation energy The electrostatic desolvation of the receptor accounts for the loss of receptor–solvent favorable electrostatic interactions due to the removal of part of the highly polarizable solvent to accommodate a nonpolarizable probe sphere. This contribution always disfavors association and can be calculated within the assumption of continuum electrostatics [13,14,30–35]. The system is partitioned into solvent and solute regions and two different dielectric constants are assigned to each region. The electrostatic energy E of the receptor in solution can be expressed in terms of the electric displacement vector $\vec{D}(\vec{x})$ and of a location dependent dielectric constant $\epsilon(\vec{x})$ as an integral over the three-dimensional (3D) space R^3 [36]:

$$E = \frac{1}{8\pi} \int_{R^3} \frac{\vec{D}^2(\vec{x})}{\epsilon(\vec{x})} d^3x \tag{3}$$

Since $\vec{D}(\vec{x})$ is additive, for point charges it can be rewritten as a sum over all charges i of the receptor:

$$\vec{D}(\vec{x}) = \sum_i \vec{D}_i(\vec{x}) \tag{4}$$

Concerning the electrostatics, docking a nonpolar sphere at the surface of the receptor has the only effect of modifying the dielectric properties in the space occupied by the sphere. Over this volume the dielectric constant changes from the solvent value (ϵ_w) to the solute value (ϵ_p). Usually, ϵ_w is set to 78.5 which is the value of water at room temperature, while the value of ϵ_p can range from 1 to 4. In the limit in which $\vec{D}(\vec{x})$ does not change significantly upon docking of the sphere, the variation of the electrostatic energy of the receptor (i.e., the desolvation) can be written according to Equation 3 as an integral over the volume occupied by the probe sphere (V_{probe}):

$$\Delta E^{\text{desolv}} = \frac{\tau}{8\pi} \int_{V_{\text{probe}}} \vec{D}^2(\vec{x})\, d^3x \tag{5}$$

where $\tau = 1/\epsilon_p - 1/\epsilon_w$. The volume occupied by the probe sphere is assumed to be a sphere of 1.7 Å radius, i.e., the van der Waals radius of the probe sphere augmented by 0.3 Å to include small voids between the probe and receptor surfaces. A 3D grid is built around the receptor and Equation 5 becomes:

$$\Delta E^{\text{desolv}} = \frac{\tau}{8\pi} \sum_{k \in V_{\text{probe}}} \vec{D}^2(\vec{x}_k) \Delta V \tag{6}$$

where ΔV is the volume of a grid cube and the index k runs over the grid points occupied by the probe sphere. The grid spacing is usually 0.5 Å. The electric displacement of every charge of the receptor can be approximated by the Coulomb field [13,34,37]:

$$\vec{D}(\vec{x}) = \sum_i q_i \frac{(\vec{x} - \vec{x}_i)}{|\vec{x} - \vec{x}_i|^3} \tag{7}$$

where $\vec{x}_i$ is the position of the receptor atom i and q_i its partial charge. Equation 7 is an analytical approximation of the total electric displacement and fulfills the condition of validity of Equations 5 and 6, i.e., $\vec{D}(\vec{x})$ is independent of the dielectric environment. The receptor desolvation in the Coulomb field approximation results from Equation 6 together with Equation 7:

$$\Delta E^{\text{desolv}} = \frac{\tau}{8\pi} \sum_{k \in V_{\text{probe}}} \left(\sum_i q_i \frac{(\vec{x}_k - \vec{x}_i)}{|\vec{x}_k - \vec{x}_i|^3} \right)^2 \Delta V \tag{8}$$

The accuracy of this approximation is discussed in the Methods section of SEED (see below).

It is important to note that the desolvation of a charged ion by a small nonpolar sphere at a distance r from the ion varies approximately as $1/r^4$ (Equation 8). This is a very short range effect compared with the ion electrostatic potential which varies as $1/r$. Hence, the potential alone cannot properly describe electrostatic desolvation.

Graphical rendering

The hydrophobicity is color-displayed over the molecular surface (MS) [38], which is traced by the surface of the probe sphere rolling over the van der Waals surface of the receptor. The MS consists of the convex receptor surface/probe contact areas and the concave (or reentrant) receptor surface/probe areas, and is preferred to the SAS because it gives a more precise description of the small details at the surface of the receptor. A smooth MS covering the receptor is generated via the molecular graphics package GRASP [23] as an ensemble of triangles. The hydrophobicity at each vertex is the value of the binding energy of the probe sphere (Equation 1) in the closest position to the vertex. This value is then visually displayed with the help of colors ranging from green (hydrophobic) through white (intermediate) to blue (hydrophilic).

Applications

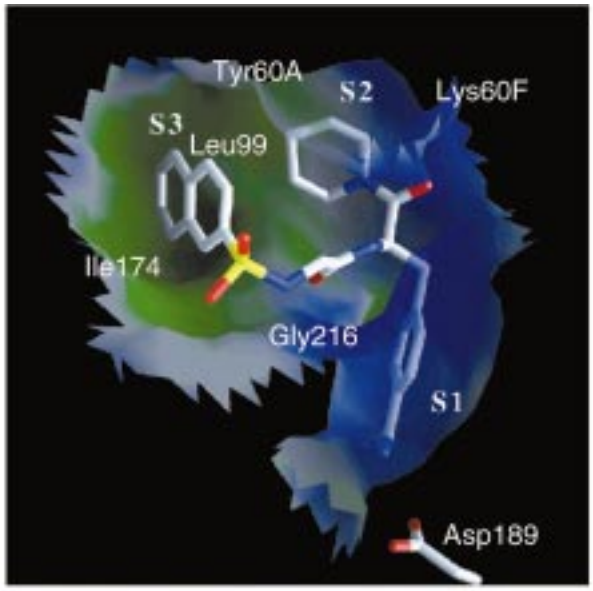
The hydrophobicity maps of thrombin and p38 mitogen-activated protein (MAP) kinase are presented here. The approach has been previously validated on 10 protein-ligand complexes [6]. In five of these complexes the ligand was a natural peptide or protein, while in the other five it was an organic compound with hydrophobic moieties. In all complexes tested up to date, the hydrophobicity maps correctly predict the regions of the receptor which are occupied by the nonpolar groups of the ligand.

Thrombin

Thrombin is a trypsin-like serine protease which fulfills an essential role in both haemostasis and thrombosis [39]. In the blood coagulation cascade, thrombin is the final enzyme that cleaves fibrinogen to release fibrinopeptides A and B and generate fibrin, which can then polymerize to form a haemostatic plug. The S3 and S2 precleavage subpockets of the active site have a hydrophobic character, whereas at the bottom of the S1 or recognition pocket the carboxyl group of Asp189 is a salt bridge partner for basic side chains. Nα-((2-naphthylsulfinyl)glycyl)-DL-*p*-amidinophenylalanylpiperidine (NAPAP) is an archetypal active site inhibitor of thrombin (Figure 1a). It fills the S3 and S2 pockets with its naphthalene and piperidine groups, respectively. Moreover, it is anchored by its basic group (benzamidine) into S1 to form a salt bridge with Asp189 [40].

The hydrophobicity map of the nonprime region of the thrombin active site is shown in Figure 1a. The side chains of Asp189 and NAPAP are also shown. S3 and part of S2 are identified as hydrophobic, while S1 shows a hydrophilic character. The naphthalene and piperidine groups of NAPAP are in contact with hydrophobic zones and bury 13 of the 100 most hydrophobic points (over a total of about 55 000) on the SAS of thrombin. The polar groups of NAPAP bind to the hydrophilic zones of the thrombin active site: the Asp189 side chain and the Gly216 backbone polar groups. The energy loss of removing water from the hydrophilic zones is compensated upon binding by favorable electrostatic ligand-receptor interactions. The lower part of S1, despite being a narrow concave cavity, is identified as hydrophilic, since electrostatic desolvation of Asp189 dominates over the favorable vdW interactions between the probe sphere and the surrounding thrombin atoms. The curvature mapped on the MS (not shown) does not take into account thrombin electrostatic desolvation and suggests that the bottom of the S1 pocket is the most hydrophobic region, in contrast with the actual binding mode of NAPAP.

(a)



(b)

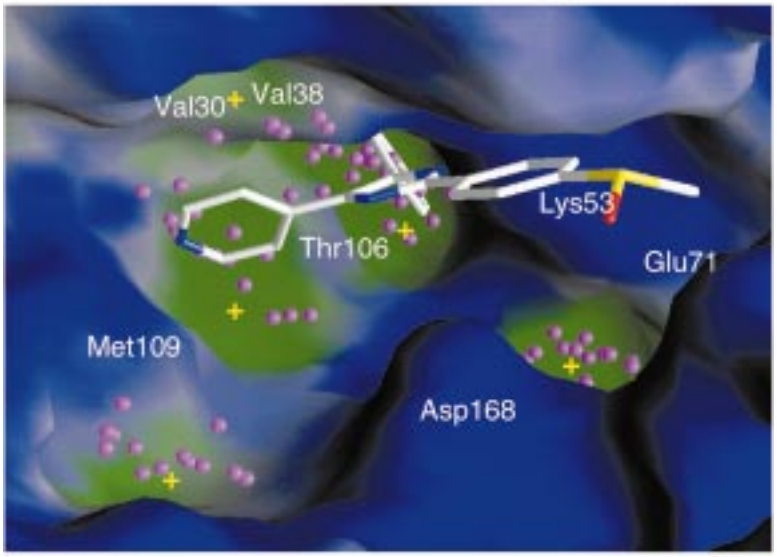


Figure 1. Hydrophobicity maps calculated with Equations 1, 2 and 8, and displayed with the program GRASP [23]. The molecular surface is displayed with colors ranging from green (hydrophobic) through white (intermediate) to blue (hydrophilic) according to the hydrophobicity map value. (a) Thrombin-NAPAP complex [40]. The transparent MS of the thrombin active site is displayed together with the side chain of Asp189 and NAPAP in a cylinder model (carbon atoms are white, nitrogen blue, oxygen red, and sulfur yellow). (b) Complex of the p38 MAP kinase and the triarylimidazole SB203580 (PDB code 1A9U) [48]. The 70 most hydrophobic points on the ATP binding site surface are displayed in magenta. Yellow crosses mark the five hydrophobic regions discussed in the text.

p38 MAP kinase

MAP kinases are essential enzymes for intracellular signalling cascades because they phosphorylate several regulatory proteins. They are responsive to hormones, cytokines, environmental stresses and other extracellular stimuli, and are activated by a dual phosphorylation of a threonine and tyrosine in the TXY motif in the so-called phosphorylation lip. p38 MAP kinase (or CSBP2) plays a role in processes as diverse as transcriptional regulation, production of interleukins, and apoptosis of neuronal cells [41–44]. Inhibitors of p38 activity could therefore be useful as a treatment strategy for inflammatory and neurodegenerative diseases. The CSAID^{TM} (cytokine suppressive anti-inflammatory drugs) class of anti-inflammatory compounds inhibits the synthesis of cytokines, such as interleukin-1 and tumor necrosis factor, by specific inhibition of the MAP kinase p38 [41,45,46]. They have a common chemical pattern: A central five-membered ring, either imidazole or pyrrole, substituted by a pyridine or a pyrimidine ring, a fluorinated or iodinated phenyl ring, and a third substituent at position 1 or 2 (Figures 1b and 6). These low-molecular weight inhibitors and their analogs bind to the ATP-binding cleft of the inactivated form of p38 and are competitive with respect to ATP. They are potent inhibitors, with IC_{50} in the nanomolar range [45,47], and highly selective for p38 compared to the other MAP kinases.

In Figure 1b the most hydrophobic points in the ATP binding site are displayed together with the triarylimidazole inhibitor SB203580 [48]. Five hydrophobic regions of concave shape are found by the computational approach described above. They are colored in green in Figure 1b and their approximate center is marked by a yellow cross. Three regions are consistent with the available structural data of p38 MAP kinase/inhibitor complexes [48,49], whereas two regions are novel. The most hydrophobic pocket is located between the Thr106 and Lys53 side chains, and is occupied by the phenyl group of the diaryl- and triarylimidazole inhibitors. The hydrophobic pocket lined by the Thr106 and Met109 side chains is occupied by the pyridine or pyrimidine cycle. In the diarylimidazole inhibitors, the N-substituent of the central imidazole is in contact with the hydrophobic region close to the Val30 and Val38 side chains [48]. Surprisingly, the hydrophobic pockets below Glu71 and Met109 are empty in the available crystal structures of the MAP kinase p38/inhibitor complexes [48,49]. For the inhibitors known to bind at the ATP site, it is expected that additional nonpolar substituents directed towards the two unoccupied hydrophobic pockets will improve the binding affinity.

Docking a fragment library with SEED

Methods

The different types of fragments are docked by SEED in the order specified by the user. After each fragment placement the binding energy is estimated. The binding energy is the sum of the van der Waals interaction and electrostatic energy with continuum solvation. The successive fragment type is docked, after all placement-energy evaluations of the preceding fragment type have been made. The fragment docking procedure and energy evaluation are outlined in this section. Further details of the method, e.g., the clustering procedure, are given in the original paper [12]. For the docking of a library of 100 fragments into a binding site of about 25 residues, the latest version of SEED requires about 5 h of CPU time on a single processor (195 MHz R10000 or PentiumIII 550 MHz). For more than one processor the speed-up is linear so that the docking of a library of 1000 fragments would require about 6 h of an 8-processor server or cluster.

Fragment docking
The binding site of the receptor is defined by a list of residues, which are selected by the user. Fragments are considered polar if they have at least one H-bond donor or acceptor. Due to this definition some 'polar' fragments can have considerable hydrophobic character (e.g., diphenylether). Therefore they can also be docked by the procedure for nonpolar fragments if specified by the user.

Docking of polar fragments These are docked so that one or more hydrogen bonds with the receptor are formed. The fragment is then rotated around the H-bond axis to increase sampling. Figure 2a shows the sampling of docked positions for pyrrole and acetone around a tyrosine side chain. Ideal and close-to-ideal hydrogen bond geometries are sampled in a discrete but exhaustive way.

Docking of nonpolar fragments The hydrophobicity maps are used to dock nonpolar fragments. The points on the receptor SAS are ranked according to the sum of van der Waals interaction and receptor desolvation (Equation 1), and the best n points (where n is an input parameter) are selected for docking. As an illustrative example, Figure 1b shows the 70 most hydrophobic points on the ATP binding site of the p38 MAP kinase.

For both the fragment and the receptor, vectors are defined by joining each point on the SAS with the corresponding atom. Finally, nonpolar fragments are docked by matching a vector of the fragment with a vector of the receptor

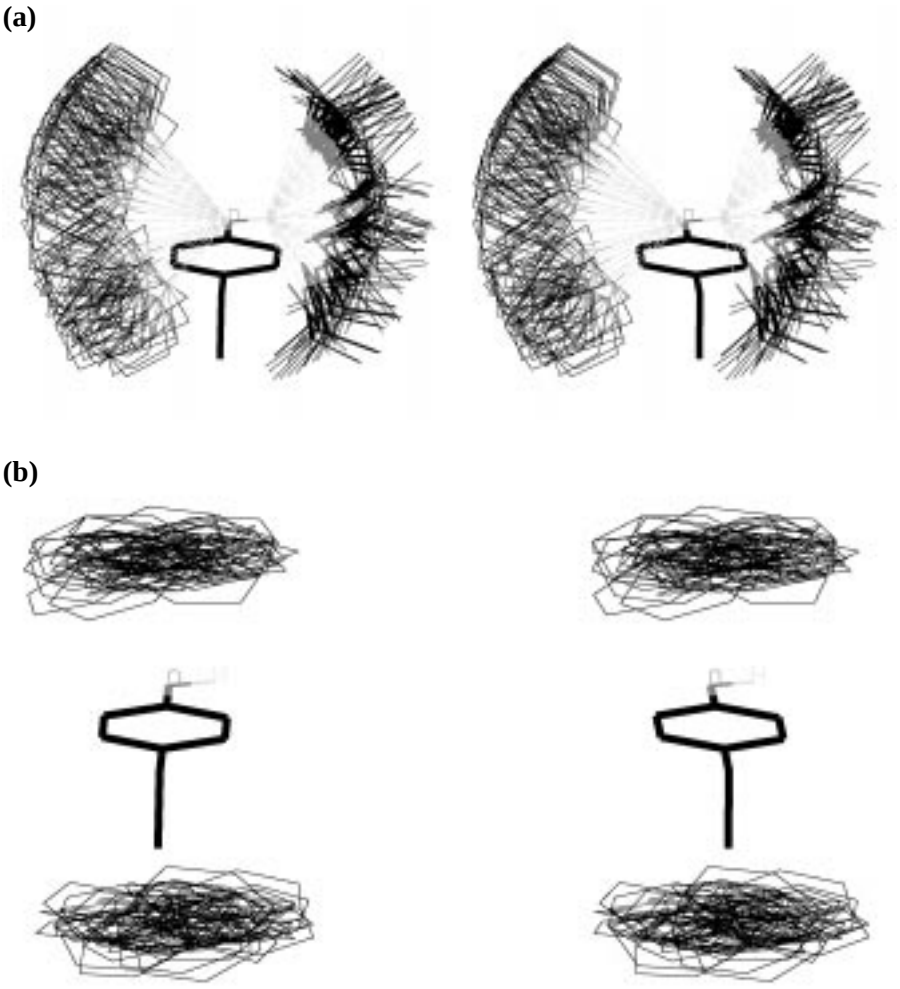


Figure 2. Relaxed-eyes stereoview of the fragments docked by SEED around a tyrosine side chain. (a) Acetone and pyrrole, (b) benzene. Carbon atoms are black, oxygen and nitrogen atoms dark gray, and hydrogen atoms light gray. Hydrogen bonds are drawn with dashed lines.

at the optimal van der Waals distance. To improve sampling, additional rotations of the fragment are performed around the axis joining the receptor atom and fragment atom (Figure 2b).

To increase efficiency, nonpolar fragments are discarded without calculation of the electrostatic energy, if the van der Waals interaction is less favorable than a threshold value.

For both polar and nonpolar fragments, the docking is exhaustive on a discrete space. The discretization originates from the limited number of preferred directions and rotations around them. Fragment symmetries are checked only once for every fragment type and are exploited to increase the efficiency in docking.

Electrostatic energy with continuum solvation
The main assumption underlying the evaluation of the electrostatic energy of a fragment-receptor complex is the description of the solvent effects by continuum electrostatics [13,14,30–35,50–52]. The system is partitioned into solvent and solute regions and different values of the dielectric constant are assigned to each region. In this approximation only the intra-solute electrostatic interactions need to be evaluated. This strongly reduces the number of interactions with respect to an explicit treatment of the solvent. Moreover, it makes feasible the inclusion of solvent effects in docking studies where the equilibration of explicit water molecules would be a major difficulty. The electrostatic effects of the solvent are relevant and it has been shown that the continuum dielectric model provides an accurate description of molecules in solution [14,53]. The difference in electrostatic energy in solution upon binding of a fragment to a receptor can be calculated as the sum of the following three terms [15,50]:

- Desolvation of the receptor: Electrostatic energy difference upon binding the uncharged (all partial charges switched off) fragment to the charged receptor in solution.
- Screened fragment-receptor interaction: Electrostatic interaction energy between the fragment and the receptor in solution.
- Desolvation of the fragment: Electrostatic energy difference upon binding the charged fragment to the uncharged (all partial charges switched off) receptor in solution.

The definition of the solute volume, i.e., the low dielectric volume, is central in the evaluation of these energy terms with a continuum model. The solute–solvent dielectric boundary is described by the molecular surface (MS) of the solute [38]. A grid covering the receptor is set up. In a first step the volume occupied by the isolated receptor is defined on the grid. Subsequently for every position of a docked fragment the volume enclosed by the MS of the fragment-receptor complex is identified.

The screened fragment-receptor interaction and the fragment desolvation are evaluated with a grid-based implementation [13,14] of the generalized Born (GB) approximation [31–35]. The GB approach would be too time con-

suming for the evaluation of the desolvation of the receptor which is calculated by the procedure described in the Methods section of the hydrophobicity maps.

Receptor desolvation It is evaluated using Equation 8 where the index k runs over the grid points in the volume occupied by the fragment. The volume occupied by a docked fragment is the part of the volume enclosed by the MS of the complex that was not occupied by the isolated receptor. It consists of the actual volume of the fragment and the interstitial volume enclosed by the reentrant surface between fragment and receptor.

Screened fragment-receptor interaction The fragment-receptor interaction in solution is calculated via the GB approximation [31]. In a solvent of dielectric constant ϵ_w, the interaction energy between two charges embedded in a solute of dielectric constant ϵ_p is

$$E_{ij}^{int} = \frac{q_i q_j}{\epsilon_p r_{ij}} - \frac{q_i q_j \tau}{R_{ij}^{GB}} \tag{9}$$

where $\tau = 1/\epsilon_p - 1/\epsilon_w$,

$$R_{ij}^{GB} = \sqrt{r_{ij}^2 + R_i^{eff} R_j^{eff} exp\left(\frac{-r_{ij}^2}{4R_i^{eff} R_j^{eff}}\right)} \tag{10}$$

and q_i is the value of the partial charge i, while r_{ij} is the distance between charges i and j. R_i^{eff} is the effective radius of charge i and it is evaluated numerically on a 3D grid covering the solute as described in Reference 13. It is a quantity depending only on the solute geometry and represents an estimate of the average distance of a charge from the solvent.

The intermolecular interaction energy is calculated as:

$$E^{int} = \sum_{\substack{i \in \text{fragment} \\ j \in \text{list}_i}} E_{ij}^{int} \tag{11}$$

where list_i contains the receptor atoms belonging to the neighbor list of atom i. The electrostatic neighbor list includes all the receptor atoms of the van der Waals neighbor list and one atom for every charged residue whose charge center is within a distance of 13 Å from the closest binding site residue. Supplementing the van der Waals neighbor list with a monopole approximation of distant charged residues dramatically reduces the error originating from the long range electrostatic interactions.

Fragment desolvation The fragment intramolecular energy in solution is calculated with the GB formula as described in Reference 13:

$$E = \sum_{i \in \text{fragment}} E_i^{\text{self}} + \sum_{\substack{i>j \\ i,j \in \text{fragment}}} \left(\frac{q_i q_j}{\epsilon_p r_{ij}} - \frac{q_i q_j \tau}{R_{ij}^{GB}} \right) \tag{12}$$

where the two sums run over the partial charges of the fragment. Equation 12 differs from Equation 11 due to the presence of the *self-energy* term $\sum_i E_i^{\text{self}}$. This term is not zero only in the case of intramolecular energies. E_i^{self} is the *self-energy* of charge i and represents the interaction between the charge itself and the solvent. It is calculated as [13,34]:

$$E_i^{\text{self}} = \frac{q_i^2}{2R_i^{\text{vdW}}\epsilon_p} - \frac{q_i^2 \tau}{2R_i^{eff}} \tag{13}$$

where R_i^{vdW} is the van der Waals radius of charge i.

The difference in the intramolecular fragment energy upon binding to an uncharged receptor in solution is:

$$\Delta E = E^{\text{docked}} - E^{\text{free}} \tag{14}$$

where E^{docked} and E^{free} are the energies in solution of the fragment bound and unbound to the receptor, respectively. They are evaluated according to Equation 12. For the unbound fragment (E^{free}) the effective radii are calculated considering as solute the volume enclosed by the molecular surface of the fragment. For the bound fragment (E^{docked}) the solute is the volume enclosed by the molecular surface of the receptor-fragment complex. E^{free} is evaluated only once per fragment type, while E^{docked} is recalculated for every fragment position in the binding site.

Validation

The approximations inherent to our continuum electrostatic approach were validated by comparison with finite difference solutions of the Poisson equation [12]. For this purpose, the three electrostatic energy terms were calculated with SEED and UHBD [52,54] for a set of small molecules and ions distributed over the binding site of thrombin and at the dimerization interface of the HIV-1 aspartic protease monomer. The molecule set included acetate ion, benzoate ion, methylsulfonate ion, methylammonium ion, methylguanidinium ion, 2,5-diketopiperazine, and benzene. The total number of fragment-receptor complexes analyzed were 1025 for thrombin (Figure 3) and 1490 for the HIV-1 protease monomer. The agreement between the two

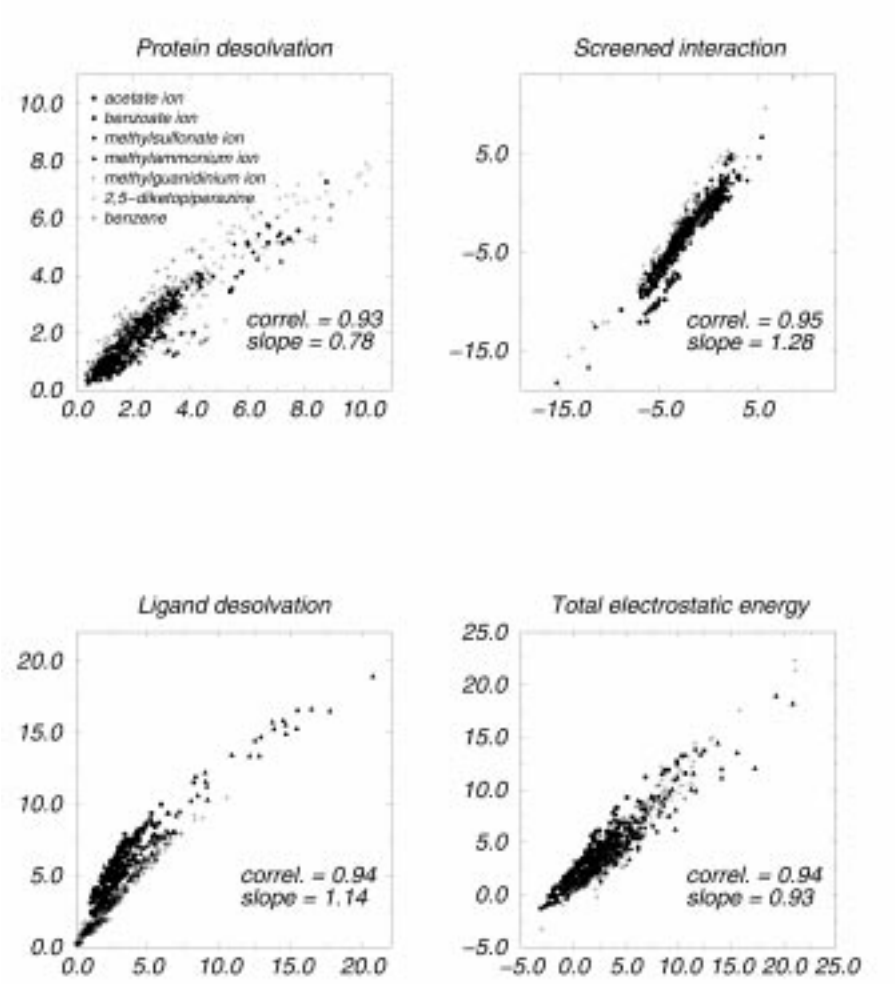


Figure 3. Correlations in the electrostatic energies calculated by finite difference solution of the Poisson equation (x-axis) and SEED (y-axis). Values are plotted for 1025 complexes of thrombin with small molecules. The total electrostatic energy is the sum of the protein desolvation, screened interaction, and ligand desolvation scaled by 0.78, 1.28, and 1.14, respectively. The finite difference calculations were performed with the program UHBD [52,54]. An interior dielectric of 4, solvent dielectric of 78.5, and grid spacing of 0.5 Å were used for both SEED and UHBD.

methods is very good, and better for a solute dielectric constant of 4.0 (Figure 3) than 1.0 (Figure 2 of Reference 12). In Table 2 of Reference 12 it was shown that systematic errors (slope $\neq 1$) are independent of the receptor and

the solute dielectric constant and consequently can be corrected by the use of appropriate scaling factors for the different energy terms.

Application to thrombin

Apart from a relatively small rigid body motion of the Tyr60A-Trp60D loop, thrombin assumes the same conformation in complexes with different inhibitors [55,56]. Figure 4 shows the most relevant results of a SEED run (interior dielectric of 1.0) on the nonprime region of the thrombin active site while a more detailed description of the SEED functionality maps of thrombin is given in Reference 12. The hydrophobic fragments bind preferentially to the S3 and S2 pockets (Figure 4a). The +1 charged groups, e.g., benzamidine (Figure 4b), 5-amidine indole (Figure 4c) and methylguanidinium, are involved in optimal hydrogen bonds with the Asp189 side chain in the S1 pocket. The SEED results are in agreement with the large amount of structural data on thrombin/inhibitor complexes [10,39,40,55–61].

The fragments docked by SEED were then connected by the program CCLD [15] as a further test of the usefulness of SEED for ligand design. CCLD generated 390 candidate ligands in 33 min on a R10000 processor. Four interesting hits, which ranked as 2nd, 37th, 42nd and 90th, are shown in Figures 5c–f.

Hits **1** and **2** (Figures 5c and d) are similar to Argatroban (Figure 5b), which is a reversible inhibitor of thrombin with a K_i of 19 nM [40,62]. They have nonpolar groups in S3 and S2 and a guanidinium in S1. The sulfonamide NH of compounds **1**, **2**, and Argatroban donates a hydrogen bond to the backbone CO of Gly216. Moreover, the carbonyl group in **2** and Argatroban accepts from the NH group of Gly216. Additional hydrogen bonds, with respect to Argatroban, are present in **1** and **2**, namely between an SO_2 oxygen and the NH group of Gly219, and between the guanidinium and the main chain CO of Gly219. Furthermore, the amide group close to the guanidinium in compound **1** donates to the carbonyl of Ser214. Argatroban has less polar interactions with thrombin than hits **1** and **2** but its double ring moiety fills the S3 pocket better than the cyclohexyl ring of compounds **1** and **2**. Hits **3** and **4** (Figures 5e and f) have a benzamidine in the S1 pocket and a benzene in S2. The benzamidine moiety of compound **3** donates to the two oxygens of the Asp189 side chain and to the carbonyl group of Gly219. In the S3 pocket the hydroxyl substituent of cyclohexane donates to the main chain CO of Glu97A. Compound **4** is similar to 4-TAPAP (Figure 5a), a reversible inhibitor of thrombin [40,63] whose racemic mixture has a K_i of 640 nM. In 4-TAPAP and hit **4** the benzamidine is involved in a salt bridge with the Asp189 side chain, the sulfonyl part accepts from the NH group of Gly219 and one NH donates to the carbonyl group of Gly216. The interaction with

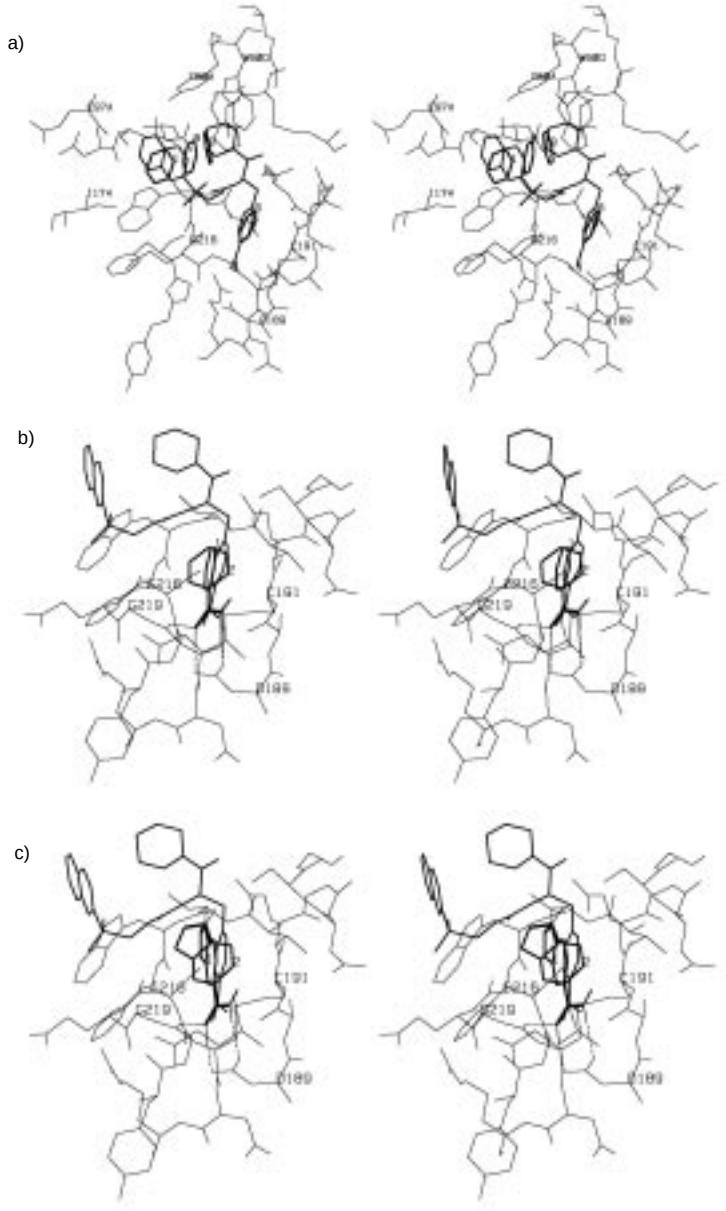


Figure 4. (a) Relaxed-eyes stereoview of the SEED cluster representatives of benzene (thick lines) in the thrombin active site (thin lines). The NAPAP inhibitor is also shown (medium lines), though it was removed during the SEED procedure. The SEED cluster representatives are labeled according to their binding energy rank within representatives of the same type. (b) Same as in (a) for benzamidine. Hydrogen bonds between protein and ligands are shown with dashed lines. (c) Same as in (b) for 5-amidine indole. Reprinted with permission from [12].

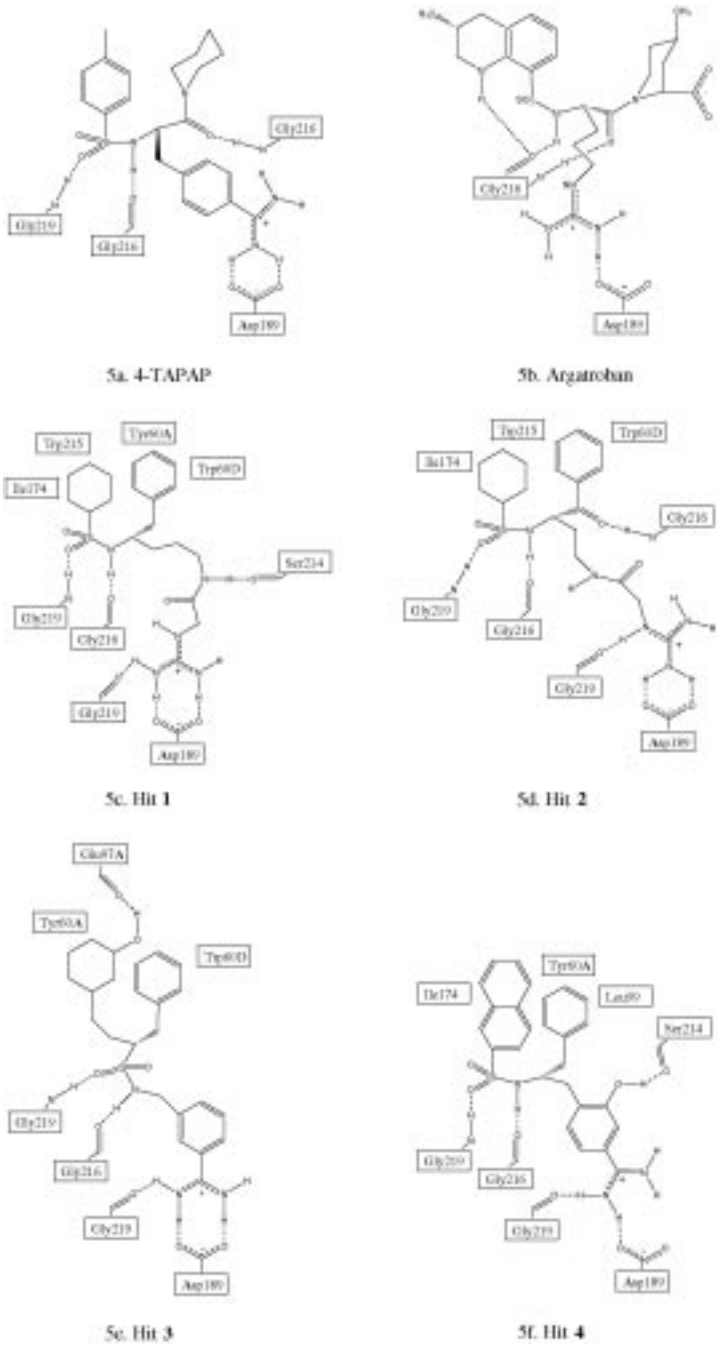


Figure 5. Schematic representations of the interactions between thrombin and (a) 4-TAPAP, (b) Argatroban, and (c–f) CCLD hits **1** to **4**. Reprinted with permission from [12].

Table 1. SEED results for the p38 MAP kinase

Rank of cluster[a]	Intermolecular		Electrostatic desolvation		$\Delta G_{binding}$[b]	Site[c]
	vdWaals	Elect	Receptor	Fragment		
Benzene						
1	–14.7	–0.6	3.4	0.3	–11.6	**Phenyl**
2	–10.1	0.0	2.0	0.3	–7.8	**Pyridine**
3	–11.3	0.3	3.1	0.3	–7.7	Phe169
4	–10.7	0.5	2.8	0.3	–7.2	Phe169
5	–11.0	0.1	3.6	0.3	–7.0	**Phenyl**
Pyridine						
1	–9.1	–0.6	1.4	0.8	–7.4	**Pyridine**
2	–8.5	–2.0	4.1	0.8	–5.6	Lys53
3	–7.6	0.1	1.3	0.8	–5.4	**Pyridine**
4	–9.2	–1.6	4.9	0.8	–5.1	Lys53
5	–9.1	–2.2	6.1	0.8	–4.4	Lys53

[a] Each cluster contains 10 fragment positions. Energy values (in kcal mol^{-1}) are given for the cluster representative which has the most favorable binding energy among the 10 members of the cluster. Cluster 1 is shown in Figure 6.
[b] Sum of the values in the four preceding columns, i.e., intermolecular interactions and electrostatic desolvation energies.
[c] The site is defined by the substituent of the triarylimidazole inhibitor (in boldface) or the closest side chain of the p38 MAP kinase.

NH of Gly216 is missing in **4** but there are two additional hydrogen bonds, with the CO groups of Ser214 and Gly219. This last hydrogen bond also occurs in the NAPAP thrombin complex. The naphthalene ring of **4** fills the S3 pocket as NAPAP [55]. These representative examples and visual analysis of other SEED-CCLD hits indicate that the present approach generates candidate ligands with interaction patterns similar to known thrombin inhibitors.

Application to the p38 MAP kinase

The SEED maps of benzene and pyridine were obtained with an interior dielectric of 4, solvent dielectric of 78.5, and grid spacing of 0.5 Å. A complete analysis of the results for a library of about 100 fragments and a comparison with the available structural data of p38 MAP kinase/inhibitor complexes will be given elsewhere (Tenette-Souaille et al., manuscript in preparation).

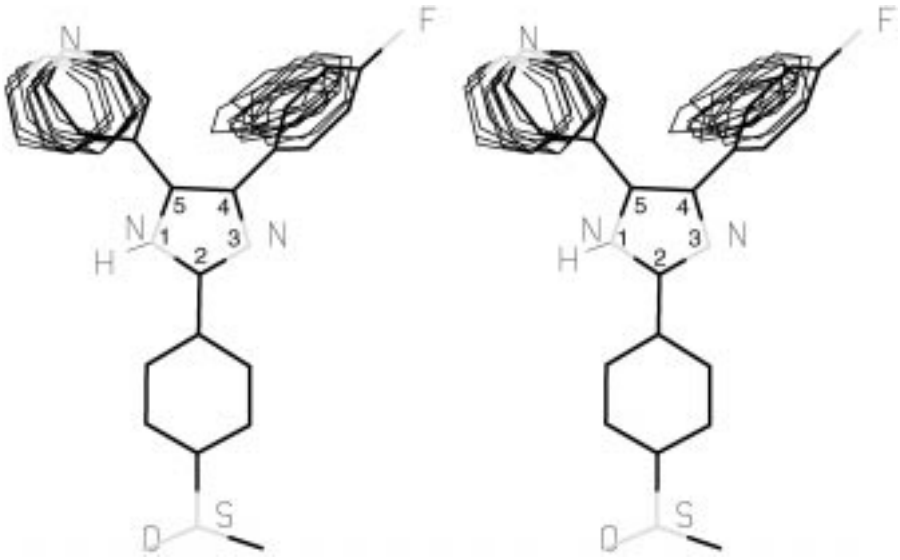


Figure 6. Relaxed-eyes stereoview of the 10 best benzenes and pyridines docked by SEED into the p38 MAP kinase. The bound conformation of the SB203580 inhibitor [48] is displayed to show that the 10 best benzenes and pyridines match the corresponding groups of SB203580.

Benzene The first cluster of benzene occupies the hydrophobic pocket, which contains the phenyl group of the known triarylimidazole inhibitors. The orientation of the members of the first cluster is similar to that observed in the crystallographic structure (Figure 6). A remarkable gap of 3.8 kcal mol^{-1} in the binding energy is found between the representative of the first cluster and the other cluster representatives (Table 1). Furthermore, even the other nine members of the first cluster have a more favorable energy than the representative of the second cluster. This difference is mainly due to the very favorable van der Waals term in the first cluster. The representatives of clusters 2 to 5 display similar energy values. Most of the fragments containing a phenyl ring (e.g., naphthalene, tetraline, N-methyl indole, and dibenzocyclohexane) match the phenyl group of the known triarylimidazole inhibitors. Moreover, there is a large energy gap (from 2.5 to 4.0 kcal mol^{-1}) between the first cluster and the other clusters.

Pyridine Pyridine, as well as other fragments containing an aromatic ring with a hydrogen bond acceptor, overlaps the pyridine substituent of the triarylimidazole inhibitors (Figure 6). The orientation of the members of the first cluster is very close to that of the known inhibitors and they are involved in a hydrogen bond with the backbone NH of Met109 as in the crystal structure [49]. The main chain NH of Met109 is indeed the privileged partner of

fragments with a hydrogen bond acceptor. As in the case of benzene, there is a significant gap (1.8 kcal mol^{-1}) between the first cluster representative of pyridine and the second cluster (Table 1).

Conclusions

A procedure to determine hydrophobic regions on the surface of a protein and a continuum electrostatic approach for the accurate and efficient docking of a fragment library have been presented. The hydrophobicity maps allow easily to discriminate between hydrophobic and hydrophilic surface regions that are close in space. This was illustrated for the thrombin–NAPAP and p38 MAP kinase–triarylimidazole inhibitor complexes, whose hydrophobicity maps can be used for de novo design and lead modification. Furthermore, for the thrombin–NAPAP and farnesyltransferase–farnesylpyrophosphate complexes it was shown previously that existing approaches based on the analysis of the surface curvature and/or the electrostatic potential are not as valuable in distinguishing regions where nonpolar and polar groups can bind [1]. The fraction of the most hydrophobic receptor regions that are buried at the binding interface is in general particularly high, suggesting that hydrophobic association is determinant in protein-ligand binding. This confirms previous findings [6,7].

A number of very efficient docking programs have been published recently. Most of them use either a scoring function with a crude approximation of solvation [64–66] or a *vacuum* energy derived from a molecular mechanics force field [67,68]. The program DOCK, which pioneered the use of geometric criteria to select ligands which best complement the shape of the receptor site [69,70], has been supplemented by the evaluation of ligand desolvation [71,72]. To efficiently screen large databases of compounds, DOCK assumes that every ligand desolvates the receptor equally and that the ligand is completely desolvated upon binding. The continuum electrostatic approach implemented in SEED does not make these assumptions.

There are two main advantages in SEED with respect to the multiple copy simultaneous search method (MCSS) which is a force field-based approach for determining optimal positions and orientations of functional groups in a protein binding site [73–78]. These are the inclusion of electrostatic solvation and the determination of all favorable binding modes. The effects of the solvent are neglected in MCSS which calculates the protein-fragment interactions with a *vacuum* potential [21]. This choice in MCSS was based on the principle that fast methods are necessary to perform effective searches of the binding site and that good candidate ligands subsequently can be ranked in terms of their binding energy. A possible difficulty is that minimized positions

may be missed or misplaced due to the lack of a solvation correction during the MCSS minimization. The best energy minima without solvation do not always turn out to be those of most interest [76]. Particularly problematic is the docking of apolar fragments which, without inclusion of solvation, are positioned in both hydrophilic and hydrophobic regions of the binding site. This problem is solved in SEED by the prioritization of apolar regions on the protein surface according to low electrostatic desolvation and favorable van der Waals interactions, as well as the efficient use during docking of a protein desolvation look-up table.

SEED samples optimal binding modes and can also find positions which do not necessarily correspond to local minima of the energy function (e.g., a favorable region with relatively flat potential energy in between two well pronounced minima). This is an advantage with respect to MCSS because not all of the molecular fragments, even in potent inhibitors, have optimal interactions with the protein [76].

Fragments are docked as rigid bodies by SEED. For larger ligands with rotatable bonds, conformational flexibility can be taken into account by docking different conformations. Programs are available for the automatic generation of diverse low-energy conformations of small molecules [79,80]. In the case of large ligands with many rotatable bonds one could use SEED to find optimal positions for the rigid moieties and then use other techniques which allow for full ligand and eventually also protein flexibility [81].

Future perspectives

The two computational approaches presented in this article, hydrophobicity maps and SEED, were developed to precisely determine hydrophobic pockets and to dock fragment libraries. The hydrophobicity maps will be useful for the characterization of binding sites for the 3D structures [82] of the large amount of sequences that are emerging from the many genome projects. The location of a binding or association site can be predicted from the clusters of most hydrophobic points on the surface, and the size and ligand type could be estimated from the area and/or the volume of the binding site cleft.

Drug design is a fully multidisciplinary research field. Methodologies and procedures from different scientific disciplines support and cross-fertilize each other. A typical example is the combinatorial strategy for fragment-based design which is common to SAR by NMR [83,84], the multiple solvent crystal structures method [85], and SEED-CCLD [2,15]. A better understanding of the physical principles of solvation is needed to help in designing drugs. Continuum models of solvation effects are particularly useful for docking; their use will grow significantly in the near future. The continuum electro-

static approach implemented in SEED allows to efficiently dock a library of molecular fragments to a receptor of known structure. The SEED-CCLD strategy uses combinatorial principles to construct candidate ligands. Possible applications are for *de novo* design, as documented here for thrombin, lead optimization, and the selection of monomers for parallel synthesis and combinatorial chemistry.

Although there is not yet a computational approach to step directly from genomes to drugs [86], we think that the methods and procedures described in this issue of *Perspectives in Drug Discovery and Design* are useful new developments for drug discovery.

Acknowledgements

We thank S. Ahmed, J. Apostolakis, N. Budin and C. Ehrhardt for helpful discussions, and A. Widmer (Novartis Pharma Inc., Basel) for the molecular modelling program WITNOTP which was used for visual analysis of the results and for preparing the Figures 2, 4, and 6. This work was supported by the Swiss National Science Foundation (Nationalfonds, Grant No. 31-53604.98), the Helmut Horten Foundation, and the EMDO Foundation. The program SEED (for SGI or PC running the Linux operating system), as well as the library of fragments in mol2 format, is available for non profit institutions from the last author (e-mail: caflisch@bioc.unizh.ch).

References

1. Janin, J. and Chothia, C., J. Biol. Chem., 265 (1990) 16027.
2. Janin, J. and Chothia, C., Biochemistry, 17 (1978) 2943.
3. Horton, N. and Lewis, M., Protein Sci., 1 (1992) 169.
4. Young, L., Jernigan, R.L. and Covell, D.G., Protein Sci., 3 (1994) 717.
5. Davis, A.M. and Teague, S.J., Angew. Chem. Int. Ed., 38 (1999) 736.
6. Scarsi, M., Majeux, N. and Caflisch, A., Proteins Struct. Funct. Genet., 37 (1999) 565.
7. Veerapandian, P., Structure-Based Drug Design, Marcel Dekker Inc., New York, NY, 1997.
8. Kubinyi, H., Curr. Opin. Drug Design Discov., 1 (1998) 4.
9. Caflisch, A., Wälchli, R. and Ehrhardt, C., News Physiol. Sci., 13 (1998) 182.
10. Böhm, H.J., Banner, D.W. and Weber, L., J. Comput.-Aided Mol. Design, 13 (1999) 51.
11. Caflisch, A. and Karplus, M., Perspect. Drug Discov. Design, 3 (1995) 51.
12. Majeux, N., Scarsi, M., Apostolakis, J., Ehrhard, C. and Caflisch, A., Proteins Struct. Funct. Genet., 37 (1999) 88.
13. Scarsi, M., Apostolakis, J. and Caflisch, A., J. Phys. Chem. A, 101 (1997) 8098.
14. Scarsi, M., Apostolakis, J. and Caflisch, A., J. Phys. Chem. B, 102 (1998) 3637.
15. Caflisch, A., J. Comput.-Aided Mol. Design, 10 (1996) 372.

16. So, S. and Karplus, M., J. Comput.-Aided Mol. Design, 13 (1999) 243.
17. Apostolakis, J. and Caflisch, A., Comb. Chem. High Throughput Screening, 2 (1999) 91.
18. Cramer, C.J. and Truhlar, D.G., Chem. Rev., 99 (1999) 2161.
19. Roux, B. and Simonson, T., Biophys. Chem., 78 (1999) 1.
20. Lee, B. and Richards, F.M., J. Mol. Biol., 55 (1971) 379.
21. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M., J. Comput. Chem., 4 (1983) 187.
22. Momany, F.A., Klimkowski, V.J. and Schäfer, L., J. Comput. Chem., 11 (1990) 654.
23. Nicholls, A., Sharp, K.A. and Honig, B., Proteins Struct. Funct. Genet., 11 (1991) 281.
24. Privalov, P.L. and Gill, S.G., Adv. Protein Chem., 39 (1988) 191.
25. Privalov, P.L. and Gill, S.G., Adv. Protein Chem., 247 (1990) 559.
26. Creighton, T.E., Curr. Opin. Struct. Biol., 1 (1991) 5.
27. Friedman, R.A. and Honig, B., Biophys. J., 69 (1995) 1528.
28. Caflisch, A., Fischer, S. and Karplus, M., J. Comput. Chem., 18 (1997) 723.
29. Vorobjev, Y.N., Almagro, J.C. and Hermans, J., Proteins Struct. Funct. Genetics, 32 (1998) 399.
30. Warwicker, J. and Watson, H.C., J. Mol. Biol., 157 (1982) 671.
31. Still, W.C., Tempczyk, A., Hawley, R.C. and Hendrickson, T., J. Am. Chem. Soc., 112 (1990) 6127.
32. Hawkins, G.D., Cramer, C.J. and Trulhar, D.G., Chem. Phys. Lett., 246 (1995) 122.
33. Hawkins, G.D., Cramer, C.J. and Trulhar, D.G., J. Phys. Chem., 100 (1996) 19824.
34. Schaefer, M. and Karplus, M., J. Phys. Chem., 100 (1996) 1578.
35. Di Qiu, Shenkin, P.S., Hollinger, F.P., and Still, W.C., J. Phys. Chem. A, 101 (1997) 3005.
36. Jackson, J.D., Classical Electrodynamics, John Wiley & Sons, New York, NY, 1975.
37. Luo, R., Moult, J. and Gilson, M.K., J. Phys. Chem. B, 101 (1997) 11226.
38. Richards, F.M., Annu. Rev. Biophys. Bioeng., 6 (1977) 151.
39. Tapparelli, C., Metternich, R., Ehrhardt, C. and Cook, N.S., Trends Pharmacol Sci., 14 (1993) 366.
40. Brandstetter, H., Turk, D., Hoeffken, H.W., Grosse, D., Stuerzebecher, J., Martin, D.P., Edwards, B.F.P. and Bode, W., J. Mol. Biol., 226 (1992) 1085.
41. Lee, J.C., Laydon, J.T., McDonnell, P.C., Gallagher, T.F., Kumar, S., Green, D., McNulty, D., Blumenthal, M.J., Heys, J.R., Landvatter, S.W. and Young, P.R., Nature, 372 (1994) 739.
42. Han, J., Lee, J.D., Bibbs, L. and Ulevitch, R.J., Science, 265 (1994) 808.
43. Rouse, J., Cohen, P., Trigon, S., Morange, M., Alonso-Llamazares, A., Zamanillo, D., Hunt T. and Nebreda, A.R., Cell, 78 (1994) 1027.
44. Freshney, N.W., Rawlinson, L., Guesdon, F., Jones, E., Cowley, S., Hsuan, J. and Saklatvala, J., Cell, 78 (1994) 1039.
45. Cuenda, A., Rouse, J., Doza, Y.N., Meier, R., Cohen, P., Gallagher, T.F., Young, P.R. and Lee, L.C., FEBS Lett., 364 (1995) 229.
46. Lee, L.C. and Adams, J.L., Curr. Opin. Biotechnol., 6 (1995) 657.
47. Gallagher, T.F., Seibel, G.L., Kassis, S., Laydon, J.T., Blumenthal, M.J., Lee, J.C., Lee, D., Boehm, J.C., Fier-Thompson, S.M., Abt, J.W., Soreson, M.E., Smietana, J.M., Hall, R.F., Garigipati, R.S., Bender, P.E., Erhard, K.F., Krog, A.J., Hofmann, G.A., Sheldrake, P.L., McDonnell, P.C., Kumar, S., Young, P.R. and Adams, J.L., Bioorg. Med. Chem., 5 (1997) 49.
48. Wang, Z., Canagarajah, B.J., Boehm, J.C., Kassisa, S., Cobb, M.H., Young, P.R., Abdel-Meguid, S., Adams, J.L. and Goldsmith, E.J., Structure, 6 (1998) 1117.

49. Tong, L., Pav, S., White, D.M., Rogers, S., Crane, K.M., Cywin, C.L., Brown, M.L. and Pargellis, C.A., Nat. Struct. Biol., 4 (1997) 311.
50. Gilson, M. K. and Honig, B. H., Proteins Struct. Funct. Genet., 4 (1988) 7.
51. Bashford, D. and Karplus, M., Biochemistry, 29 (1990) 10219.
52. Davis, M. E., Madura, J. D., Luty, B. A. and McCammon, J. A., Comput. Phys. Commun., 62 (1991) 187.
53. Marrone, T. J., Gilson, M. K. and McCammon, J. A., J. Phys. Chem., 100 (1996) 1439.
54. Davis, M.E. and McCammon, J.A., J. Comput. Chem., 10 (1989) 386.
55. Banner, D.W. and Hadvary, P., J. Biol. Chem., 266 (1991) 20085.
56. Stubbs, M. T. and Bode, W., Perspect. Drug Discov. Design, 1 (1993) 431.
57. Hilpert, K., Ackermann, J., Banner, D.W., Gast, A., Gubernator, K., Hadvary, P., Labler, L., Müller, K., Schmid, G., Tschopp, T. and van de Waterbeemd, H., J. Med. Chem., 37 (1994) 3889.
58. Lyle, T.A., Perspect. Drug Discov. Design, 1 (1993) 453.
59. Tabernero, L., Chang, C.Y., Ohringer, S., Lau, W.F., Iwanowicz, E.J., Han, W.C., Wang, T.C., Seiler, S.M., Roberts, D.G.M. and Sack, J.S., J. Mol. Biol., 246 (1995) 14.
60. Obst, U., Gramlich, V., Diederich, F., Weber, L. and Banner, D.W., Angew. Chem., 107 (1995) 1874.
61. Wagner, J., Kallen, J., Ehrhardt, C., Evenou, J. and Wagner, D., J. Med. Chem., 41 (1998) 3664.
62. Kikumoto, R., Tamao, Y., Tezuka, T., Tonomura, S., Hara, H., Ninomiya, K., Hijikata, A. and Okamoto, S., Biochemistry, 23 (1984) 85.
63. Stürzebecher, J., Walsmann, P., Voigt, B. and Wagner, G., Thromb. Res., 36 (1984) 457.
64. Rarey, M., Kramer, B., Lengauer, T. and Klebe, G., J. Mol. Biol., 261 (1996) 470.
65. Eldridge, M.D., Murray, C.W., Auton, T.R., Paolini, G.V. and Mee, R.P., J. Comput.-Aided Mol. Design, 11 (1997) 425.
66. Baxter, C.A., Murray, C.W., Clark, D.E., Westhead, D.R. and Eldridge, M.D., Proteins Struct. Funct. Genet., 33 (1998) 367.
67. Jones, G., Willett, P., Glen, R.C., Leach, A.R. and Taylor, R., J. Mol. Biol., 267 (1997) 727.
68. Lorber, D.M. and Shoichet, B.K., Protein Sci., 7 (1998) 938.
69. Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. and Ferrin, T.E., J. Mol. Biol., 161 (1982) 269.
70. DesJarlais, R.L., Sheridan, R.P., Dixon, J.S., Kuntz, I.D. and Venkataraghavan, R., J. Med. Chem., 29 (1986) 2149.
71. Shoichet, B.K., Stroud, R.M., Santi, D.V., Kuntz, I.D. and Perry, K.M., Science, 259 (1993) 1445.
72. Shoichet, B.K., Leach, A.R. and Kuntz, I.D., Proteins Struct. Funct. Genet., 34 (1999) 4.
73. Miranker, A. and Karplus, M., Proteins Struct. Funct. Genet., 11 (1991) 29.
74. Caflisch, A., Miranker, A. and Karplus, M., J. Med. Chem., 36 (1993) 2142.
75. Miranker, A. and Karplus, M., Proteins Struct. Funct. Genet., 23 (1995) 472.
76. Grootenhuis, P.D.J. and Karplus, M., J. Comput.-Aided Mol. Design, 10 (1996) 1.
77. Joseph-McCarthy, D., Hogle, J.M. and Karplus, M., Proteins Struct. Funct. Genet., 29 (1997) 32.
78. Leclerc, F. and Karplus, M., Theor. Chem. Acc., 101 (1999) 131.
79. Kolossvary, I. and Guida, W.C., J. Am. Chem. Soc., 118 (1996) 5011.
80. Klebe, G. and Mietzner, T., J. Comput.-Aided Mol. Design, 8 (1994) 583.
81. Apostolakis, J., Plückthun, A. and Caflisch, A., J. Comput. Chem., 19 (1998) 21.
82. Sánchez, R. and Sali, A., Proc. Natl. Acad. Sci. USA, 95 (1998) 13597.

83. Shuker, H.B., Hajduk, P.J., Meadows, R.P. and Fesik, S.W., Science, 274 (1996) 1531.
84. Hajduk, P.J., Meadows, R.P. and Fesik, S.W., Science, 278 (1997) 497.
85. Mattos, C. and Ringe, D., Nat. Biotechnol., 14 (1996) 595.
86. Dill, K.A., Nature, 400 (1999) 309.