PROTEINS: Structure, Function, and Genetics 42:256–268 (2001)

Efficient Electrostatic Solvation Model for Protein-Fragment Docking

Nicolas Majeux, Marco Scarsi, and Amedeo Caflisch*
Department of Biochemistry, University of Zürich, Zürich, Switzerland

ABSTRACT **A method is presented for the fast evaluation of the binding energy of a protein-small molecule complex with electrostatic solvation. It makes use of a fast preprocessing step based on the assumption that the main contribution to electrostatic desolvation upon ligand binding originates from the displacement of the first shell of water molecules. For a rigid protein, the precomputation of the energy contributions on a set of grids allows the estimation of the energy in solution of about 300 protein-fragment binding modes per second on a personal computer. The docking procedure is applied to five rigid binding sites whose size ranges from 17 residues to a whole protein of 107 amino acids. Using a library of 70 mainly rigid molecules, known micromolar inhibitors or close analogs are docked and prioritized correctly. The docking based rank-ordering of the library requires about 5 h and is proposed as a complementary approach to structure-activity relationships by nuclear magnetic resonance. Proteins 2001;42:256–268.**

Key words: binding energy; library docking; first solvation shell; FKBP-12; MDM2; p38 MAP kinase; thrombin; ICE

INTRODUCTION

An accurate treatment of solvation effects is required to determine the most probable binding mode of a protein-ligand complex as well as for ranking a library of organic compounds according to binding affinity.[1–3] An aqueous solution has a twofold effect on the electrostatics of ligand binding to a macromolecular receptor: 1. it screens the interactions between solute charges, and 2. it interacts directly with each solute charge (direct solvation or Born term). The screening can be favorable to binding (e.g., in the case of partial charges of same sign) but is in most cases unfavorable because of charge complementarity between protein and ligand. Direct electrostatic desolvation, i.e., displacement of high dielectric solvent by low dielectric solute, is always detrimental to protein-ligand binding.

The electrostatic effects of the solvent can be described by continuum electrostatics (see Refs. 4 and 5 for a review). The system is partitioned into solvent and solute regions and appropriate values of the dielectric constant are assigned to each region. Only the solute is described at the microscopical level as a spatial distribution of point charges. Hence, 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 and does not require any equilibration of the water molecules, which is usually computationally very expensive.

For a spatial discontinuity in the dielectric, the Coulomb law does not appropriately describe the electrostatic interactions between the solute partial charges. An exact solution of the problem is obtained by solving the Poisson equation for the electrostatic potential on a three-dimensional grid with a finite difference algorithm.[6] The results obtained with this numerical approach agree well with experimental data,[7–9] with free energy calculations,[10,11] and with molecular dynamics simulations that include explicit water molecules.[12–14] Even if the gain in computation time is remarkable with respect to explicit water calculations, for a wide screening of ligand-receptor conformations, this methodology is still too slow. For this reason, approximations have been introduced to simplify the original formulas of continuum electrostatics.[15–19] The accuracy of these approximated models can be tested by comparison with the finite difference solution of the Poisson equation[17,20] or explicit solvent simulations.[15,21,22]

Recently, we described a computational method for docking molecular fragments to a rigid protein with evaluation of the binding energy.[23] The approach is called SEED (solvation energy for exhaustive docking). For docking small- to medium-sized fragments, the program SEED makes use of an approximated continuum electrostatic model partially based on the generalized Born equation.[15,20] The agreement with finite difference solutions of the Poisson equation has been shown to be very good.[23]

Herein, we present a new electrostatic solvation approximation for the fast evaluation of binding energy. It is based on the assumption that electrostatic desolvation energies can be modeled by the first solvation shell at the binding interface. The new model has been implemented in SEED and is used as a preprocessing step together with a look-up

Grant sponsor: Helmut Horten Foundation; Grant sponsor: Swiss National Science Foundation; Grant number: 31-53604.98.

The program SEED (for SGI or PC running the Linux operating system) as well as the library of fragments in mol-2 format are available for nonprofit institutions from A.C.

*Correspondence to: Amedeo Caflisch, Department of Biochemistry, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland. E-mail: caflisch@bioc.unizh.ch

Received 18 July 2000; Accepted 22 September 2000

table evaluation of the van der Waals and electrostatic interaction energies. In the Results section, the validations of the electrostatic solvation approximation and the new SEED are presented. This is followed by the docking of a library of 70 compounds, including hydrophobic, polar and charged molecules, into five proteins with rigid binding site of different polar character and size.

METHODS

The docking approach implemented in the program SEED[23] determines optimal positions and orientations of small- to medium-sized molecular fragments in the binding site of a protein (called receptor henceforth). Apolar fragments are docked into hydrophobic regions of the receptor whereas polar fragments are positioned such that at least one intermolecular hydrogen bond is formed. Each fragment is placed at several thousand different positions with multiple orientations (for a total of approximately 10^6 conformations) and the binding energy is estimated whenever severe clashes are not present (usually about 10^5 conformations). The binding energy is the sum of the van der Waals interaction and the electrostatic energy. The latter consists of screened receptor-fragment interaction, as well as receptor and fragment desolvations. Depending on the fragment size, between 10 and 20 s of central processing unit (CPU) time on a 550 MHz Pentium III processor were required to calculate the binding energy of 100 receptor-fragment binding modes in the first version of SEED.[23]

An efficient way of speeding up the fragment screening by SEED is to divide the selection of favorable binding modes into two different steps. In the first step, a fast evaluation of the binding energy is performed ("fast model"). The binding energy in the fast model consists of the van der Waals interaction and electrostatic energy with an approximated treatment of solvation. The main assumption is that differences in electrostatic solvation energies upon binding can be approximated by the displacement of the first shell of water molecules from the binding interface. The fast energy has been implemented in a new version of SEED to quickly discard unfavorable conformations. This is done first by a van der Waals potential on a look-up table to prevent steric clashes and then by the electrostatic energy with fast solvation. The solvation approach presented in this article allows calculation of the binding energy of about 300 receptor-fragment conformations per second of CPU time on a 550 MHz Pentium III processor. This represents a speedup of a factor of about 40 with respect to the first version of SEED. The positions are then sorted according to binding energy and clustered by using a criterion based on distances between similar atom types.[24] In the second step, the n best binding modes ($n = 10$ is used in Results) within each cluster are evaluated with the more accurate solvation model.[23] The two-step approach has been implemented in a new version of SEED. The following sections focus on the evaluation of the binding energy with the fast model whereas the details of the more accurate solvation model can be found in our previous articles.[23,25]

Binding Energy

The binding energy can be partitioned into electrostatic and nonelectrostatic contributions.[4,5]

Electrostatic Energy in Solution

The electrostatic component of the binding energy is approximated by the sum of the following terms[26,27]:

- Screened receptor-fragment interaction: intermolecular electrostatic energy between the fragment and the receptor in the solvent.
- Partial desolvation of the receptor: electrostatic energy difference caused by the displacement of high dielectric solvent by the fragment volume.
- Partial desolvation of the fragment: electrostatic energy difference caused by the displacement of high dielectric solvent by the receptor volume.

The fast model is based on the linear distance-dependent approximation of the screened interaction and on the Coulombic approximation of the electric displacement[17,20] as explained in the following subsections.

Screened receptor-fragment interaction. The screened interaction between fragment and receptor ($E_{\text{elect}}^{\text{interm}}$ in kcal/mol) is calculated according to a linear distance-dependent dielectric model[28]:

$$E_{\text{elect}}^{\text{interm}} = 332 \sum_{\substack{i \in \text{fragment} \\ j \in \text{receptor}}} \frac{q_i q_j}{\epsilon_{\text{int}} r_{ij}^2} \tag{1}$$

where q_i and q_j are the partial charges (in electronic units) of atoms i and j, r_{ij} is the distance between them (in Å), and ϵ_{int} is the value of the interior, i.e., solute, dielectric constant (see below). There is no sound physical justification in favor of the distance-dependent dielectric function even if the agreement with more sophisticated models is remarkable (see Results). Nevertheless, it is a simple and useful approximation, because it yields a shorter range interaction than the Coulomb law. Recently, distance-dependent dielectric models have been used for docking and ligand design,[29–31] and molecular dynamics simulations of protein folding[32,33] and unfolding.[34]

Partial desolvation of the receptor. A preliminary step consists of the evaluation of the receptor desolvation caused by a low dielectric probe sphere of 1.4 Å radius rolling over the van der Waals surface of the receptor.[25] The center of the sphere spans the solvent accessible surface (SAS).[35] A number of points are distributed uniformly on the SAS of the receptor with a given surface density (usually 0.5 points per Å^2; see below) to describe the different positions of the center of the probe sphere. Furthermore, a cubic grid of 0.5 Å spacing is used to discretize the volume surrounding the receptor. The volume occupied by the probe sphere is then approximated on the cubic grid. The receptor desolvation resulting from the probe sphere at a point p on the SAS of the receptor (see Fig. 1A) is evaluated according to the Coulomb approximation of the electric displacement[25]:

$$\Delta G^{p}_{\text{desolv}} = \frac{1}{8\pi}\left(\frac{1}{\epsilon_{\text{int}}} - \frac{1}{\epsilon_{\text{w}}}\right) \sum_{k_p \in V_{\text{probe}}} \left(\sum_{j \in \text{receptor}} q_j \frac{(\vec{x}_{k_p} - \vec{x}_j)}{|\vec{x}_{k_p} - \vec{x}_j|^3} \right)^2 \Delta V \quad (2)$$

where the index k_p runs over the cubic grid elements occupied by the probe sphere and ΔV is the volume of a cube. Further, $\vec{x}_j$ is the coordinate vector of the receptor atom j, $\vec{x}_{k_p}$ the position of the cube included in the probe sphere, and ϵ_{int} and ϵ_{w} are the solute and solvent dielectric constants, respectively. The receptor desolvation caused by the probe sphere is calculated only once at the beginning of SEED for every point on the SAS. It is always positive, i.e., unfavorable, because $\epsilon_{\text{int}} < \epsilon_{\text{w}}$.

The receptor desolvation upon binding is approximated by the sum of the values of the desolvation operated by the probe sphere over the SAS receptor points that are included within the SAS of the fragment (see Fig. 1B):

$$\Delta G^{\text{receptor}}_{\text{desolv}} = \sum_{p \in \text{SAS}^{\text{buried}}_{\text{receptor}}} \Delta G^{p}_{\text{desolv}}. \quad (3)$$

Because the adjacent positions of the sphere can partially overlap, the total receptor desolvation is scaled by a multiplicative factor (see Results). The assumption underlying this model is that the main contribution to the receptor desolvation results from the removal of the first shell of water.[17] This approximation is justified by the fact that the desolvation of a spherical ion by a small low dielectric sphere at a distance r from the ion varies as $1/r^4$.[20,25]

Partial desolvation of the fragment. The desolvation of the fragment ($\Delta G^{\text{fragment}}_{\text{desolv}}$) is evaluated in a way that is specular to the receptor desolvation (see Fig. 1A and B). First, the fragment desolvation caused by a probe sphere rolling over the fragment SAS is calculated once for every fragment type. Subsequently, the fragment desolvation upon binding is approximated by the sum of the desolvation values associated with the points on the SAS of the fragment that are included within the SAS of the receptor. The same scaling factor as for the receptor desolvation is used (see Results).

Nonelectrostatic Contributions

These are a fine balance between the solute-solute van der Waals interactions (favorable to binding), the loss of solute-solvent van der Waals interactions (unfavorable), and the disruption of water structure that is a favorable entropic effect at room temperature.[36] A number of approaches have been proposed to approximate these contributions.[36–41] Interestingly, similar results have been obtained from models resulting from different assumptions (e.g., nonelectrostatic contributions represented by solute-solute dispersion interactions or by a buried surface curvature-dependent model).[40] In accord with previous findings (Caflisch et al.[41] and Figure 6 of Vorobjev et al.[42]), it is assumed here that solute-solvent van der Waals interactions and disruption of water structure compensate each

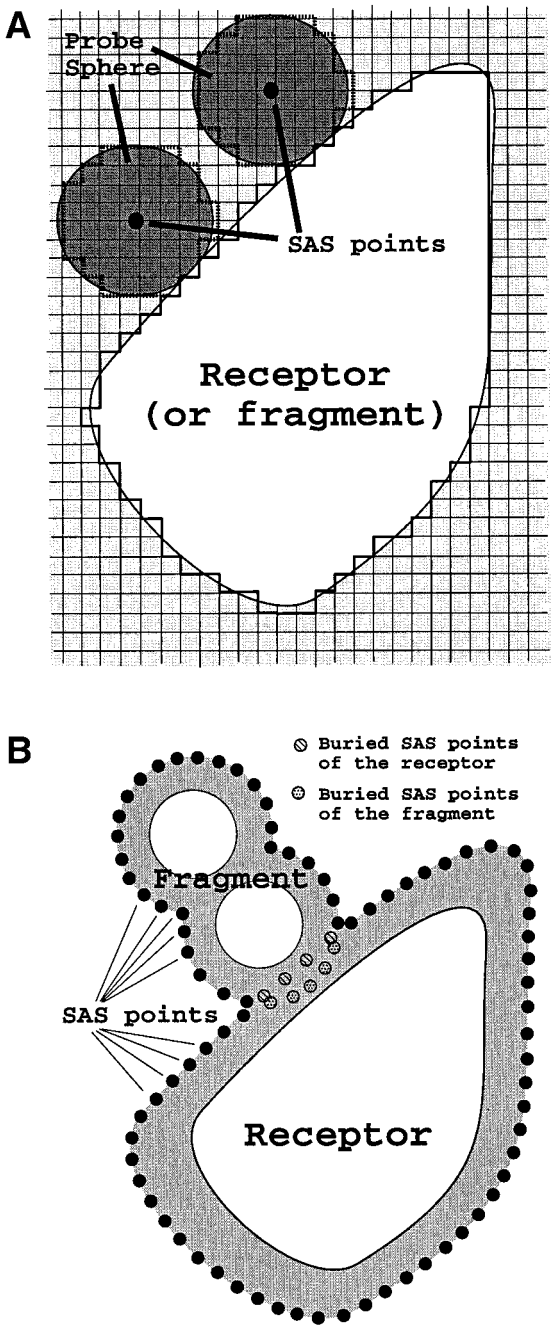

Fig. 1. **A:** Evaluation of the receptor (fragment) desolvation caused by a probe sphere rolling over the van der Waals surface of the receptor (fragment). A grid box of 0.5-Å spacing (light gray) is built around the molecule and the desolvation resulting from the occupation of every grid element is evaluated as described in Ref. 23. The desolvation caused by the probe sphere in a given position is approximated by the sum of the values of the desolvation caused by the cubes within the sphere (dark gray). **B:** Evaluation of the receptor and fragment desolvations upon binding. The points on the SAS of the receptor and the SAS of the fragment represent the positions of the rolling probe sphere. The receptor desolvation is approximated by the sum of the desolvation values associated with the SAS points of the receptor buried within the SAS of the fragment. The fragment desolvation is approximated by the sum of the desolvation values associated with the SAS points of the fragment buried within the SAS of the receptor.

other, and that solute-solute van der Waals interactions can account for the nonelectrostatic part of the binding energy.[25] The neglect of a term proportional to the loss in solvent accessible surface upon binding is justified by its strong correlation with the solute-solute van der Waals interaction.[43]

The van der Waals interaction between a fragment and the receptor is described as the sum of a steep repulsion and an attractive dispersion term with the 6-12 Lennard-Jones model:

$$E_{\mathrm{vdW}} = \sum_{i \in \mathrm{fragment}} \sum_{j \in \mathrm{receptor}} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) \quad (4)$$

where r_{ij} is the distance between atoms i and j, A_{ij} and B_{ij} are van der Waals repulsion and attraction parameters. The assumption that the receptor is rigid favors the use of a grid-based evaluation of the interaction. To make the fragment and receptor terms in Eq. (4) factorizable, the geometric mean approximation is used[29,44–46]: $A_{ij} = \sqrt{A_i A_j}$ and $B_{ij} = \sqrt{B_i B_j}$, with $A_i = \varepsilon_i (2R_i^{\mathrm{vdW}})^{12}$ and $B_i = 2\varepsilon_i (2R_i^{\mathrm{vdW}})^{6}$. R_i^{vdW} is the van der Waals radius of atom i and ε_i is the minimum of the van der Waals potential between two atoms of type i at optimal distance of $2R_i^{\mathrm{vdW}}$. A grid is spanned over the binding site of the receptor and the grid spacing is usually 0.2 Å or 0.3 Å. When the program starts, for every grid point p the two following "receptor potentials" are calculated and stored in look-up tables:

$$\phi^A(p) = \sum_{j \in \mathrm{receptor}} \frac{\sqrt{A_j}}{r_{pj}^{12}} \quad \text{and} \quad \phi^B(p) = \sum_{j \in \mathrm{receptor}} \frac{\sqrt{B_j}}{r_{pj}^{6}} \quad (5)$$

where the sums run over the receptor atoms which are within a 10 Å cutoff distance of the grid point. The contribution of fragment atom i with coordinates $\vec{x}_i$ is evaluated by multiplying its van der Waals parameters ($\sqrt{A_i}$ and $\sqrt{B_i}$) with the "receptor potentials" (ϕ^A and ϕ^B, respectively). The value of the potential is derived from the eight points of the grid surrounding $\vec{x}_i$ by the trilinear interpolation method.[47]

Total Binding Energy

The total binding energy in the fast model is

$$\Delta G_{\mathrm{binding}}^{\mathrm{fast}} = E_{\mathrm{vdW}}^{\mathrm{grid}} + E_{\mathrm{elect}}^{\mathrm{interm}} + \Delta G_{\mathrm{desolv}}^{\mathrm{receptor}} + \Delta G_{\mathrm{desolv}}^{\mathrm{fragment}}. \quad (6)$$

The four contributions on the right-hand side are calculated according to Eqs. (1) to (5). The fast van der Waals energy ($E_{\mathrm{vdW}}^{\mathrm{grid}}$) is used also to detect clashes: fragments are discarded without evaluation of the electrostatic contributions if $E_{\mathrm{vdW}}^{\mathrm{grid}}$ is less favorable than a threshold value (usually 1 kcal/mol).

Differences in the intrasolute entropy of binding are neglected in SEED. They are assumed to be sufficiently similar for the different fragment types and binding modes.[41] The solute entropic penalty, which is not accounted for in the present model, is expected to partially counterbalance the very favorable binding energy values (see Results).

Fragment Library and Docking

A library of 70 mainly rigid fragments ranging in size from 7 to 31 atoms was used in this study. It contains 17 apolar fragments (no hydrogen bond donors or acceptors), 39 polar and neutral compounds, and 14 fragments with one or two formal charges (Table I and Fig. 2). Many of the molecular frameworks found frequently in known drugs[48] are included (e.g., benzene, pyridine, naphthalene, 5-phenyl-1,4-benzodiazepine, etc.) and some of them can be used for the synthesis of combinatorial libraries in the solid phase[49] or by portioning and mixing.[50] Fragment structures were generated with the molecular modeling program WITNOTP (A. Widmer, Novartis Pharma Basel, unpublished). For each fragment type, all of the low-energy conformations are included in the library (e.g., cis and trans for 2-butene). Partial charges were assigned with the MPEOE method,[51,52] whose WITNOTP implementation reproduces the all-hydrogen CHARMm22 parameter set (Molecular Simulations Inc.) for proteins and proteinaceous fragments. Fragment coordinates were minimized with the program CHARMM[53] and the CHARMm22 parameter set to an average value of the energy gradient of 0.01 kcal/mol Å using a linear distance-dependent dielectric function.

The fragment library was docked by SEED into five proteins whose coordinates were taken from the Protein DataBank (PDB).[54] To be consistent with the partial atomic charges of the fragments, partial charges for the protein atoms were taken from the CHARMm22 parameter set (Molecular Simulations Inc). Hydrogen atoms for the proteins were generated with WITNOTP and then minimized with CHARMM by keeping the heavy atoms fixed. For docking, a grid spacing of 0.3 Å was used for the evaluation of $E_{\mathrm{vdW}}^{\mathrm{grid}}$, and the same input parameters as in Table I of the original SEED article[23] were used except for the following three. The interior dielectric constant was set to four to model the electronic polarizability and dipolar reorientation effects of the solute.[55] The number of apolar points was increased from 100 to 150 because of the very large binding sites (see below). Finally, the radius of the probe sphere for the definition of the SAS (used for the selection of the apolar vectors) was set to 1.4 Å (instead of 1.8 Å) to better define microcavities and small crevices.

RESULTS

Validation of the Fast Electrostatics Model

The fast model for the evaluation of the electrostatic component of the binding energy is validated by comparison with finite difference solutions of the Poisson equation obtained by the program UHBD.[56,57] For this purpose, the three electrostatic energy terms were calculated with the fast model of SEED and with UHBD 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 proteinase (HIV-1 PR) monomer. The molecule set included acetate ion, benzoate ion, methylsulfonate ion, methylammonium ion, methylguanidinium ion, 2,5-diketopiperazine, and benzene. The total number of receptor-fragment complexes analyzed was 1,025 for thrombin

TABLE I. Fragment Library

Fragment[a]	No. of atoms		Number of conformations	H-bond		Formal charges
	Heavy	Total		Acceptors	Donors	
Apolar fragments						
Ethane	2	8	1	—	—	0
Propane	3	11	1	—	—	0
Cyclopropane	3	9	1	—	—	0
2-Methylpropane	4	14	1	—	—	0
1-Butene	4	12	4	—	—	0
2-Butene	4	12	2	—	—	0
2-Methyl-2-butene	5	15	1	—	—	0
2,2-Dimethylpropane	5	17	1	—	—	0
Cyclopentane	5	15	1	—	—	0
Benzene	6	12	1	—	—	0
Cyclohexane	6	18	1	—	—	0
Adamantane	10	26	1	—	—	0
Dekaline	10	28	1	—	—	0
Naphthalene	10	18	1	—	—	0
N-Methylindole	10	19	1	—	—	0
Tetraline	10	22	1	—	—	0
Dibenzocyclohexane (**2**)	14	26	1	—	—	0
Polar and neutral fragments						
Dimethylsulfoxyde	4	10	1	1	—	0
Isopropanol	4	12	2	1	1	0
Imidazole	5	9	1	1	1	0
N-Methylacetamide	5	12	1	1	1	0
Pyrrole	5	10	1	—	1	0
N-Methyl-methylsulfonamide	6	13	1	2	1	0
Oxazolidinone	6	11	1	2	1	0
Pyridine	6	11	1	1	—	0
Pyrimidine	6	10	1	2	—	0
2-Pyrrolidinone	6	13	1	1	1	0
4-Thiazolidinone	6	11	1	1	1	0
Delta-valero-lactam	7	16	2	1	1	0
3,4-Dihydroxy-tetrahydrofurane	7	15	4	3	2	0
Phenol	7	13	1	1	1	0
Tetrahydro-2-pyrimidinone	7	15	1	1	2	0
Cytosine	8	13	1	2	2	0
1,2-Dihydroxy-benzene	8	14	1	2	2	0
1,2-Dihydroxy-cyclohexane	8	20	4	2	2	0
2,5-Diketo-1,4-piperazine	8	14	2	2	2	0
Uracil	8	12	1	2	2	0
Indole	9	16	1	—	1	0
2-Methyl-3-amino-N-methylbutanamide	9	23	18	1	2	0
Adenine	10	15	1	3	2	0
3,6-Dimethyl-2,5-diketo-1,4-piperazine	10	20	3	2	2	0
Isoquinoline	10	17	1	1	—	0
N-Formyl-L-proline (**5**)	10	19	8	3	1	0
Quinazoline	10	16	1	2	—	0
Quinoline	10	17	1	1	—	0
Tetrahydro-quinoline	10	21	1	—	1	0
Guanine	11	16	1	3	3	0
Meso-inositol	12	24	1	6	6	0
Alpha-carboline	13	21	1	1	1	0
Beta-carboline	13	21	1	1	1	0
Diphenylether (**1**)	13	23	1	1	—	0
5-Methyl-3-methylsulfoxyde-acetophenone	13	25	8	2	—	0
2,3,4-Furantricarboxylic-acid	14	18	8	7	3	0
5-Phenyl-1,4-benzodiazepine (**3**)	17	29	2	2	—	0
5-Phenyl-1,4-benzodiazepine-2-one (**4**)	18	30	2	2	1	0
4-(4-Fluorophenyl)-1-methyl-5-(4-pyridyl)-imidazole (**6**)	19	31	2	2	—	0
Charged fragments						
Methylammonium	2	8	1	—	1	+
Methylamidine	4	11	1	—	2	+
Methylguanidine	5	13	1	—	3	+
Tetrahydropyrrole	5	15	1	—	1	+
Piperidine	6	18	1	—	1	+
Benzamidine (**8**)	9	18	2	—	2	+
5-Amidine-indole (**7**)	12	22	2	—	3	+
Acetate (**10**)	4	7	1	2	—	−
Methylsulfonate (**9**)	5	8	1	3	—	−
Benzoic acid (**11**)	9	14	1	2	—	−
L-Proline	8	17	2	2	1	+−
D-Proline	8	17	2	2	1	+−
Piperazine	6	18	1	—	2	++
Phosphate	6	9	1	4	—	−−

[a]The numbers in boldface correspond to the structures shown in Figure 2.

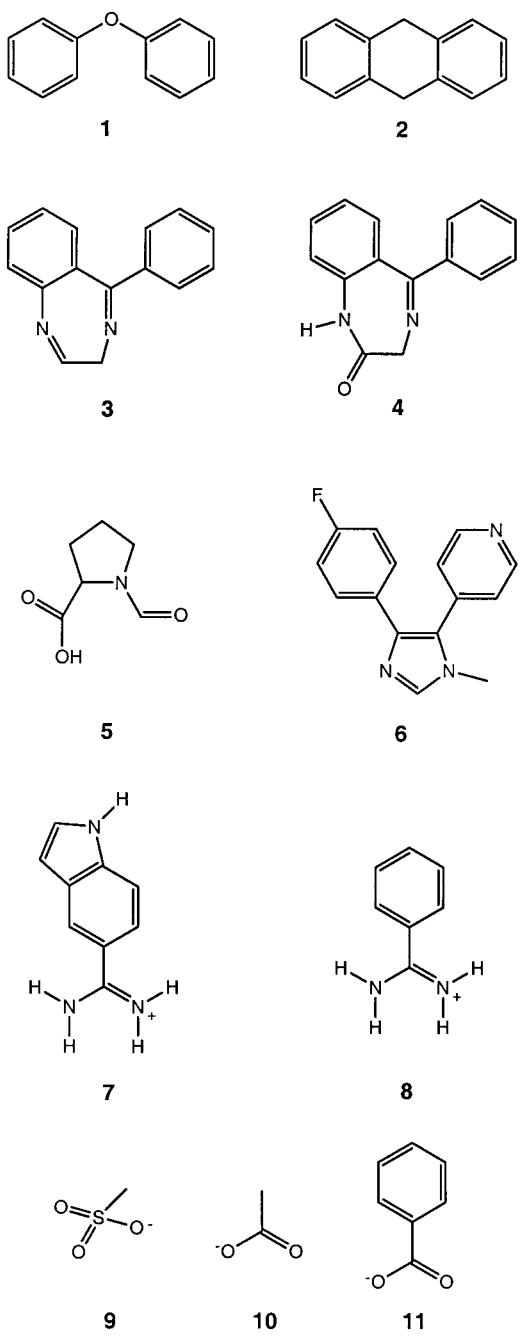

Fig. 2. Chemical structure of 11 compounds in the fragment library.

and 1,490 for the HIV-1 PR monomer. A surface density of 0.5 points per Å^2 was used to discretize the SAS of both receptor and fragment. A surface density of two points per Å^2 did not show any improvement in accuracy but required a 4-times larger computational cost. A cubic grid with a spacing of 0.5 Å was used to discretize the volume surrounding the receptor.

The results are shown in Table II for two different values of the solute dielectric constant (1 and 4). Figure 3 shows the results for thrombin assuming a solute dielectric constant of 1 and a solvent dielectric constant of 78.5. The agreement between the two methods is good. Systematic errors (slope $\neq$ 1) can be corrected by the use of appropriate scaling factors for the different energy terms. Two scaling factors are adopted independently of the value of the solute dielectric constant. One applies to both the receptor and fragment desolvations and one to the screened interaction. The scaling factor for the desolvations (2.13) results from an average of the slopes of the fitting lines of the ligand and receptor desolvations in Table II (eight values in total). The scaling factor for the screened interaction (1.43) results from an average of the slopes of the fitting lines of the screened interactions in Table II (four values in total). This scaling factor is in accord with previous studies, in which it was found that the interaction energies calculated with the linear distance-dependent dielectric function are approximately 1.4-times larger than the values obtained by the finite difference solution of the Poisson equation.[29,55]

Validation of Fast van der Waals

The errors caused by the use of the geometric mean approximation and trilinear interpolation on the cubic grid were validated by comparison with the standard off-grid van der Waals potential with a 10 Å cutoff. As a test set, 1,025 small molecules were distributed over the active site of thrombin using SEED. The agreement between the van der Waals energy and the geometric mean approximation based on the grid is very good with a grid spacing of 0.2 Å. The correlation factor (R) and slope (s) of the linear fitting are equal to 0.99 and 0.95, respectively. The largest discrepancies are found for the positions with steric clashes because of the steepness of the $1/r^{12}$ repulsive term. For the 609 positions with $E_{\mathrm{vdW}}^{\mathrm{grid}} < 1$ kcal/mol, R = 0.95, s = 1.14, and the average and maximal errors are 0.5 and 4.9 kcal/mol, respectively. A grid spacing of 0.3 Å yields R = 0.98 and s = 0.75 for the 1,025 complexes, and for the 556 positions with $E_{\mathrm{vdW}}^{\mathrm{grid}} < 1$ kcal/mol, R = 0.95, s = 1.04, and average and maximal errors are 0.5 and 3.2 kcal/mol, respectively.

Validation of the New SEED on Thrombin

The thrombin functionality maps calculated with the new version of SEED, which uses the fast evaluation of the binding energy as preprocessing step, are compared with the results obtained previously with the old version of SEED.[23] For both values of the solute dielectric constant ($\epsilon_{\mathrm{int}} = 1$ and $\epsilon_{\mathrm{int}} = 4$), the new version of SEED reproduces the most favorable binding modes of the 13 fragment types (four nonpolar, four polar and neutral, three cations, one dication, and one anion) used in the previous study. This is because the fast model does not miss favorable binding modes, and false positives are then given low priority in the second step. The largest difference in ranking is found for cluster 3 of benzene (total binding energy of -2.9 kcal/mol) which is ranked as number 10 (total binding energy of -2.4 kcal/mol) with the new version of SEED. This is probably because of two reasons. First, the fast model ranking might yield a different clustering than in the old SEED. Second, the accurate energy is evaluated

TABLE II. Comparison of Electrostatic Energies Calculated by SEED (Fast Model) and UHBD

	Thrombin		HIV-protease monomer	
	R^a	α^b	R^a	α^b
Solute dielectric constant = 1				
Receptor desolvation	0.86	1.92	0.91	1.69
Ligand desolvation	0.83	1.93	0.91	2.95
Screened interaction	0.88	1.42	0.91	1.61
Total fitted electrostatic energy[c]	0.85	0.91	0.84	1.12
Solute dielectric constant = 4				
Receptor desolvation	0.87	1.98	0.92	1.75
Ligand desolvation	0.87	1.94	0.94	2.85
Screened interaction	0.92	1.18	0.94	1.49
Total fitted electrostatic energy[c]	0.88	0.86	0.89	1.08

[a]Correlation coefficient for the fit of the SEED to the UHBD energies.
[b]Slope of the fitting line of the SEED to the UHBD energies.
[c]Sum of the three energy components, each scaled according to the multiplicative factors described in Results.

only for the first 10 members of each cluster in the new SEED, whereas the remaining positions are discarded. Depending on the fragment type, the old version of SEED (without fast preprocessing) required between 0.5 and 1.5 h CPU time of a 550 MHz Pentium III processor for docking a fragment in a medium-size binding site (10 Å sphere). For the same binding site and processor type, the time needed with the new two-step approach ranges from 1 (benzamidine) to 5 min (benzene, *N*-methyl-methylsulfonamide).

Library Docking

Three of the five proteins investigated (FKBP-12, MDM2, and p38 MAP kinase) have a mainly hydrophobic binding site, whereas the remaining two (thrombin and interleukin-1β converting enzyme [ICE]) have a partially hydrophilic character because of the presence of charged side-chains in the active site. The binding site definition ranged from 17 (ICE) to 107 (FKBP-12) residues. The new version of SEED was used for docking the library of fragment. The CPU time required was approximately the same in all binding sites (about 5 h) because the same amount of hydrophobic points (150) and polar vectors (3,000) was selected for docking in all of the five test cases. The binding modes and energy values presented below refer to the more accurate solvation model,[23] i.e., after clustering and postprocessing.

FKBP-12

FKBP-12 is a peptidylprolyl isomerase of 107 residues and 12 kDa which is widely distributed in almost all tissues. It catalyzes the interconversion between cis and trans rotamers of peptidylprolyl amide bonds in peptide substrates.[58,59] Furthermore, FKBP-12 binds the natural product FK506, a macrocyclic immunosuppressant drug. For the SEED run, the 1.85 Å resolution X-ray structure of unliganded FKBP-12 (PDB code 1d6o) was used, and the binding site definition included all of the 107 residues. The hydrophobic compounds have very favorable SEED binding energy. Considering only the best binding mode for every fragment type, the binding energy ranges from −14.3 kcal/mol to −7.4 kcal/mol for the 17 apolar fragments. Fifteen of these are docked in the active site, which consists of several aromatic and aliphatic side chains. The substrate analog *N*-formyl-L-proline has the best binding energy followed by two 1,4-benzodiazepines (Table III). The proline ring of *N*-formyl-L-proline overlaps the piperidine ring of FK506 and is involved in the same interactions with FKBP-12 (Fig. 4A). These consist of van der Waals contacts with the side chains of Tyr26, Phe46, Val55, and Trp59, and a hydrogen bond with the hydroxyl group of Tyr82. The van der Waals interaction is the main contribution to the binding energy, whereas the electrostatic desolvation counterbalances the intermolecular electrostatic term (Table III).

MDM2

The MDM2 oncoprotein is a cellular inhibitor of the p53 tumor suppressor. It can associate to the transactivation domain of p53 and downregulate its ability to activate transcription of genes responsible for growth arrest or apoptosis after genotoxic events.[60] The 2.3 Å crystal structure of the *X. laevis* MDM2 protein from the complex with a 15-residue transactivation domain of p53 (PDB code 1ycq[61]) was used in the SEED calculations. The p53 peptide was removed and the binding site was defined by taking into account all of the 88 residues of MDM2 that are in the PDB file. All of the apolar fragments have very favorable binding energy (values ranging from −17.4 to −7.6 kcal/mol for the best binding modes) and are docked in an extended pocket, which in the MDM2-p53 complex is occupied by the side-chains of a triad of hydrophobic residues of p53 (Phe19, Trp23, and Leu26). The pivotal role of these three residues of p53 for binding to MDM2 is supported by crystallographic studies[61] and phage display experiments.[62] The fragments with the best binding energy are two 1,4-benzodiazepines and dibenzocyclohexane (Table III). Their benzene rings overlap the Phe19 and Trp23 side-chains of p53 (Fig. 4b). Furthermore, the NH group of the 1,4-benzodiazepine-2-one matches the indolic NH of Trp23 and is involved in the same hydrogen bond with the backbone CO group of the MDM2 residue Leu54. As in the case of FKBP-12, the van der Waals interaction term is dominant. The SEED results suggest that C-3 substituted derivatives of 5-phenyl-1,4-benzodiazepine-2-one might bind to MDM2 by filling the extended hydrophobic pocket occupied by the p53 residues Phe19, Trp23, and Leu26 (Fig. 4B).

p38 MAP Kinase

Mitogen-activated protein (MAP) kinases are essential enzymes for intracellular signaling cascades because they phosphorylate several regulatory proteins. p38 MAP kinase plays a role in processes as diverse as transcriptional regulation, production of interleukins, and apoptosis of

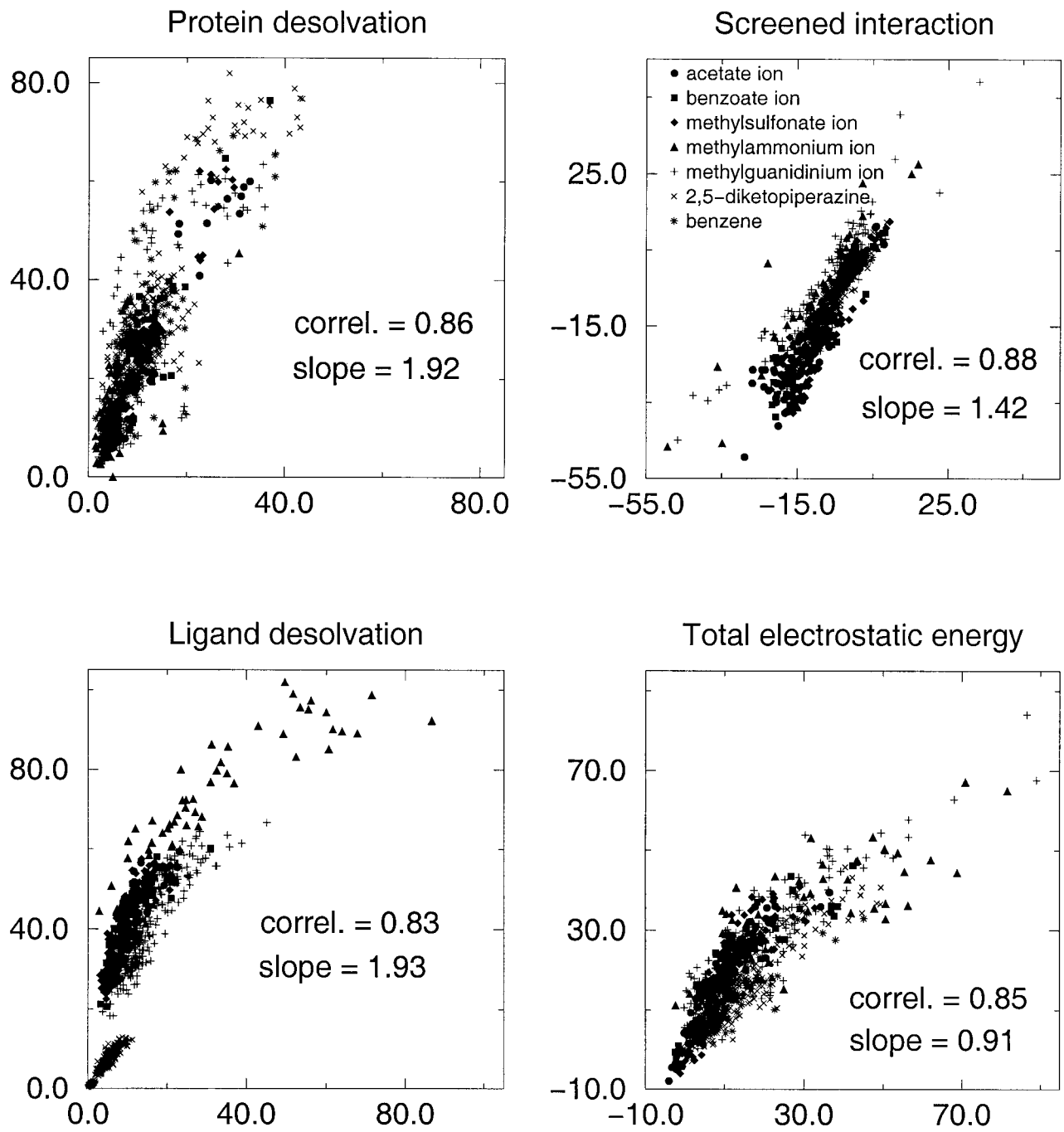

Fig. 3. Correlations in the electrostatic energies (in kcal/mol) calculated by finite difference solution of the Poisson equation (*x*-axis) and by SEED with the fast model (*y*-axis). Values are plotted for 1,025 complexes of thrombin with small molecules. The total electrostatic energy is the sum of the protein desolvation (Eq. 3), ligand desolvation, and screened interaction (Eq. 1), scaled by 2.13, 2.13, and 1.43, respectively, as described in Results. The finite difference calculations were performed with the program UHBD.[56,57] Values of $\epsilon_{int} = 1$, $\epsilon_w = 78.5$, and grid spacing of 0.5 Å were used for both SEED and UHBD.

neuronal cells. Inhibitors of p38 activity could therefore be useful as a treatment strategy for inflammatory and neurodegenerative diseases. The library used in this study contains a disubstituted imidazole (compound **6**, Fig. 2) that is a close analog of a class of potent inhibitors of p38. These 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.[63–65] Compound **6** does not have the third substituent. Figure 4C shows that SEED docks compound **6** in the right orientation (heavy atom root mean square deviation of 0.9 Å from the position of the inhibitor SB203580, PDB code 1a9u[66]) and ranks it as best fragment with a very favor-

TABLE III. Results of the Docking of the Fragment Library[†]

Protein/fragment[a]	$\Delta G_{binding}$[b]	Intermolecular		Electrostatic desolvation	
		van der Waals	Electrostatic	Receptor	Fragment
FKBP-12 (1d6o; 107)[c]					
N-Formyl-L-proline (**5**)[d]	−16.1	−17.9	−2.3	2.0	2.0
1,4-Benzodiazepine-2-one (**4**)	−16.0	−19.5	−0.3	3.1	0.7
1,4-Benzodiazepine (**3**)	−15.7	−19.2	−0.1	3.1	0.5
MDM2 (1ycq; 88)					
1,4-Benzodiazepine-2-one (**4**)	−21.0	−22.5	−0.4	0.8	1.2
1,4-Benzodiazepine (**3**)	−19.4	−20.9	0.2	0.8	0.5
Dibenzocyclohexane (**2**)	−17.4	−18.1	−0.4	0.7	0.4
p38 MAP kinase (1a9u; 32)					
Compound **6**	−17.1	−26.0	−1.9	9.7	1.1
Diphenylether (**1**)	−16.2	−21.3	−0.9	5.6	0.5
Dibenzocyclohexane (**2**)	−16.0	−22.8	−0.5	6.8	0.4
Thrombin (1hgt; 20)					
5-Amidine-indole (**7**)	−26.6	−21.3	−19.2	3.1	10.7
Benzamidine (**8**)	−25.0	−19.8	−18.0	2.4	10.3
1,4-Benzodiazepine-2-one (**4**)	−19.8	−24.5	−1.2	4.8	1.1
ICE (1ice; 17)					
Methylsulfonate (**9**)	−14.7	−9.6	−17.3	2.4	9.9
Acetate (**10**)	−13.7	−9.5	−17.1	2.4	10.6
Benzoic acid (**11**)	−12.0	−11.6	−16.7	5.1	11.2

[†]All energy values are in kcal/mol.
[a]The protein name is in boldface.
[b]Sum of the values in the four following columns, i.e., intermolecular and electrostatic desolvation energies calculated as described in Ref. 23.
[c]PDB code and number of residues used to define the binding site in SEED.
[d]The numbers in boldface correspond to the structures shown in Figure 2.

able van der Waals energy (Table III). Furthermore, the aromatic rings of the second and third best fragments, diphenylether and dibenzocyclohexane, overlap the pyridine and phenyl rings of the inhibitor SB203580.

Thrombin

Thrombin is a trypsin-like serine protease that plays an essential role in the blood coagulation cascade.[67] It is one of the best characterized enzymes from a structural viewpoint and binds a series of diverse inhibitors without major rearrangements of its conformation.[68–71] The natural substrate, i.e., fibrinogen, and most of the known inhibitors of thrombin have a basic group that is involved in a salt bridge with Asp189 at the bottom of the S1 or recognition pocket.[72] There are two known micromolar inhibitors of thrombin in the library used in this study: 5-amidine-indole (IC_{50} = 22 μM[73]) and benzamidine (K_i = 300 μM[68]). SEED docks them correctly in the S1 pocket of thrombin (heavy atoms root mean square deviation of 0.8 Å between benzamidine docked by SEED and the corresponding moiety of the NAPAP inhibitor, PDB code 1dwd[70]) and ranks them as first and second with a gap in binding energy of about 5 kcal/mol with respect to the third best fragment (Table III). Because of the favorable salt-bridge energy, the intermolecular electrostatic energy is much more favorable for 5-amidine-indole and benzamidine than for the remaining fragments. Despite the high desolvation penalty of both 5-amidine-indole (10.7 kcal/mol) and benzamidine (10.3 kcal/mol), the total electrostatic energy is favorable by about −5 kcal/mol. The poor ranking of the piperazine dication (no. 67 of 70, $\Delta G_{binding}$ = −5.2 kcal/mol, E_{elect}^{interm} = −24.8 kcal/mol, and $\Delta G_{desolv}^{fragment}$ = 33.6 kcal/mol) indicates that an accurate description of fragment desolvation is required to avoid that compounds with many formal charges are preferred. In fact, a model that neglects the desolvation of the fragment would rank piperazine as best fragment followed by 5-amidine-indole and benzamidine.

ICE

ICE is a cysteine protease whose substrates have an aspartic acid adjacent and N-terminal to the scissile peptide bond. The 2.6-Å resolution X-ray structure of ICE, from a covalent complex with the tetrapeptide aldehyde inhibitor acetyl-Tyr-Val-Ala-Asp-H (PDB code 1ice[74]), was used in the SEED calculations. The three compounds with the best SEED binding energy have a formal charge of −1 (Table III) in agreement with the substrate preference. They are involved in a salt bridge with Arg179 and Arg341 which are the salt-bridge partners of the aspartic acid in the substrate.[74] The deviation of the acetate from the corresponding atoms in the Asp side chain of the inhibitor is 1.7 Å.

DISCUSSION

We have presented a computational approach for the fast evaluation of electrostatic desolvation energies of

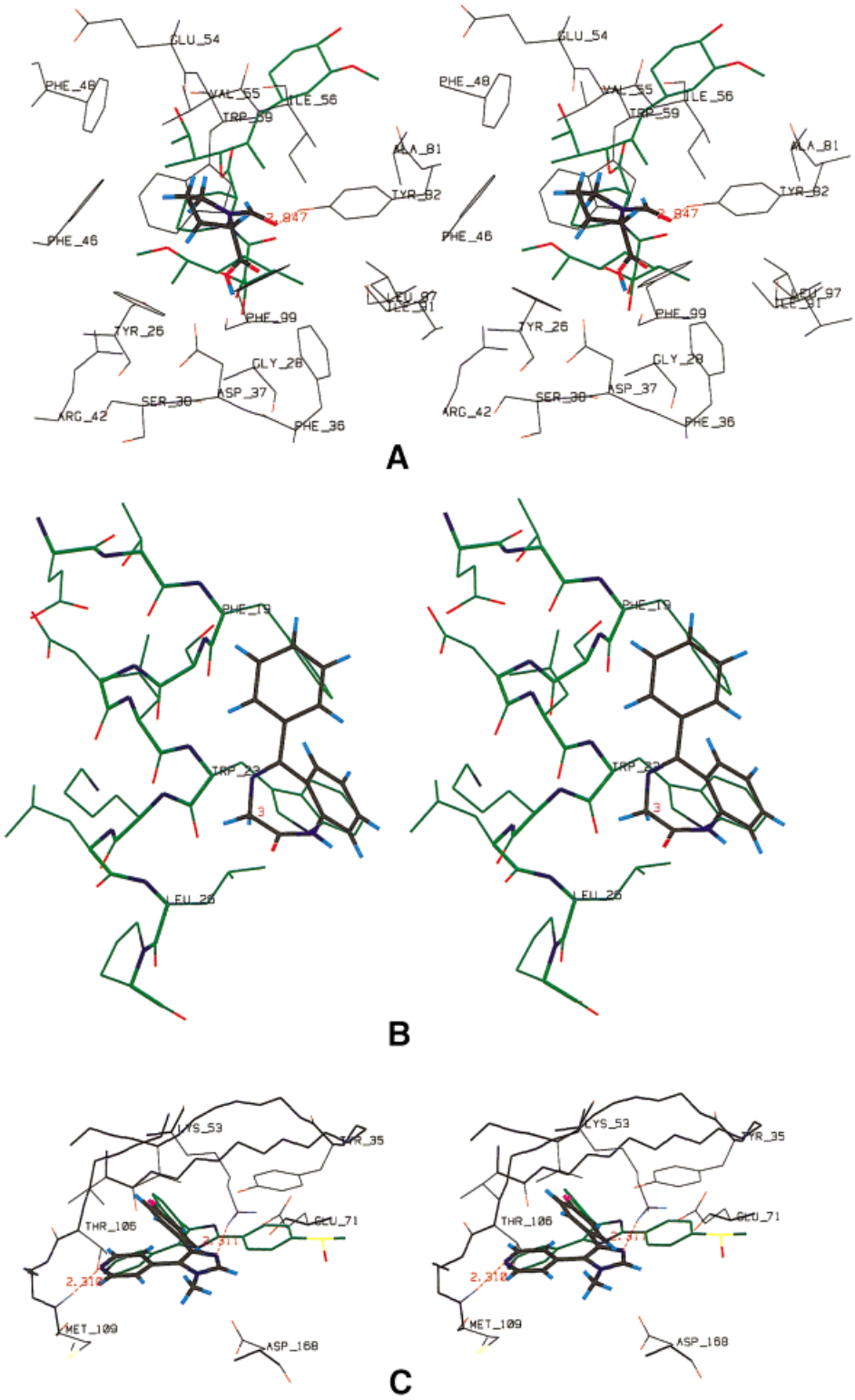

Fig. 4. Relaxed-eyes stereopictures of SEED results. **A:** The FKBP binding site is shown in thin lines with part of the FK506 inhibitor (medium lines, green carbon atoms), and the *N*-formyl-L-proline (thick lines) docked by SEED. Oxygen atoms are colored in red, nitrogen in blue, and hydrogen in cyan. Hydrogen bonds are shown by a red dashed line. **B:** The p53 inhibitor is shown with green carbon atoms superimposed to the 1,4-benzodiazepine-2-one (black carbon atoms) docked by SEED into the MDM2 oncoprotein (not shown). The C-3 atom of 1,4-benzodiazepine-2-one is labeled; it could be substituted with a hydrophobic group to mimic the Leu26 side-chain of p53 (see text). **C:** The p38 MAP kinase binding site is shown in thin lines with the SB203580 inhibitor (medium lines, green carbon atoms), and a close analog, i.e., compound **6** (thick lines), docked by SEED.

receptor and ligand upon binding. It is based on the continuum model and the assumption that electrostatic desolvation can be approximated by the removal of the first layer of water molecules at the binding interface. The fast model is used in a new version of the program SEED as preprocessing step. Two-step approaches have already been shown to be both efficient and sufficiently accurate for docking and qualitative ranking.[27,75] The fast model allows evaluation of the binding energy with electrostatic solvation of about 10^5 conformations of a complex between a rigid receptor and a small molecule in about 5 min of CPU time on a personal computer. The new SEED was tested by docking a library of 70 fragments into five protein binding sites of different sizes and characteristics, ranging from almost completely hydrophobic to mainly hydrophilic. Known micromolar inhibitors or close analogs were docked and rank-ordered correctly.

Recently, our numerical approach for the calculation of electrostatic energy in solution[20] was implemented in the DOCK program by others.[1] Two important differences of the new SEED with respect to DOCK are: 1. the first shell approximation of desolvation, and 2. a more accurate description in SEED of the difference in the volume occupied by the solute upon binding. The embedded volume is automatically recalculated for every receptor-fragment conformation in SEED, whereas in DOCK it has to be defined a priori by the user and is assumed to be constant for different fragment types and orientations.[1] The precise description of the boundary between low-dielectric and high-dielectric regions is extremely critical for obtaining accurate electrostatic energies in solution.[20]

Two limitations of SEED are the use of a rigid protein target and the neglect of conformational entropy. Even for a protein that shows the same overall conformation in complexes with different inhibitors, a minor rotation of a charged or polar side-chain in the binding site may result in very different values of the electrostatic desolvation. Moreover, a slightly different orientation of any type of side-chain can affect the calculated binding energy dramatically because of the steepness of the repulsive term in the van der Waals potential. This is clearly a main limitation of SEED. A partial solution to this problem is to run SEED separately on multiple binding-site conformations. Although the required time scales linearly on multiple processors, more human time is required for the analysis of the functionality maps obtained on different protein conformations. The neglect of conformational entropy is another limitation, because especially for the side-chains in the binding site, this contribution could vary widely even for the same ligand in two different binding pockets. A possible solution might be the use of an entropic penalty proportional to the number of frozen rotatable bonds.

SEED is a computational tool that has certain aspects in common with structure-activity relationships by nuclear magnetic resonance (SAR by NMR), an experimental method for the identification of small molecules that bind to proximal subsites of a protein.[76] Using NMR-derived structural data, the small ligands are subsequently linked together resulting in potent inhibitors whose binding energy can be more favorable than the sum of the components.[77,78] The computational screening by SEED could be used as a filter to design molecular libraries for SAR by NMR experiments. Furthermore, SAR by NMR, as well as other methods used for detecting weakly bound ligands, requires that the compounds be soluble at millimolar concentrations in aqueous solutions. In the case of unsoluble hydrophobic fragments, one could use the results of SEED to complement the SAR by NMR data for the identification and/or optimization of high affinity ligands.

ACKNOWLEDGMENTS

We thank S. Ahmed, N. Budin, and Dr. C. Tenette-Souaille for helpful discussions. We also thank Dr. C. Ehrhardt (Novartis Pharma, Basel) for preparing the fragment library and for critical reading of the manuscript, and A. Widmer (Novartis Pharma, Basel) for the molecular modeling program WITNOTP, which was used to generate the library and Figure 4.

REFERENCES

1. Zou X, Sun Y, Kuntz ID. Inclusion of solvation in ligand binding free energy calculations using the generalized-Born model. J Am Chem Soc 1999;121:8033–8043.
2. Apostolakis J, Caflisch A. Computational ligand design. Comb Chem High Throughput Screen 1999;2:91–104.
3. Scarsi M, Caflisch A. Comment on the validation of continuum electrostatics models. J Comput Chem 1999;14:1533–1536.
4. Roux B, Simonson T. Implicit solvent models. Biophys Chem 1999;78:1–20.
5. Cramer CJ, Trulhar DG. Implicit solvation models: equilibria, structure, spectra, and dynamics. Chem Rev 1999;99:2161–2200.
6. Warwicker J, Watson HC. Calculation of the electric potential in the active site cleft due to α-helix dipoles. J Mol Biol 1982;157:671–679.
7. Mohan V, Davis ME, McCammon JA, Pettitt BM. Continuum model calculations of solvation free energies: accurate evaluation of electrostatic contributions. J Phys Chem 1992;96:6428–6431.
8. Sharp K, Jean-Charles A, Honig B. A local dielectric constant model for solvation free energies which accounts for solute polarizability. J Phys Chem 1992;96:3822–3828.
9. Sitkoff D, Sharp KA, Honig B. Accurate calculation of hydration free energies using macroscopic solvent models. J Phys Chem 1994;98:1978–1988.
10. Jean-Charles A, Nicholls A, Sharp K, et al. Electrostatic contributions to solvation energies: Comparison of free energy perturbation and continuum calculation. J Am Chem Soc 1991;113:1454–1455.
11. Nina M, Beglov D, Roux B. Atomic radii for continuum electrostatics calculations based on molecular dynamics free energy simulations. J Phys Chem B 1997;101(26):5239–5248.
12. Marrone TJ, Gilson MK, McCammon JA. Comparison of continuum and explicit models of solvation: potential of mean force for alanine dipeptide. J Phys Chem 1996;100:1439–1441.
13. Resat H, Marrone TJ, McCammon JA. Enzyme-inhibitor association thermodynamics: explicit and continuum solvent studies. Biophys J 1997;72:522–532.
14. Berneche S, Nina M, Roux B. Molecular dynamics simulation of melittin in a dimyristoylphosphatidylcholine bilayer membrane. Biophys J 1998;75(4):1603–1618.
15. Still WC, Tempczyk A, Hawley RC, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. J Am Chem Soc 1990;112:6127–6129.
16. Hawkins GD, Cramer CJ, Trulhar DG. Pairwise solute descreening of solute charges from a dielectric medium. Chem Phys Lett 1995;246:122–129.
17. Schaefer M, Karplus M. A comprehensive analytical treatment of continuum electrostatics. J Phys Chem 1996;100:1578–1599.
18. Hawkins GD, Cramer CJ, Trulhar DG. Parametrized models of aqueous free energies of solvation based on pairwise descreening

of solute atomic charges from a dielectric medium. J Phys Chem 1996;100:19824–19839.
19. Qiu Di, Shenkin PS, Hollinger FP, Still WC. The GB/SA continuum model for solvation: a fast analytical method for the calculation of approximate Born radii. J Phys Chem A 1997;101: 3005–3014.
20. Scarsi M, Apostolakis J, Caflisch A. Continuum electrostatic energies of macromolecules in aqueous solutions. J Phys Chem A 1997;101:8098–8106.
21. Scarsi M, Apostolakis J, Caflisch A. Comparison of a GB solvation model with explicit solvent simulations: potentials of mean force and conformational preferences of alanine dipeptide and 1,2-dichloroethane. J Phys Chem B 1998;102:3637–3641.
22. Dominy BN, Brooks CL. Development of a generalized Born model parametrization for proteins and nucleic acids. J Phys Chem B 1999;103(18):3765–3773.
23. Majeux N, Scarsi M, Apostolakis J, Ehrhard C, Caflisch A. Exhaustive docking of molecular fragments on protein binding sites with electrostatic solvation. Proteins 1999;37:88–105.
24. Kearsley SK, Smith GM. An alternative method for the alignment of molecular structures: maximizing electrostatic and steric overlap. Tetrahedron Comput Methodol 1990;3:615–633.
25. Scarsi M, Majeux N, Caflisch A. Hydrophobicity at the surface of proteins. Proteins 1999;37:565–575.
26. Gilson MK, Honig BH. Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. Proteins 1988;4:7–18.
27. Caflisch A. Computational combinatorial ligand design: application to human α-thrombin. J Comput Aided Mol Design 1996;10: 372–396.
28. Gelin BR, Karplus M. Sidechain torsional potentials and motion of amino acids in proteins: bovine pancreatic trypsin inhibitor. Proc Natl Acad Sci USA 1975;72:2002–2006.
29. Luty BA, Wasserman ZR, Stouten PFW, Hodge CN, Zacharias M, McCammon JA. A molecular mechanics/grid method for evaluation of ligand-receptor interactions. J Comput Chem 1995;16:454–464.
30. Caflisch A, Ehrhardt C. Structure-based combinatorial ligand design. In: Veerapandian P, editor. Structure-based drug design. New York: Marcel Dekker; 1997. p 541–558.
31. Horvath D. A virtual screening approach applied to the search for trypanothione reductase inhibitors. J Med Chem 1997;40:2412–2423.
32. Ferrara P, Apostolakis J, Caflisch A. Computer simulations of protein folding by targeted molecular dynamics. Proteins 2000;39: 252–260.
33. Ferrara P, Caflisch A. Folding simulations of a three-stranded antiparallel β-sheet peptide. Proc Natl Acad Sci USA 2000;97: 10780–10785.
34. Lazaridis T, Karplus M. "New view" of protein folding reconciled with the old through multiple unfolding simulations. Science 1997;278:1928–1931.
35. Lee B, Richards FM. The interpretation of protein structures: estimation of static accessibility. J Mol Biol 1971;55:379–400.
36. Nicholls A, Sharp KA, Honig B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 1991;11:281–296.
37. Privalov PL, Gill SG. Stability of protein structure and hydrophobic interactions. Adv Protein Chem 1988;39:191–234.
38. Privalov PL, Gill SG. Common features of protein unfolding and dissolution of hydrophobic compounds. Adv Protein Chem 1990; 247:559–561.
39. Creighton TE. Stability of folded conformations. Curr Opin Struct Biol 1991;1:5–16.
40. Friedman RA, Honig B. A free energy analysis of nucleic acid base stacking in aqueous solution. Biophys J 1995;69:1528–1535.
41. Caflisch A, Fischer S, Karplus M. Docking by Monte Carlo minimization with a solvation correction: application to an FKBP-substrate complex. J Comput Chem 1997;18:723–743.
42. Vorobjev YN, Almagro JC, Hermans J. Discrimination between native and intentionally misfolded conformations of proteins: ES/IS, a new method for calculating conformational free energy that uses both dynamics simulations with an explicit solvent and an implicit solvent continuum model. Proteins 1998;32:399–413.
43. So S, Karplus M. A comparative study of ligand-receptor complex binding affinity prediction methods based on glycogen phosphorylase inhibitors. J Comput Aided Mol Design 1999;13:243–258.
44. Hagler AT, Huler E, Lifson S. Energy functions for peptides and proteins. I. Derivation of a consistent force field including the hydrogen bond from amide crystals. J Am Chem Soc 1974;96:5319–5327.
45. Pattabiraman N, Levitt M, Ferrin TE, Langridge R. Computer graphics in real-time docking with energy calculation and minimization. J Comput Chem 1985;6:432–436.
46. Meng EC, Shoichet BK, Kuntz ID. Automated docking with grid-based energy evaluation. J Comput Chem 1992;13:505–524.
47. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical recipes in fortran. New York: Cambridge University Press; 1992.
48. Bemis GW, Murcko MA. The properties of known drugs. 1. Molecular frameworks. J Med Chem 1996;39:2887–2893.
49. Andres CJ, Denhart DJ, Deshpande MS, Gillman KW. Recent advances in the solid phase synthesis of drug-like heterocyclic small molecules. Comb Chem High Throughput Screen 1999;2:191–210.
50. Furka A, Bennett WD. Combinatorial libraries by portioning and mixing. Comb Chem High Throughput Screen 1999;2:105–122.
51. Gasteiger J, Marsili M. Iterative partial equalization of orbital electronegativity: a rapid access to atomic charges. Tetrahedron 1980;36:3219–3288.
52. No KT, Grant JA, Scheraga HA. Determination of net atomic charges using a modified partial equalization of orbital electronegativity method. 1. Application to neutral molecules as models for polypeptides. J Phys Chem 1990;94:4732–4739.
53. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 1983;4: 187–217.
54. Bernstein FC, Koetzle TF, Williams GJB, et al. The protein data bank: a computer-based archival file for macromolecular structures. J Mol Biol 1977;112:535–542.
55. Gilson MK, Honig BH. Energetics of charge-charge interactions in proteins. Proteins 1988;3:32–52.
56. Davis ME, McCammon JA. Solving the finite difference linearized Poisson-Boltzmann equation: a comparison of relaxation and conjugate gradient methods. J Comput Chem 1989;10:386–391.
57. Davis ME, Madura JD, Luty BA, McCammon JA. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian dynamics program. Comput Phys Commun 1991;62:187–197.
58. Siekierka JJ, Hung HY, Poe M, Lin CS, Sigal NS. A cytosolic binding protein for the immunosuppressant FK506 has peptidyl-prolyl isomerase activity but is distinct from cyclophilin. Nature 1989;341:755–757.
59. Harding MW, Galat A, Uehling DE, Schreiber SL. A receptor for the immunosuppressant FK506 is a cis-trans peptidyl-proly isomerase. Nature 1989;341:758–760.
60. Momand J, Zambetti GP, Olson DC, George D, Levine AJ. The MDM-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. Cell 1992;69:1237.
61. Kussie PH, Gorina S, Marechal V, et al. Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. Science 1996;274:948–953.
62. Böttger A, Böttger V, Garcia-Echeverria C, et al. Molecular characterization of the hdm2-p53 interaction. J Mol Biol 1997;269: 744–756.
63. Cuenda A, Rouse J, Doza YN, et al. SB 203580 is a specific inhibitor of a map kinase homologue which is stimulated by cellular stresses and interleukin-1. FEBS Lett 1995;364:229–233.
64. Lee JC, Laydon JT, McDonnell PC, et al. A protein kinase involved in the regulation of inflammatory cytokine biosynthesis. Nature 1994;372:739–746.
65. Lee JC, Adams JL. Inhibitors of serine/threonine kinases. Curr Opin Biotechnol 1995;6:657–661.
66. Wang Z, Canagarajah BJ, Boehm JC, et al. Structural basis of inhibitor selectivity in MAP kinases. Structure 1998;6:1117–1128.
67. Tapparelli C, Metternich R, Ehrhardt C, Cook NS. Synthetic low-molecular weight thrombin inhibitors: molecular design

and pharmacological profile. Trends Pharmacol Sci 1993;14:366–376.
68. Hilpert K, Ackermann J, Banner DW, et al. Design and synthesis of potent and highly selective thrombin inhibitors. J Med Chem 1994;37:3889–3901.
69. Bode W, Mayr I, Baumann U, Huber R, Stone SR, Hofsteenge J. The refined 1.9-Å crystal structure of human α-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. EMBO J 1989;8:3467–3475.
70. Banner DW, Hadvary P. Crystallographic analysis at 3.0-Å resolution of the binding to human thrombin of four active site-directed inhibitors. J Biol Chem 1991;266:20085–20093.
71. Obst U, Gramlich V, Diederich F, Weber L, Banner DW. Design neuartiger, nichtpeptidischer Thrombin-Inhibitoren und Struktur eines Thrombin-Inhibitor-Komplexes. Angew Chem 1995;107:1874–1877.
72. Lyle TA. Small-molecule inhibitors of thrombin. Perspect Drug Disc Design 1993;1:453–460.
73. Iwanowicz EJ, Lau WF, Lin J, Roberts DGM, Seiler SM. Derivatives of 5-amidine indole as inhibitors of thrombin catalytic activity. Bioorg Med Chem Lett 1996;6:1339–1344.
74. Wilson KP, Black FJ, Thomson JA, et al. Structure and mechanism of interleukin-1β converting enzyme. Nature 1994;370:270–275.
75. Hoffmann D, Kramer B, Washio T, Steinmetzer T, Rarey M, Lengauer T. Tow-stage method for protein-ligand docking. J Med Chem 1999;42:4422–4433.
76. Shuker HB, Hajduk PJ, Meadows RP, Fesik SW. Discovering high affinity ligands for proteins: SAR by NMR. Science 1996;274:1531–1534.
77. Hajduk PJ, Shepperd G, Nettesheim DG, et al. Discovery of potent nonpeptide inhibitors of stromelysin using SAR by NMR. J Am Chem Soc 1997;119:5818–5827.
78. Olejniczak ET, Hajduk PJ, Marcotte PA, et al. Stomelysin inhibitors designed from weakly bound fragments: effects of linking and cooperativity. J Am Chem Soc 1997;119:5828–5832.