

Empirical free energy as a target function in docking and design: application to HIV-1 protease inhibitors

Benjamin L. King, Sandor Vajda, Charles DeLisi*

Molecular Engineering Research Laboratory, Department of Biomedical Engineering, Boston University College of Engineering, Boston, MA 02215, USA

Received 19 February 1996; revised version received 7 March 1996

Abstract Structure-based drug design requires the development of efficient computer programs for exploring the structural compatibility of various flexible ligands with a given receptor. While various algorithms are available for finding docked conformations, selecting a target function that can reliably score the conformations remains a serious problem. We show that the use of an empirical free energy evaluation method, originally developed to characterize protein-protein interactions, can substantially improve the efficacy of search algorithms. In addition to the molecular mechanics interaction energy, the function takes account of solvation and side chain conformational entropy, while remaining simple enough to replace the incomplete target functions used in many drug docking and design procedures. The free energy function is used here in conjunction with a simple site mapping-fragment assembly algorithm, for docking the MVT-101 non-peptide inhibitor to HIV-1 protease. In particular, we predict the bound structure with an all atom RMSD of 1.21 Å, compared to 1.69 Å using an energy target function, and also accurately predict the free energy shifts obtained with a series of five trimeric hydroxyethylene isostere analogs.

Key words: Free energy method; Free energy mapping; Protein-ligand docking and design; HIV-1 protease inhibitor

1. Introduction

The ability to limit the ways in which a putative ligand might bind a known receptor site is crucial to the success of structure-based drug design [1]. Automated docking methods position the ligand and receptor molecules together in many different ways and then 'score' each configuration according to a target function [2,3]. The main issues in developing computational tools for docking are the selection of an efficient search algorithm, and the choice of a target function that can reliably identify the most likely conformations of the ligand-receptor complex [1,4].

The number of search algorithms used in docking is exploding, with methods ranging from the simple but very successful geometric search in the DOCK program [5] to various combinatorial [6–8], Monte Carlo [9–13], and genetic [14,15] algorithms. Another large class of methods is based on the site

mapping-fragment assembly strategy [16] used in a variety of programs [17–23]. The basic idea is to construct maps of the binding site by placing individual functional groups on a grid, and determining the most favorable positions. These positions are then used to build entire ligand molecules in the assembly stage of the procedure. A variety of drug design programs based on this principle have been developed.

While a large number of search strategies are available, the development of an effective target function remains a major obstacle to achieving automatic and reliable procedures for docking and design [1,4]. The ideal target function would be the binding free energy. However, computationally intensive free energy evaluation through free energy perturbation and thermodynamic integration techniques [24] is not viable, and hence at present docking and design are based on the use of simplified target functions such as shape complementarity, electrostatic or total interaction energy [25]. These are useful under certain conditions, but are generally unable to correctly rank the docked conformations, largely because they fail to properly account for solvation and entropy change [26].

Here we present results indicating that the use of a recently introduced empirical binding free energy function [27] can substantially improve the accuracy and efficiency of site mapping and fragment assembly algorithms in docking and design. Although results are presented only for the HIV-1 protease inhibitor MVT-101 and for five hydroxyethylene isostere analogs, the free energy evaluation model is general and has been applied to a variety of other systems [28]. The example shown here demonstrates that, starting with the structural motif seen in other peptide-based protease inhibitors, MVT-101 can be docked to within 1.21 Å all atom RMSD (root-mean square deviation) from the crystal structure. For comparison, the same procedure with a target function based on the CHARMM energy (i.e., neglecting solvation and entropic effects) yields a structure with 1.69 Å RMSD. We also show that starting with predicted structures of the HIV-1 PR[Aba_{469,497,469,497}]-MVT-101 [29] inhibitor complex, the free energy correctly rank orders five trimeric hydroxyethylene isostere analogs.

2. Methods

2.1. Free energy evaluation

The binding free energy, ΔG , is written as a sum of the various free energy contributions

$$\Delta G = E_{\text{el}}^{\text{cl}} + \langle \Delta G_{\text{h}} \rangle + \langle \Delta E_{\text{i}} \rangle - T\Delta S_{\text{sc}} - T\Delta S_{\text{bb}} + \text{const} \quad (1)$$

where $E_{\text{el}}^{\text{cl}}$ is the electrostatic interaction energy between the receptor and the ligand, and $\langle \Delta E_{\text{i}} \rangle$ is the difference in the internal energy of the ligand between its bound (E_{i}^{b}) and free (E_{i}^{f}) states. The operation $\langle \dots \rangle$ denotes Boltzmann-weighted averaging over an appropriate sample of free ligand conformations. The terms $T\Delta S_{\text{sc}}$ and $T\Delta S_{\text{bb}}$ represent con-

*Corresponding author. Fax: (1) (617) 353 5929.
E-mail: delisi@enga.bu.edu

Abbreviations: HIV-1 PR[Aba_{469,497,469,497}]-MVT-101, HIV-1 protease, with Aba (L- α -amino-*n*-butyric acid) substituted for Cys at positions 69 and 97 in both halves of the dimer, complexed with the peptide inhibitor MVT-101 with sequence Ace-Thr-Ile-Nle- Ψ [CH₂-NH]-Nle-Gln-Arg-amide, where Nle represents norleucine and Ace is N-acetyl.

Table 1

Site mapping with the free energy target function (Eq. 2); positions with the lowest and second lowest values of the target function for each residue

Res.	Rot. ^a	Translation (Å)			Rotation (degrees)			RMSD (Å)	
		x	y	z	x	y	z	Mapped ^b	Final ^c
Thr ₁	1	0.0	0.3	0.0	0.0	−30.0	0.0	1.78	—
Thr ₁	2	0.0	0.3	0.0	0.0	−30.0	0.0	1.26	1.32
Ile ₂	3	0.3	0.0	0.0	0.0	0.0	0.0	1.51	1.00
Ile ₂	3	0.0	0.3	0.0	0.0	0.0	−30.0	1.79	—
Nle ₃	1	0.0	−0.3	0.0	0.0	0.0	0.0	1.10	—
Nle ₃	1	−0.3	0.0	0.0	0.0	0.0	0.0	1.06	0.77
Nle ₄	1	0.0	0.0	−0.3	−30.0	0.0	0.0	1.68	—
Nle ₄	1	0.0	0.0	−0.3	0.0	0.0	−30.0	1.70	2.03
Gln ₅	4	0.0	0.0	−0.3	0.0	30.0	0.0	1.29	1.19
Gln ₅	4	0.0	−0.3	0.0	0.0	30.0	0.0	1.24	—
Arg ₆	2	0.0	−0.3	0.0	0.0	0.0	0.0	1.29	—
Arg ₆	2	0.0	0.0	−0.3	0.0	0.0	0.0	1.08	0.61

^aNumber of rotamer as listed by Ponder and Richards [39]. Rotamers for Met were used for Nle₃ and Nle₄.

^bAll atom RMSD between the best mapped position of the individual residue, and the corresponding conformation in the X-ray structure of MVT-101 [29].

^cAll atom RMSD between the residue in the best docked conformation following concatenation, and the corresponding conformation in the X-ray structure of MVT-101 [29].

tributions to the free energy due to the loss of side chain (sc) and backbone (bb) entropy. $\langle \Delta G_h \rangle$ represents the hydrophobic contribution to the binding free energy and is defined as

$$\langle \Delta G_h \rangle = \Delta G_h^{\text{rl}} - \langle \Delta G_h^{\text{r}} \rangle - \langle \Delta G_h^{\text{l}} \rangle$$

where the terms on the right hand side denote free energies of transfer from water into a (partially) nonpolar environment (in this paper octanol), and the superscripts rl, r, and l indicate the receptor-ligand complex, the free receptor, and the free ligand, respectively. $\langle \Delta G_h^{\text{r}} \rangle$ and $\langle \Delta G_h^{\text{l}} \rangle$ are Boltzmann-weighted averages over the free states. Finally, 'const' is a number (9 kcal/mol) that is independent of the details of the interaction, and reflects the loss of rotational, vibrational and cratic free energies.

The origin and testing of this function were discussed extensively in Vajda et al. [27]. Briefly, the hydrophobic free energy change is evaluated using a modified Eisenberg-McLachlan atomic solvation parameter model [30]; the electrostatic term using a Coulombic expression with CHARMM parameters; and the conformational entropy term by scaling the side chain entropy losses listed by Pickett and Sternberg [31]. Prior to free energy evaluation, all van der Waals clashes are removed with an energy minimization protocol. The model assumes that the solute-solute interface is as well packed as the solute-solvent interface; i.e., the van der Waals interactions lost when the ligand and receptor are separated, are compensated by van der Waals interactions gained between the exposed surfaces and the solvent, and hence the total change in the van der Waals interactions is small. This assumption, used in a number of semi-empirical free energy functions [32,33], is clearly violated by packing defects, and provides only a first-order approximation. Nevertheless, it is better than restricting considerations to protein-ligand interaction energy, and thereby neglecting all interactions with the solvent.

Free energy terms that do not vary from one bound conformation to another need not be considered in docking, and hence the target functions becomes

$$\Phi \equiv \Delta G - \langle \Delta E_l \rangle + T\Delta S_{\text{bb}} - \text{const} + \langle \Delta G_h^{\text{r}} \rangle + \langle \Delta G_h^{\text{l}} \rangle = E_{\text{rl}}^{\text{el}} + \Delta G_h^{\text{rl}} + E_l^{\text{b}} - T\Delta S_{\text{sc}} \quad (2)$$

For comparison we repeat the calculations with the traditional target function of the form

$$\Phi_{\text{CHARMM}} = E_{\text{rl}} + E_l^{\text{b}} \quad (3)$$

where E_{rl} is the CHARMM interaction energy between the receptor and the ligand, and E_l^{b} is the CHARMM conformational energy of the ligand. The interaction energy consists of electrostatic and van der Waals terms, whereas the conformational energy also includes bond angle, bond length, dihedral and improper angle energies [34].

For the calculation of target functions of trimeric analogs in which

the central side chains are substituted, we will always use the backbone of the three central residues of the docked MVT-101 inhibitor. Since we are interested only in target function differences when we compare analogs, the internal energy E_l^{b} will cancel. We also take the backbone conformational entropy of the free trimers to be the same. Hence for comparing analogs, we need only evaluate

$$\Omega = E_{\text{rl}}^{\text{el}} + \Delta G_h^{\text{rl}} - T\Delta S_{\text{sc}} \quad (4)$$

and

$$\Omega_{\text{CHARMM}} = E_{\text{rl}} \quad (5)$$

2.2. Docking the MVT-101 inhibitor to HIV-1 protease

Protease bound peptide-based inhibitors have a well-defined structural motif, and structural conservation is even stronger within a particular class of proteases. For example, the calculated α -carbon RMSD is 0.56 Å for HIV-1 protease inhibitors MVT-101 [29], JG-365 [35], U-85548e [36], A-74704 [37], and acetylpepstatin [38]. This strong overall backbone homology reduces the search space considerably and suggests that a very simple search algorithm should be effective.

The free energy of the receptor is mapped by translating and rotating individual amino acid residues with each side chain conformation from the rotamer library of Ponder and Richards [39]. We allow limited C^α movement (± 0.3 Å along each axis) and limited reorientation of the C^α - C^β bond ($\pm 30^\circ$ about each axis). The procedure yields 49 possible translational/rotational positions for each rotamer. The origin in all translations and rotations corresponds to the α -carbon positions from the crystal; however, the rotamer conformations used are different from the crystal, thus avoiding the imposition of bias. For each state thus obtained ($49 \times$ the number of rotamers for a given side chain), the probe-receptor system is subjected to 120 steps of CHARMM minimization, in order to define a space that avoids steric clashes, and the target function is evaluated. The parameters used throughout in all steps are a distance dependent dielectric constant (4r), a non-bonded cutoff of 17.0 Å, and mass-weighted harmonic constraints of 20 kcal/mol/Å² on atomic positions. The mapping procedure creates a list of favorable positions and conformations for each residue, ranked according to its free energy of binding to the receptor.

The sequence of MVT-101, with the number of rotamers used for each residue, is as follows: Thr(2), Ile(3), Nle(3), Nle(3), Gln(6), and Arg(5). In spite of the small rotamer library [39] and the limited number of translations and rotations, the combinations of discrete states result in about 2.24×10^{10} conformations. Several procedures short of an exhaustive search, which is not feasible, can be used to generate complete structures from the mapping data, including extended dynamic programming methods [8]. Here we select only the two best states for each side chain, giving a total of 64 structures for the 6 residue long inhibitor. Prior to free energy evaluation, the energy

Table 2

Site mapping with the CHARMM energy target function (Eq. 3); positions with the lowest and second lowest values of the target function for each residue

Res.	Rot. ^a	Translation (Å)			Rotation (degrees)			RMSD (Å)	
		x	y	z	x	y	z	Mapped ^a	Final ^a
Thr ₁	2	0.0	0.0	0.3	0.0	0.0	0.0	0.33	1.24
Thr ₁	2	0.0	−0.3	0.0	0.0	−30.0	0.0	1.28	–
Ile ₂	1	0.0	0.3	0.0	0.0	30.0	0.0	1.66	–
Ile ₂	1	0.3	0.0	0.0	0.0	30.0	0.0	1.63	1.14
Nle ₃	1	0.0	0.0	0.3	0.0	−30.0	0.0	0.97	0.92
Nle ₃	1	0.0	−0.3	0.0	0.0	−30.0	0.0	0.97	–
Nle ₄	2	0.3	0.0	0.0	0.0	0.0	−30.0	1.10	2.41
Nle ₄	2	0.0	0.3	0.0	0.0	0.0	−30.0	0.98	–
Gln ₅	3	0.0	0.0	0.3	0.0	0.0	30.0	2.26	2.17
Gln ₅	3	−0.3	0.0	0.0	0.0	0.0	30.0	2.47	–
Arg ₆	2	−0.3	0.0	0.0	−30.0	0.0	0.0	1.85	–
Arg ₆	2	0.0	0.0	0.3	−30.0	0.0	0.0	1.92	1.72

^aNumber of rotas described in Table 1.

of each of the 64 structures is minimized with a mass-weighted harmonic constraint of 20 kcal/mol/Å² for 200 steps to remove possible steric clashes and strains that can arise during concatenation. In order to provide a frame of reference for comparison, 100 additional structures were generated by randomly selecting a conformation from the mapping data for each of the 6 residues. In particular, a rotamer for the residue was randomly selected, then one of the translational/rotational positions was randomly chosen. Following concatenation, the minimization and free energy evaluation procedure, explained above, was repeated for each of the 100 structures.

As a further test, we repeated the entire calculation (i.e., both the mapping of the site and the construction of 64 'best' structures), using the CHARMM energy target function defined by Eq. 3 instead of the free energy target function defined by Eq. 2.

2.3. Free energy evaluation of hydroxyethylene isostere inhibitor analogs

In a related calculation, we evaluated the binding free energies for five trimeric hydroxyethylene isostere analogs, for which inhibition constants were measured by Dreyer et al. [40]. The structures of the five trimeric inhibitor analogs with sequences (Boc)Phe-(CH₂)X-Val(NH₂), where X=Gly, Ala, Npr, Ile, and Phe, were generated by homologous extension, starting with the MVT-101 structure that minimized the free energy target function. The substitutions were carried out with the QUANTA protein design tool, which maximally overlaps the common atoms of the new side chain with those of the template side chain [41]. Each structure was minimized for a total of 120 steps subject to a mass-weighted harmonic constraint of 20 kcal/mol/Å², and then the free energy was evaluated.

The calculations were repeated with the CHARMM energy target function, starting with the MVT-101 structure that minimized the same CHARMM target function. Since the substitution of Phe at position P3 results in steric clashes that are not completely removed by the constrained minimization, we do not constrain the atoms at this position. After minimization, the target function was calculated according to Eq. 3.

3. Results and discussion

3.1. Docking the MVT-101 inhibitor to HIV-1 protease

Sixty-four structures for MVT-101 were constructed by concatenating the two lowest energy conformations at each position (Tables 1 and 2). To provide perspective on the range of RMSD possibilities, we generated an additional 100 structures obtained by concatenating randomly selected states for different residues using the free energy target function. The 64 MVT-101 structures obtained by concatenating and refining the two lowest free energy structures have all atom RMSDs from the minimized crystal structure, ranging from 1.15 Å to 1.57 Å with an average of 1.36 Å (Fig. 1). By contrast, the RMSDs of the additional 100 structures obtained by combin-

ing randomly selected conformations, ranged from 1.47 Å to 2.93 Å with an average of 2.24 Å. Only 7 of these had RMSD values within the range of those chosen by free energy. The 64 structures generated with the CHARMM target function, Φ_{CHARMM} , in both mapping and assembly stages of the algorithm, range from 1.53 Å to 1.78 Å; have an average of 1.68 Å and show no correlation between RMSD and target function value (Fig. 2). Specifically, there is a group of structures which have very low CHARMM energy, but a high RMSD range of 1.69 to 1.72 Å.

We also calculated the CHARMM objective function for the 64 structures generated using free energy, and found that CHARMM was able to rank these conformations almost as well as the free energy. However, as shown in Fig. 2, the CHARMM potential performs poorly in the mapping stage, and using only the two best mapped positions, the fragment assembly stage fails to yield any good conformation. In other words, the best positions found for the individual residues are substantially different from their positions in the X-ray structure of the ligand if solvation and entropic effects are not taken into account.

3.2. Design of hydroxyethylene isostere inhibitor analogs

The relation between $\ln K_i$ for five hydroxyethylene isostere

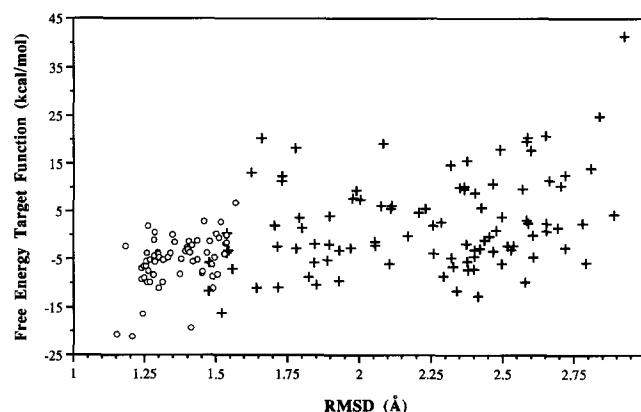


Fig. 1. Plot of free energy target function against the all atom RMSD from the minimized crystal for the 64 structures generated by concatenating the best and second best mapped residue positions (circles), and 100 structures generated by concatenation of randomly selected residue positions (crosses).

Table 3

The free energy values, Ω , and total CHARMM interaction energies computed for the respective set of analogs constructed from MVT-101 templates that had the lowest value of the target function

Free energy target function					
X^a	ΔG_h	E_{rl}^c	$T\Delta S_{sc}$	Ω^b	$\ln K_i$
Gly	-21.87	-1.91	5.41	-18.38	-11.94
Ala	-23.73	-2.00	5.61	-20.12	-15.16
Npr	-25.71	-2.12	6.71	-21.12	-16.81
Ile	-26.21	-2.20	6.68	-21.73	-17.59
Phe	-27.57	-1.98	6.60	-22.96	-20.39

CHARMM energy target function					
X^a	E_{VDW}^c	E_{ELC}^d	E_{rl}^e	$\ln K_i$	
Gly	-24.18	-1.91	-26.10	-11.94	
Ala	-26.73	-2.00	-28.73	-15.16	
Npr	-30.98	-2.12	-33.09	-16.81	
Ile	-28.89	-2.20	-31.08	-17.59	
Phe	-30.39	-1.98	-30.37	-20.39	

^aInhibitors with sequence: (Boc)Phe-(CH₂)X-Val(NH₂).

^bFree energy target function, Eq. 5.

^cVan der Waals component of the CHARMM interaction energy.

^dElectrostatic component of the CHARMM interaction energy.

^eCHARMM energy target function, Eq. 5.

analogs, studied by Dreyer et al., and free energies calculated according to Eq. 2 for the structures estimated by homologous extension are shown in Fig. 3 and listed in Table 3 [40]. Assuming the inhibition constants, K_i , were obtained under the quasi-equilibrium conditions of Michaelis-Menten kinetics [42,43], we expect a plot of ΔG against $\ln K_i$ to be a straight line, with a slope given by the thermal energy, $RT=0.6$ kcal/mol (at 300 K).

When free energy shifts are evaluated for the analogs, using the predicted structure as template, the slope is 0.80, and r exceeds 0.97 (Fig. 3, Table 3). When the CHARMM interaction energy is used, starting with its predicted structure as template (Fig. 3, Table 3), the slope is 1.03 and the correlation coefficient is 0.97. These results show that correct ranking of the analogs can be achieved with both traditional and free energy target functions. However, the use of free energy yields two substantial advantages. First, due to the assumed cancellation of van der Waals terms, the free energy is a 'smooth function' providing well defined, reliable values that are not

sensitive to small changes in the atomic coordinates. By contrast, the value of the interaction energy is heavily affected by the details of calculation such as the termination condition in the minimization. Second, the spacing between analogs (i.e. the slope of the line in Fig. 3) is more accurate using free energy. Indeed, because what is measured (i.e., $\ln K_i$ or pIC_{50}) is related to the binding free energy, the correlation of the inhibition constant and interaction energy must be in error by some combination of solvation and entropy, which apparently varies from analog to analog. However, since the error does not vary greatly, an empirical correction is possible. For example, Holloway et al. used a training set of 33 inhibitor structures to relate calculated interaction energies (from the MM2X force field) to experimental inhibition constants [44]. This relation, which maps interaction energy into pIC_{50} values, has been applied to 16 additional inhibitors. The correlation between predicted pIC_{50} values and measured inhibition constants was 0.937, which must be considered good agreement.

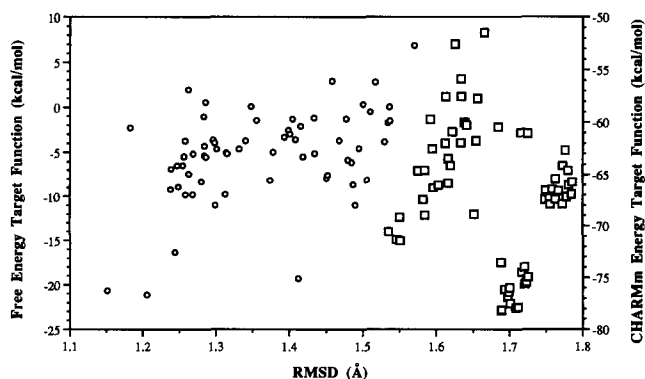


Fig. 2. Plot of free energy target function (Eq. 2) versus all atom RMSD for the 64 structures generated by free energy mapping (circles), and CHARMM energy target function (Eq. 3) versus all atom RMSD for the 64 structures constructed by CHARMM energy mapping (solid squares).

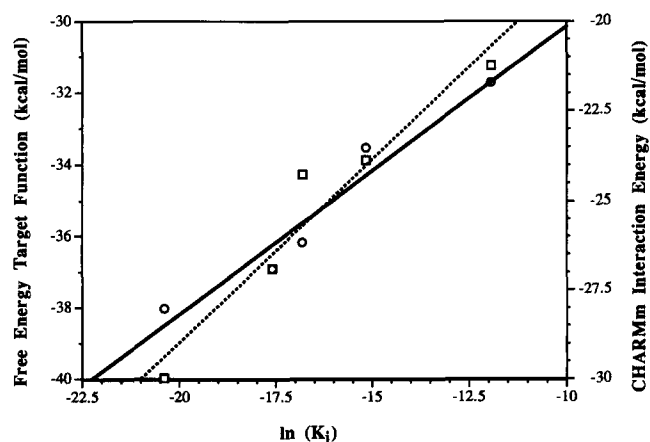


Fig. 3. Plot of free energy target function and $\ln K_i$ with linear regression (circles and solid line) [slope = 0.80, $r = 0.970$], and the plot of the CHARMM interaction energy and $\ln K_i$ with linear regression (squares and small dashed line), [slope = 1.03, $r = 0.967$].

References

- [1] Verlinde, C.L.M.J. and Hol, W.G.J. (1994) *Structure* 2, 577–587.
- [2] Kuntz, I.D. (1992) *Science* 257, 1078–1082.
- [3] Meng, E.C., Gschwend, D.A., Blaney, J.M. and Kuntz, I.D. (1993) *Proteins* 17, 266–278.
- [4] Rosenfeld, R., Vajda, S. and DeLisi, C. (1995) *Annu. Rev. Biophys. Biomol. Struct.* 24, 677–700.
- [5] Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. and Ferrin, T.E. (1982) *J. Mol. Biol.* 161, 269–288.
- [6] Brucoleri, R.E. and Karplus, M. (1987) *Biopolymers* 26, 137–168.
- [7] Leach, A.R. (1992) *J. Comp. Chem.* 13, 733–748.
- [8] Gulukota, K., Vajda, S. and DeLisi, C. (1996) *J. Comp. Chem.*, in press.
- [9] Caflisch, A., Niederer, P. and Anliker, M. (1992) *Proteins* 13, 223–230.
- [10] Goodsell, D.S. and Olson, A.J. (1990) *Proteins* 8, 195–202.
- [11] Hart, T.N. and Read, N.J. (1992) *Proteins* 13, 206–222.
- [12] Yue, S. (1990) *Protein Eng.* 4, 177–184.
- [13] Totrov, M. and Abagyan, R. (1994) *Nat. Struct. Biol.* 1, 259–263.
- [14] Judson, R.S., Jaeger, E.P. and Treasurywala, A.M. (1994) *J. Mol. Struct. (Theochem.)* 308, 191–206.
- [15] Jones, G., Willett, P. and Glen, R.C. (1995) *J. Mol. Biol.* 245, 43–53.
- [16] Goodford, P.J. (1985) *J. Med. Chem.* 28, 849–857.
- [17] Lawrence, M.C. and Davis, P.C. (1992) *Proteins* 12, 31–41.
- [18] Bohm, H.J. (1992) *J. Comp.-Aid. Mol. Des.* 6, 131–147.
- [19] Moon, J.B. and Howe, J. (1991) *Proteins* 11, 314–328.
- [20] Verlinde, C.L.M.J., Rudenko, G. and Hol, W.G.J. (1992) *J. Comp.-Aid. Mol. Des.* 6, 131–147.
- [21] Caflisch, A., Miranker, A. and Karplus, M. (1993) *J. Med. Chem.* 36, 2142–2167.
- [22] Rotstein, S.H. and Murcko, M.A. (1993) *J. Med. Chem.* 36, 1700–1710.
- [23] Eisen, M.B., Wiley, D.C., Karplus, M. and Hubbard, R.E. (1994) *Proteins* 19, 199–221.
- [24] Straatsma, T.P. and McCammon, J.A. (1991) *Methods Enzymol.* 202, 497–511.
- [25] Shoichet, B.K., Bodian, D.L. and Kuntz, I.D. (1992) *J. Comp. Chem.* 3, 380–397.
- [26] Shoichet, B.K. and Kuntz, I.D. (1991) *J. Mol. Biol.* 221, 237–246.
- [27] Vajda, S., Weng, Z., Rosenfeld, R. and DeLisi, C. (1994) *Biochemistry* 33, 13977–13988.
- [28] Vajda, S., Weng, Z. and DeLisi, C. (1996) *Protein Eng.* (in press).
- [29] Miller, M., Schneider, J., Sathyanarayana, B.K., Toth, M.V., Marshall, G.R., Clawson, L., Kent, S.B. and Wlodawer, A. (1989) *Science* 246, 1149–1152.
- [30] Eisenberg, D. and McLachlan, A.D. (1986) *Nature* 319, 199–203.
- [31] Pickett, S.D. and Sternberg, M.J.E. (1993) *J. Mol. Biol.* 231, 825–839.
- [32] Novotny, J., Brucoleri, R.E. and Saul, F.A. (1989) *Biochemistry* 28, 4735–4749.
- [33] Jackson, R.M. and Sternberg, M.J.E. (1995) *J. Mol. Biol.* 250, 258–275.
- [34] Brooks, B., Brucoleri, R.E., Olafson, B., States, D.J., Swaminathan, S. and Karplus, M. (1983) *J. Comp. Chem.* 4, 187–217.
- [35] Swain, A.L., Miller, M.M., Green, J., Rich, D.H., Schneider, J., Kent, S.B.H. and Wlodawer, A. (1990) *Proc. Natl. Acad. Sci. USA* 87, 8805–8809.
- [36] Jaskolski, M., Tomasselli, A.G., Sawyer, T.K., Staples, D.G., Heinrichson, R.L., Schneider, J., Kent, S.B.H. and Wlodawer, A. (1991) *Biochemistry* 30, 1600–1609.
- [37] Erickson, J., Neidhart, D.J., VanDrie, J., Kempf, D.L., Wang, X.C., Norbeck, D.W., Plattner, J.J., Rittenhouse, J.W., Turon, M., Wideburg, N., Kohlbrenner, W.E., Simmer, R., Helfrich, R., Paul, D.A. and Knigge, M. (1990) *Science* 249, 527–533.
- [38] Fitzgerald, P.M.D., McKeever, B.M., Van Middlesworth, J.F., Springer, J.P., Heimbach, J.C., Leu, C., Herber, W.K., Dixon, R.A.F. and Darke, P.L. (1990) *J. Biol. Chem.* 265, 14209–14219.
- [39] Ponder, J. and Richards, F. (1987) *J. Mol. Biol.* 193, 775–791.
- [40] Dreyer, G.B., Lambert, D.M., Meek, T.D., Carr, T.J., Tomaszek Jr., T.A., Fernandez, A.V., Bartus, H., Cacciavillani, E., Hassell, A.M., Minnich, M., Petteway Jr., S.R., Metcalf, B.W. and Lewis, M. (1992) *Biochemistry* 31, 6646–6659.
- [41] QUANTA version 4.0. Molecular Simulations Inc. 16 New England Executive Park, Burlington, MA 01803–5297.
- [42] Stryer, L. (1988) *Biochemistry*, 3rd edn., W.H. Freeman and Co., New York, pp. 187–195.
- [43] Creighton, T.E. (1993) *Proteins: Structures and Molecular Properties*, 2nd edn., W.H. Freeman and Co., New York, pp. 386–389.
- [44] Holloway, M.K., Wai, J.M., Halgren, T.A., Fitzgerald, P.M.D., Vacca, J.P., Dorsey, B.D., Levin, R.B., Thompson, W.J., Chen, L.J., deSolms, S.J., NleGaffin, N., Ghosh, A.K., Giuliani, E.A., Graham, S.L., Guare, J.P., Hungate, R.W., Lyle, T.A., Sanders, W.M., Tucker, T.J., Wiggins, M., Wiscourt, C.M., Woltersdorf, O.W., Young, S.D., Darke, P.L. and Zugay, J.A. (1995) *J. Med. Chem.* 38, 305–317.