PROTEINS: Structure, Function, and Genetics 4:31–47 (1988)

Structure and Energetics of Ligand Binding to Proteins: *Escherichia coli* Dihydrofolate Reductase-Trimethoprim, A Drug-Receptor System

Pnina Dauber-Osguthorpe, Victoria A. Roberts, David J. Osguthorpe, Jon Wolff, Monique Genest, and Arnold T. Hagler

The Agouron Institute, La Jolla, CA 92037

ABSTRACT **A study of the binding of the antibacterial agent trimethoprim to *Escherichia coli* dihydrofolate reductase was carried out using energy minimization techniques with both a full, all-atom valence force field and a united atom force field. Convergence criteria ensured that no significant structural or energetic changes would occur with further minimization. Root-mean-square (RMS) deviations of both minimized structures with the experimental structure were calculated for selected regions of the protein. In the active site, the all-atom minimized structure fit the experimental structure much better than did the united atom structure. To ascertain what constitutes a good fit, the RMS deviations between crystal structures of the same enzyme either from different species or in different crystal environments were compared. The differences between the active site of the all-atom minimized structure and the experimental structure are similar to differences observed between crystal structures of the same protein.**

Finally, the energetics of ligand binding were analyzed for the all-atom minimized coordinates. Strain energy induced in the ligand, the corresponding entropy loss due to shifts in harmonic frequencies, and the role of specific residues in ligand binding were examined. Water molecules, even those not in direct contact with the ligand, were found to have significant interaction energies with the ligand. Thus, the inclusion of at least one shell of waters may be vital for accurate simulations of enzyme complexes.

Key words: protein simulation, dihydrofolate reductase

INTRODUCTION

With the advent of supercomputer resources, theoretical energetic based simulation methods, such as minimization, molecular dynamics, and Monte Carlo, can be applied to the design of biological systems. Applications range from the design of therapeutic agents such as peptide hormone analogs[1] to the design of specific ligands to mediate physiological enzyme reactions and extend even to the design of proteins themselves. There are two basic levels at which such theoretical simulations can contribute to the design process. The first is a semiquantitative fashion in which results of structural variations in the molecular system of interest are used to supplement X-ray crystallographic information, to provide insight into the conformational properties of the system, and to allow visualization of interactions of interest. These structural insights obtained in general from observation on an interactive graphics system with perhaps some information on energetics of interaction are used by the chemist in what is basically an iterative synthetic design process. The second level is to use the results in a quantitative manner, actually calculating and comparing binding energies and free energies, differences in conformational free energies, and quantitative structural changes on molecular modification. More fundamentally, in the latter application, we are attempting to understand these processes at the level of the underlying intermolecular forces. Implicit in both levels is the ability to simulate *reliably* the systems of interest. At the first level, quantitative accuracy is not as crucial, although clearly the qualitative behavior of the properties being investigated must be reproduced. The quantitative analysis requires still more rigor.

We have recently embarked on such studies both for smaller systems such as the design of peptide hormone analogs,[1,2] where the nature of the problem basically determines that the approach be at the "insight" level, and for protein-ligand systems. In the latter case, extensive experimental data allow us to attempt to ultimately describe and understand the system at the level of the fundamental interatomic forces. As the first step in this study, we have carried out an extensive study of the *Escherichia coli* complex of dihydrofolate reductase (DHFR) with trimethoprim in which the adequacy of various techniques and approximations, such as the reliability of the united

Received September 9, 1987; revision accepted May 3, 1988.

Address reprint requests to Arnold T. Hagler, The Agouron Institute, 505 Coast Blvd. South, La Jolla, CA 92037.

Victoria A. Roberts' present address is: Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, California 92037; Jon Wolff's present address is: Department of Pediatrics, School of Medicine, University of California at San Diego, La Jolla, California 92093; Prina Dauber-Osguthorpe and David J. Osguthorpe's present address is: School of Chemistry, University of Bath, Bath, BA2 7AY, England; Monique Genest's present address is: Centre de Biophysique, Moléculaire (CNRS), Université d' Orléans, 45071 Orléans Cedex 2, France.

atom approximation and the convergence criteria required, are explored. In addition, we have addressed the question of what constitutes agreement with experiment in the simulation of a protein system. Finally, we have interpreted the results obtained in these preliminary simulations at the "semiquantitative" level to begin to understand the factors involved in ligand binding such as the strain induced by the protein on the ligand, and the role of "structural" water molecules in the system.

We have chosen the dihydrofolate system for several reasons. First, DHFR is a vital enzyme in the metabolic pathway of all organisms. It catalyzes the reduced nicotinamide-adenine dinucleotide phosphate (NADPH)-dependent reduction of dihydrofolate to tetrahydrofolate, which is then utilized in the biosynthesis of thymidylate, purines, and amino acids.[3–5] Inhibition of DHFR results in depletion of tetrahydrofolate and consequent cell death. Although there is high homology between DHFRs from different species, several inhibitors demonstrate a high degree of species selectivity. One of these is trimethoprim, which binds to bacterial DHFRs about 3,000 times more strongly than to vertebrate DHFRs.[6,7] This has led to its widespread use as an antibiotic. In addition, since efficient DHFR action is particularly important in rapidly dividing cell lines, inhibitors of vertebrate DHFR such as methotrexate have found use in the treatment of various types of cancer.[8–10]A second important reason for our choice of DHFR is the availability of a large number of highly refined structures of reductases from various species in several liganded states, including among them the structure of the *E. coli* DHFR·trimethoprim binary complex.[11–17] With this wealth of structural information, we are in a position to extend our simulations to DHFRs of many species and to assess rigorously the validity of the theoretical techniques by detailed comparison with the experimental structures.

Use of supercomputers such as the Cyber 205 has allowed simulations of large systems to be carried out that would have been impossible just a few years ago. This calculation constitutes the first simulation in which the same force field used for small molecules, including cross-terms and the explicit inclusion of all hydrogen atoms, has been used for a protein system. Up to this time, large protein simulations have used the united atom approximation, where nonpolar hydrogen atoms are omitted and the van der Waals radius of the heavy atoms to which they were attached is increased accordingly. This approximation reduces the number of atoms by about 30%, the resulting calculation time per iteration by about 50%, and also reduces the total number of iterations that are required to achieve a given convergence criterion. We were concerned with the adequacy of this approximation for application to ligand binding, and we have taken advantage of the power of the Cyber 205 to compare the all-atom simulation with the united-atom simulation.

METHODS AND MATERIALS

Theoretical Techniques

The potential energy of a system can be expressed in terms of an analytical representation of all internal degrees of freedom and van der Waals and electrostatic interactions. From this expression, such properties as the minimum energy structure, vibrational spectra, or dynamic trajectory of a molecular system may be calculated.[2,18] In addition to previously determined parameters for nonbonded interactions of amides [19,20] and acids, [21–23] we have optimized the parameters used in the potential energy equation for the functional groups in the protein and ligand giving a *complete* "first-generation" force field for this work (see Appendix).[24] The conformational entropy of the system can be obtained from the vibrational frequencies in the harmonic approximation from the Einstein equations.[25,26]

System Setup

Simulations were carried out on the *E. coli* binary complex of DHFR with trimethoprim (TMP) and 259 waters. (The initial protein coordinates were obtained from molecule 2 of the asymmetric unit in the crystal structure of the binary complex.) Crystallographically determined water molecules within 7 Å of TMP and those that form hydrogen bonds to at least two protein residues were included. The latter may be important in the maintenance of the structural integrity of the protein. The local environment around each crystal water was examined to determine possible hydrogen bonds with the protein and TMP, and hydrogen atoms on the water molecules were added to fulfill these hydrogen bonds. A box of water molecules was generated around the protein from which waters were selected so that all solvent within 7 Å of TMP and solvent within 3 Å of all charged groups was included. In the complete model, *all* hydrogen atoms, polar and nonpolar, were added to the protein and TMP in standard geometries to complete the setup of the system. In the united atom simulation, the hydrogen atoms of CH, CH_2, and CH_3 groups were merged with the parent carbon. (All parameters were the same for the united atom model except that the van der Waals radius of the heavy atom was increased to compensate for the lack of hydrogens.)

Minimization Procedure

In order to prevent artifactually large forces because of initial bad contacts[27] from causing major structural perturbations at the outset, the protein complex was gradually annealed. First, the energy was minimized allowing only the waters to relax with the oxygen atoms gently constrained in order to adjust the positions of the water hydrogens and relieve clashes between some of the poorly determined crystal waters and the protein. In the next stage, all atoms of the system were allowed to relax with the heavy atoms lightly constrained to relieve unfavor-

able interactions between poorly defined crystallographic atoms or between the hydrogen atoms built onto the protein and TMP. Finally, the energy of the system was minimized with respect to all 3N Cartesian coordinates of the system until the average derivative was less than 0.0006 kcal $mol^{-1}Å^{-1}$ (corresponding to a maximum derivative of 0.022 kcal $mol^{-1}Å^{-1}$ and a change in the total energy of less than 0.00005 kcal mol^{-1} in the last step). This unconstrained minimization required 10,800 iterations, which took about 15 hours of CPU time on the Cyber 205 (4.7 seconds per iteration). The minimization of the united atom model took 2.5 seconds per iteration and required 9,460 steps to converge.

Convergence Criteria

The supercomputer has allowed us to achieve convergence in this minimization as indicated by the size of derivatives and by the average RMS movement of approximately 10^{-5} Å per step during the final 1,000 steps of the minimization. This convergence criterion is substantially more rigorous than those that have been used standardly in protein minimizations (e.g., 0.2 kcal mol^{-1} change in energy per step[28] or the RMS deviation of derivatives less than 0.1 kcal $mol^{-1}Å^{-1}$).[29] For the all-atom minimization, an average derivative of 0.1 kcal $mol^{-1}Å^{-1}$ corresponded to an energy 140 kcal mol^{-1} greater than that of the fully minimized complex, and the protein had an RMS deviation of 0.7 Å from the fully minimized complex at this time. Preliminary results on minimizations of other DHFR complexes show similar RMS and energy differences at this point in the minimization. Thus, significant changes in both energy and structure can occur between a structure minimized to an average derivative of 0.1 kcal $mol^{-1}Å^{-1}$ and full convergence. Therefore, if we want to compare the structures of different minimized complexes or determine how much the minimized structure varies from the experimental complex, it appears that a more rigorous convergence criterion than 0.1 kcal $mol^{-1}Å^{-1}$ is required.

COMPARISON OF SIMULATED DHFR COMPLEXES WITH EXPERIMENTAL STRUCTURES–RMS DEVIATIONS

Both the minimized all-atom structure and the minimized united-atom structure were compared with the experimental structure in order to assess overall fit and the effect of omitting nonpolar hydrogens. This was achieved by the superposition of each minimized structure onto the experimental structure using a least-squares fit of the heavy atoms of the secondary structures, i.e., the α helices and β strands. The use of the structurally conserved framework of the protein for superposition avoids problems with poorly defined regions or loops at the surface, as discussed below. The root mean square (RMS) deviation between several structural segments of interest was then calculated. Figure 1 gives ribbon representations of the *E. coli* DHFR molecule with bound TMP, depicting schematically the RMS deviations of the C_α atom of each residue in the minimized all-atom and united-atom structures.

The All-Atom Model

The minimized *E. coli* binary complex has an overall RMS deviation of the heavy atoms, excluding water oxygens, of 1.23 Å when compared with the experimental structure (Table I). Even more important than the overall RMS value is the distribution of the deviations within the complex. The largest deviations (RMS > 2.0 Å) occur in regions such as the external loop between βC and βD (residues 64–68 and 71) or in the large external loop between βF and βG (residues 121, 122, and 130–132) both of which exhibit substantial disorder in the experimental structure (Fig. 1). Virtually all residues with RMS deviations greater than 1 Å are on the surface of the protein with their side chains extending outward or along the surface of the protein. Thus, the areas affected by crystal disorder and the crystal environment are those that are the least well reproduced in the minimization, which is expected since the crystal environment has not been included in this simulation. On the other hand, the active-site region, which consists of all residues within 4 Å of TMP, has an RMS deviation of 0.58 Å. The β strands and α helices, which comprise the structural scaffold of the protein, also have significantly lower RMS deviations than the overall protein. This pattern of deviations is also observed in the crystal between the two molecules in the asymmetric unit as discussed below. The minimized structure retains the conserved structures and active site region present in both molecules in the asymmetric unit.

These considerations serve to reinforce the need to examine the components of the RMS deviation obtained between a theoretical simulation and the observed structure. A single number is inadequate as a measure of reliability; it can mask a fairly good fit to well-ordered structural data in important regions of the protein such as the active site because of extremely poor fit in loop regions.

The United-Atom Model

As seen from Tables I and II and Figure 1, the all-atom structure fits the experimental structure consistently better than the united-atom calculation. The fit to the experimental structure at the active site is of special importance because one of our major objectives is to understand ligand binding to DHFR. Whereas deviations of the residues of the protein within 4 Å of the ligand are ~0.6 Å for the complete-atoms model, the deviations are almost twice as large

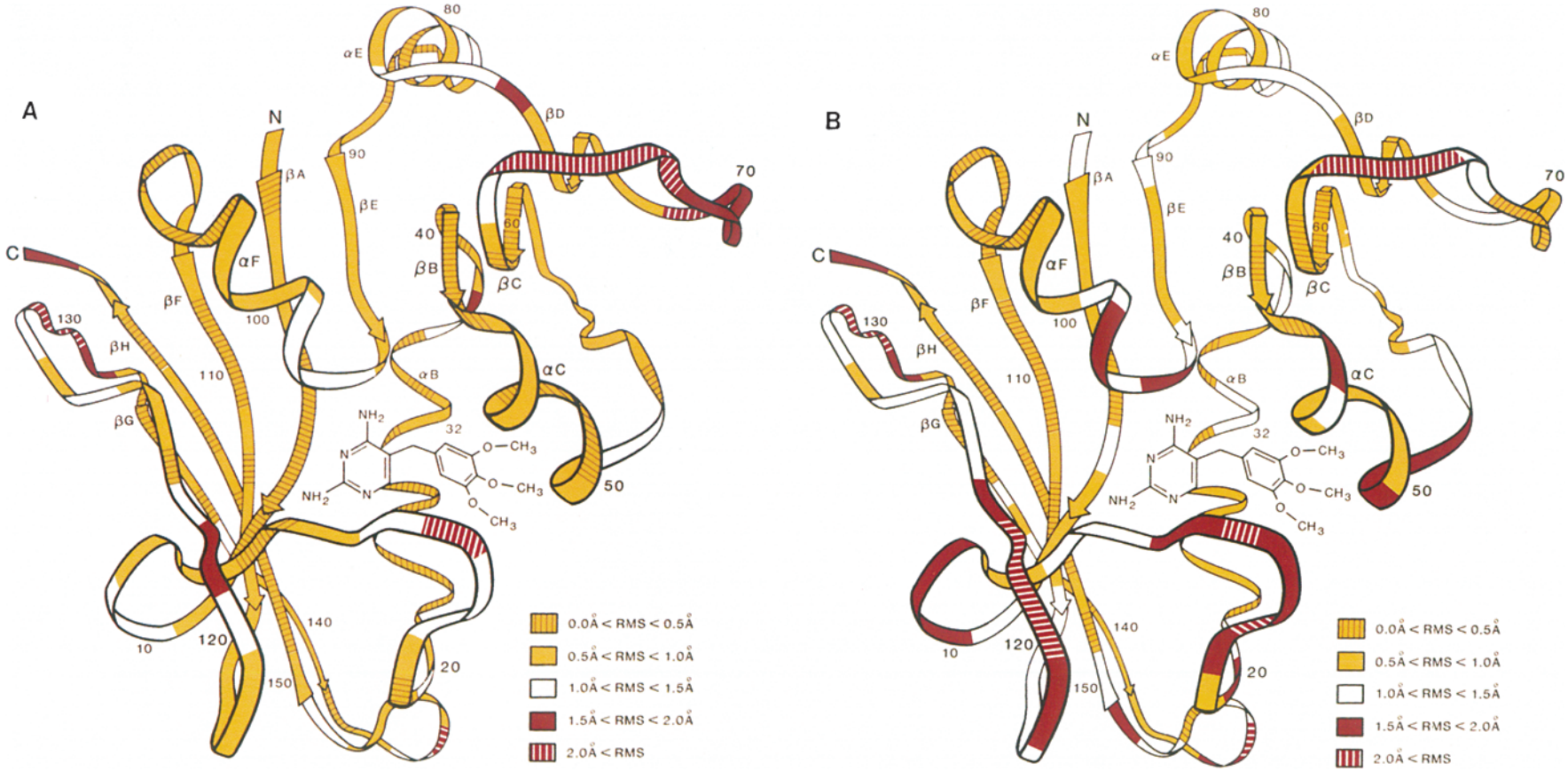


Fig. 1. Schematic representation of the *E. coli* DHFR molecule, with a bound trimethoprim (TMP) molecule, depicting the secondary structure in the enzyme, showing RMS deviations of Cα atoms between the experimental structure and the all-atom minimized structure (the heavy atoms of the secondary structures were fitted). The all-atom structure (**a**) and united-atom structure (**b**) are shown. In the all-atom calculated structure (**a**) most of the residues with large RMS values (greater than 2 Å lie either in an external loop between βC and βD (residues 64–68 and 71) or in a large external loop between βF and βG (residues 121, 122, and 130–132), both of which exhibit substantial disorder in the experimental structure. Virtually all residues with RMS deviations greater than 1 Å in this structure are on the surface of the protein with their side chains extending outward or along the surface of the protein. The active site region, which consists of all residues within 4 Å of TMP, consists of residues with RMS deviations less than 1 Å. Secondary structures near the active site—most of βA, αB, βF, βG, and βH—have similarly small deviations. These last three β strands are hydrogen bonded to N2 of the pyrimidine ring of the inhibitor through a water molecule. Thus, the areas affected by crystal disorder and the crystal environment are those that are the least well reproduced in the minimization, which is expected since the crystal environment has not been included in this simulation, while the areas around the active site are well reproduced. The united-atom structure (**b**) has only 29 residues with Cα RMS deviations less than 0.5 Å, far fewer than the all-atom structure (47), and a much larger number of residues with deviations greater than 1.5 Å (33 vs. 18), showing a general overall poorer fit to the experimental structure than the all-atom minimized structure has. The only region where the united-atom structure fits the experimental structure better than the all-atom structure is in the disordered exterior loop between βC and βD (which, in fact, makes the "better fit" meaningless). Several residues near the active site have Cα deviations greater than 1.0 Å.

TABLE I. RMS Deviations for Heavy Atoms and C_α Atoms Between the Minimized and the Observed Structures of the *E. coli* Binary Complex (in Å)*

Region	Experimental	
	With minimized all hydrogens**	With minimized united atoms**
All heavy atoms†	1.23 (1.07)**	1.48 (1.22)**
TMP	0.81 0.00	0.80 0.00
Active site‡	0.58 (0.54)	1.03 (0.92)
α helices	1.10 (0.72)	1.50 (0.95)
β strands	0.82 (0.66)	1.02 (0.80)

*Disordered regions of the experimental structure where atomic positions could not be determined (residues Pro^{66}, Glu^{129}, and Asp^{131}) have been omitted from all RMS calculations.
**Nos. in parentheses are the RMS deviations of C_α atoms. †Deviations of water molecules are not included. ‡Protein residues with 4 Å of TMP.

TABLE II. RMS Deviations Between the Heavy Atoms of the Minimized and Observed Structures of Residues Within 4 Å of Trimethoprim (Deviations Are in Å)

Region	Residue	All hydrogens			United-atom model		
		RMS	Max. dev.	C^α dev.	RMS	Max. dev.	C^α dev.
βA	Ile^5	0.56	0.97	0.39	0.77	1.45	0.49
	Ala^6	0.35	0.66	0.15	1.18	1.50	1.16
	Ala^7	0.30	0.47	0.23	0.62	0.73	0.73
αB	Asp^{27}	0.42	0.83	0.35	0.58	0.71	0.59
	Leu^{28}	0.44	0.59	0.35	0.99	1.47	0.73
	Trp^{30}	0.27	0.33	0.27	0.55	0.99	0.20
	Phe^{31}	0.50	0.88	0.67	1.10	1.49	1.25
αC	Thr^{46}	0.93	1.37	0.87	1.34	2.40	0.71
	Ser^{49}	0.64	1.17	0.41	1.47	2.59	0.93
	Ile^{50}	0.84	1.24	0.66	1.82	3.07	1.54
β-turn	Leu^{54}	0.62	1.05	0.60	1.08	1.18	1.12
βE	Ile^{94}	0.85	1.02	0.93	1.24	1.44	1.34
βF	Tyr^{100}	0.56	0.71	0.57	0.68	0.92	0.52
	Thr^{113}	0.38	0.47	0.39	0.46	0.57	0.51
Total*		0.58	1.37	0.54	1.03	3.07	0.92

*The total RMS, maximum deviation, and C^α deviations are the RMS deviation of all heavy atoms of the protein in the active site, the largest atomic deviation of the protein in the active site, and the RMS deviation of all C^α atoms of the protein in the active site, respectively.

for the united-atom model (Table II). The large discrepancies in the active site region resulting from the use of the united-atom model are a severe drawback and raise serious concerns about using this approximation in the study of ligand binding.

Variations in Experimental Structures

Use of isolated protein molecule in ligand-binding and drug-design studies

A problem that arises in the above analysis as well as in all studies of this type, where the crystalline environment has not been taken into account in the calculation and where, in addition, there are several sources of experimental uncertainty, is what is a good or reasonable fit, i.e., are we reproducing the essential features of the enzyme that is being simulated. A related, crucial question of practical importance is what is the minimal system we need to include in order to study the energetics of ligand binding and understand the fundamental protein-ligand interactions (e.g., to obtain reliable correlations in designing ligands for drug applications). Crystal disorder and the crystal environment affect not only our comparisons of the minimized structure with the experimental structure but also cause variations in different molecules of the same protein in the asymmetric unit in a crystal. In addition, the experimental error for the atomic positions is of the order of 0.2 Å[11] and problems in interpretation and fitting of regions of strongly anisotropic electron density distributions can lead to still larger deviations. In order to quantitate these effects and obtain some objective measure of a "reasonable" fit to an observed protein structure, we found it useful to examine the relationship between identical proteins in different crystalline environments (two molecules in the asymmetric unit or in different space groups), between homologous pro-

TABLE III. RMS Deviations Between Experimental Structures of Proteins*

Protein	Resolution	R-factor	Whole structure*	Secondary structure**
1. *E. coli* DHFR·TMP†,[11]	2.3	0.19	1.15(0.76)	0.63(0.30)
2. *E. coli* DHFR·MTX†,[16]	1.7	0.16	1.08(0.68)	0.58(0.25)
3. *E. coli* DHFR·TMP[11]	2.3	0.19		1.01(0.97)‡
L. casei DHFR·MTX·NADPH[17]	1.7	0.15		
4. *E. coli* DHFR·TMP[11]	2.3	0.19		1.61(1.46)‡
Chicken liver DHFR·TMP·NADPH[11]	2.2	0.21		
5. *L. casei* DHFR·MTX·NADPH[17]	1.7	0.15		1.60(1.49)‡
Chicken liver DHFR·TMP·NADPH[11]	2.2	0.21		
6. Trypsinogen[30]	1.8	0.23	1.15(0.59)	0.63(0.47)
Trypsinogen[31]	1.65	0.22		
7. Myoglobin[32]	2.0	0.23	0.67(0.38)	0.41(0.21)
Myoglobin[33]	1.4			
8. Ribonuclease[34]	2.0	0.16	0.71(0.20)	0.61(0.18)
Ribonuclease[35]	1.5	0.24		
9. Trypsin, pH 5.0[36]	1.7	0.18	0.39(0.45)	0.33(0.39)
Trypsin, pH 8.0[37]	1.55	0.19		
10. Trypsinogen, CH_3OH-H_2O[37]	1.8	0.18	0.38(0.44)	0.30(0.39)
Trypsinogen, 2.4M $MgSO_4$[37]	1.8	0.20		
11. Trypsin, orthorhombic[37]	1.55	0.19	0.63(0.51)	0.51(0.45)
Trypsin, trigonal[37]	1.7	0.20		
12. Mercury trypsinogen, 170K[36]	1.7	0.19	0.44(0.46)	0.35(0.39)
Mercury trypsinogen, 103K[36]	1.7	0.21		

*The binary *E. coli* complex contains two DHFR molecules in the asymmetric unit. The RMS deviation between the heavy atoms of the two proteins in the asymmetric unit was 1.15 Å when the secondary structures were fit (entry 1). If only the secondary structures, which do not include disordered external loops affected by crystal environment, are considered, the RMS deviation drops to 0.63 Å. Similar results are found for the comparison of the two independent molecules in the asymmetric unit of *E. coli* complex with the ligand methotrexate (entry 2). Therefore, experimental uncertainty and actual differences in the experimental structures (or packing effects, less likely for the internal residues) leads to overall deviations of ≈ 0.6 Å for heavy atoms of the secondary structure, of the same order as those found for residues about the active site in the all-atom calculated structure. The RMS deviation between selected main chain atoms of Chymotrypsinogen A (excluding residues 33–38, 142–153, and 215–222) has been calculated[38] and is similar to the RMS deviations of the C_α atoms in the reductase crystals (1 & 2). Larger deviations are found in the comparison of the other pairs of dihydrofolate reductases (entries 3–5). Proteins solved independently (entries 6–8) and the same proteins crystallized under different conditions with the crystal structures refined by the same methods (entries 9–12) show RMS deviations in their secondary structures similar to those obtained in the comparisons of molecules in the asymmetric unit discussed above (0.30–0.63 Å). The RMS deviations of the whole structure are somewhat larger reflecting the effects of crystal environment and disorder.

**Nos. in parentheses are the RMS deviation of the C_α atoms.

†Comparison of the two molecules in the asymmetric unit.

‡The RMS deviation is for all common heavy atoms of homologous residues of the secondary structure.

teins, between identical proteins solved using different refinement methods, and between identical proteins crystallized in different physical environments (pH, solvent, temperature).

A comparison of such structures presented in Table III reveals that differences in secondary structures of these identical proteins range from approximately 0.3 Å to 0.6 Å, while differences in the entire protein are slightly larger. A comparison of the results presented in Table III with those in Table I reveals that the simulation of the binary *E. coli* complex, with all hydrogen atoms included explicitly, shows a reasonable fit with the experimental structure. Areas buried in the complex have RMS deviations not significantly different from the ~0.5 Å deviations of secondary structures between independent DHFR molecules in the asymmetric unit or from differences in structures obtained from two methods of refinement of the same protein. The outer areas of the DHFR complexes show larger deviations, but this is expected where the experimental structure itself is poorly defined or where the crystal environment has not been simulated. For investigations of the buried active site, these results suggest that inclusion of the crystal environment may not be necessary. This is encouraging, if generally true for drug-design studies, but clearly many more such comparisons of calculated vs. experimental structures need to be carried out before we can have confidence in this as a general procedure. To compare our results rigorously with the crystal structure, however, the crystal environment will have to be included.

DETAILED ANALYSIS OF THE ACTIVE SITE

Ligand Environment

In the experimental structure, the TMP pyrimidine ring is observed to hydrogen bond to residues Asp^{27}, Ile^{5}, and Ile^{94} of the protein and the structurally conserved crystallographic water 33. These hydrogen

bonds are maintained in the minimized structure: heavy atoms that participate in the hydrogen bonds move by at most 0.3 Å from their locations in the observed structure (Fig. 2). In the minimized structure, two new hydrogen bonds have formed to N4 of the pyrimidine ring. One is with the oxygen atom of crystallographic water 39, which is 4.5 Å from N4 in the experimental structure but is 3.3 Å from N4 in the minimized structure and 2.3 Å from one of the hydrogens on N4. The other hydrogen bond involves the hydroxyl oxygen of Tyr^{100}, which is 3.5 Å from N4 in the experimental structure but 3.0 Å away in the minimized structure. Thus, the calculated minimum energy configuration shows a more elaborate hydrogen bond pattern around the pyrimidine ring than the experimental structure.

The pyrimidine ring is also involved in an extensive hydrogen bond network to the protein through water 33. Based on the experimental interatomic distances of the heavy atoms, this water can potentially hydrogen bond to residues Ala^6, Trp^{30}, Tyr^{111}, and Thr^{113}. Even though water molecules are not part of the "sequence" of the protein, the crystal water, 33, can be considered a "conserved" residue in the sense that a well-defined water molecule is observed in a similar position in all DHFR structures that have been determined. The hydrogen atoms of this water were originally built to interact with the carbonyl oxygens of Ala^6 and Tyr^{111}. In the calculated structure, the hydrogen atoms on the water have adjusted their positions so that one hydrogen maintains its hydrogen bond with Tyr^{111} and, in addition, forms a weak bond with the Ala^6 carbonyl, while the other hydrogen now is directed toward the oxygen atom of the Thr^{113} side chain (Fig. 2). Thus, the minimization has allowed the reorientation of the hydrogen atoms of this water molecule (not available from experiment) so that the hydrogen bond requirements of this region can be fulfilled while the relative positions of the water and surrounding protein as determined experimentally are maintained.

In the experimental structure, the trimethoxyphenyl moiety of the ligand is not involved in any hydrogen bonds to the protein, but lies in a water-filled cavity that extends out to the surface of the protein. Several residues of the protein have van der Waals contacts with this group (Leu^{28}, Phe^{31}, Ser^{49}, Ile^{50}, and Leu^{54}). This moiety does move substantially relative to the protein (RMS deviation of 0.97 Å) in the calculated structure, but the nonspecific van der Waals interactions between the ligand and these residues of between 3 to 4 Å are maintained.

Energetics of Specific Interactions of the Ligand With Neighboring Residues

The structural features discussed above are reflected in the energetics of interaction of the ligand with its neighboring residues in the active site cleft. Thus, in addition to the electrostatic interaction between the protonated pyrimidine ring and $Asp^{27,39}$ the coulombic interaction of the hydrogen bonded residues Ile^5, Ile^{94}, Tyr^{100}, water 33, and water 39 with the pyrimidine ring are -5.3, -4.2, -2.0, -8.1, and -0.2 kcal mol^{-1}, respectively. These are specific, directional interactions. In contrast, the most favorable electrostatic interaction between a residue of the protein and the trimethoxyphenyl moiety is -0.6 kcal mol^{-1} (Leu^{28}). Our quantitative results support the conclusions drawn from crystallographic evidence: favorable van der Waals interactions between the protein and the trimethoxyphenyl ring are an important contribution to the binding of the trimethoprim to *E. coli* DHFR.[11] The van der Waals interactions of the trimethoxyphenyl ring with Leu^{28}, Phe^{31} (a conserved residue), Ser^{49}, Ile^{50}, and Leu^{54} contribute -2.8, -3.2, -1.7, -3.6, and -1.1 kcal mol^{-1}, respectively, to the binding energy.

Ligand-Solvent Interactions

The largest ligand-water interaction is between the pyrimidine ring and water 33, as discussed above. This interaction is of the same magnitude as the strongest ligand-residue interactions (excluding the interaction with the charged Asp^{27}) and thus further supports the structural role of this water as being a conserved residue integral to the protein structure. It also serves to emphasize the importance of including this water and analogous "structural water molecules" in simulations; they can be as important as the protein residue themselves. Three more waters, 39, 63, and 91, hydrogen bond to trimethoprim with total interactions of -2.3, -2.3, and -5.6 kcal mol^{-1}. These interaction energies reflect the differences in distance and in relative orientation of the waters that arise from the balance between the hydrogen bond interactions with the inhibitor and interactions with the protein, with other water molecules, or with other parts of the ligand. In addition, there are several water molecules that do not form direct hydrogen bonds with the ligand but still exhibit significant (mainly electrostatic) interactions (waters 34, 52, 58, 59, 102, and 103 with energies of -2.8, -2.4, -1.5, -2.5, -4.1, and -4.4 kcal mol^{-1}). These results indicate that it is not sufficient to include only one "layer" of waters in order to simulate the ligand environment; waters near to, but not actually forming hydrogen bonds with, the ligand also may play a crucial role.

THERMODYNAMIC PROPERTIES

In general, strain energy is induced in a ligand on binding. The ligand loses vibrational, rotational, and translational free energy in exchange for the energy it gains by interaction with the enzyme. We have assessed several of these factors for the dihydrofolate reductase system in an attempt to extend further our understanding of the factors underlying the ligand binding in this system.

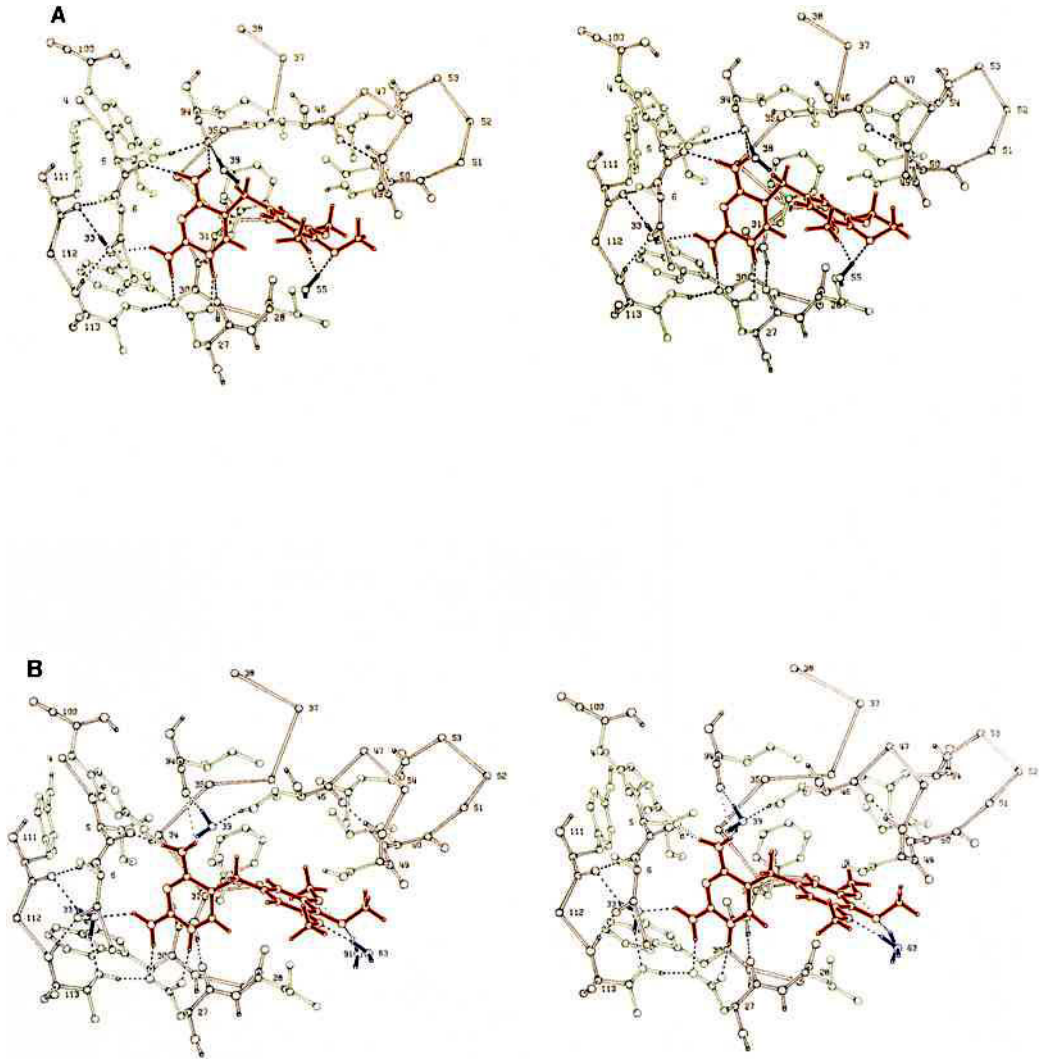
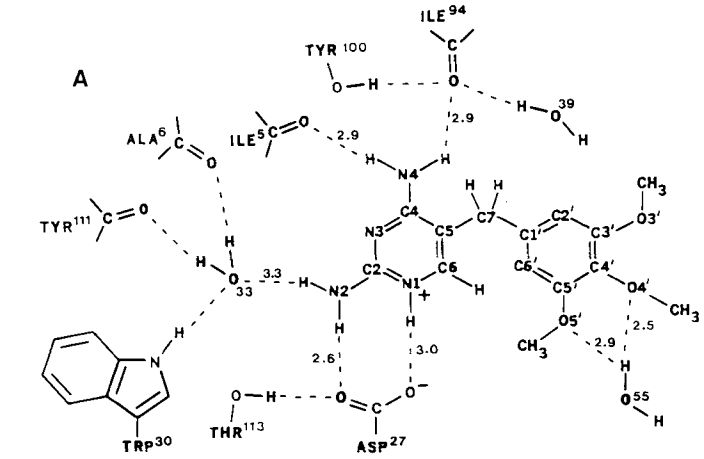


Fig. 2. Stereo plots of the active site of the observed structure (**A**) and the minimized structure (**B**) accompanied by schematic diagrams demonstrating the overall retention of the structure in the active site in the minimized structure. Adjustments of the hydrogen bond pattern around N4 of the pyrimidine ring of the inhibitor and around water 33, which is hydrogen bonded to N2 of the pyrimidine ring, are indicated. Trimethoprim and waters 33, 39, 55, 63, and 91 contain solid bonds in the stereo plots, while the bonds in the protein are open. Hydrogen bonds are indicated by dashed lines in both the stereo plots and the schematic diagrams. The numbers on these bonds in the schematic diagrams refer to the distance (in Å) between the heavy atoms (N and O).


A
ILE 94
TYR 100
ALA 6
ILE 5
TYR 111
TRP 30
THR 113
ASP 27
2.9
3.3
3.0
2.6
2.5

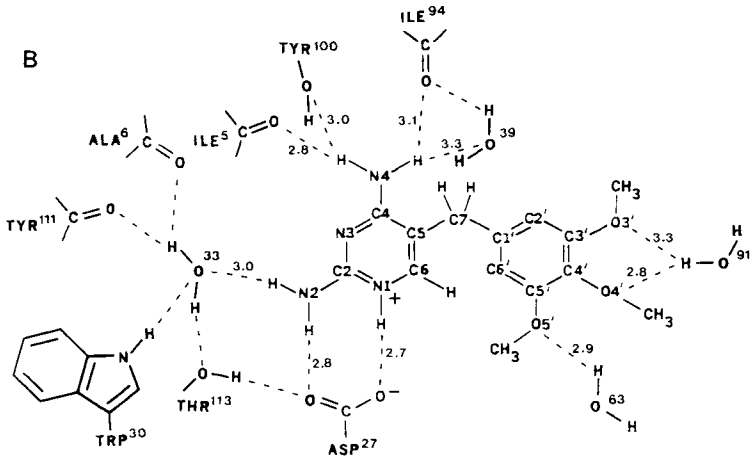

B
ILE 94
TYR 100
ALA 6
ILE 5
TYR 111
TRP 30
THR 113
ASP 27
3.0
3.1
3.3
2.8
2.7
2.9

TABLE IV. Energy and Entropy of Trimethoprim, Isolated and Bound in the Binary *E. coli* Complex (Energies Are in kcal mol^{-1})

Energy component	Isolated*	Bound
Valence		
Bond	27.9	27.8
Angle	16.6	16.7
Torsion	7.2	8.1
Out of plane	0.0	0.0
Cross terms	1.9	2.3
Intramolecular		
VDW	76.0	79.8
Coulomb	−92.3	−92.0
Total nonbond	−16.3	−12.2
Total intramolecular	37.5	42.7
Vibrational		
Enthalpy	213.4	218.7
T × entropy	24.2	24.2**
Free energy	189.2	194.4
Rotational		
Enthalpy	0.9	—
T × entropy	10.4†	—
Free energy	−9.5	—
Translational		
Enthalpy	0.9	—
T × entropy	10.9‡	—
Free energy	−10.0	—
Total		
Enthalpy	215.2	218.7
T × entropy	45.5	24.2
Free energy	169.7	194.4

*The conformation of the isolated ligand was determined by the minimization of TMP *in vacuo* starting from the conformation of the bound ligand. The final conformation is close to the lowest energy conformation predicted by Koetzle *et al.*[40] and one of the four lowest energy conformations predicted by the energy surface map using flexible geometry we have calculated.[41]

**The entropy for the bound ligand is calculated from all 3N normal modes.[26] The eigenvalues for these normal modes are all non-zero with the lowest frequency equal to 50.8cm^{-1}. The normal modes of the 14 lowest frequencies include translational and rotational components, reflecting the conversion of these motions into vibrational motions for the bound ligand. Translational movement was determined by a change in the center of mass upon displacement along each normal mode, and rotational movement was determined by calculation of the total angular momentum, the sum of R_i X mr_i for all atoms of the trimethoprim molecule, where R_i is the vector from the center of mass of trimethoprim to atom i, and r_i is the displacement of atom i along the normal mode (minus any translational movement). Another reduction for the bound ligand arises from the selection of one of the four low-energy conformations for the isolated ligand as mentioned in the first footnote, above.

†The rotational entropy was calculated using the rotational partition function for a nonlinear molecule with a temperature of 298°K. The product of the principal moments of inertia for the isolated trimethoprim molecule was determined to be 5.45×10^{-131}kg^3·m^6.

‡The translational entropy was calculated using the translational partition function assuming a standard state concentration of one mole per liter at 298°K for TMP.

Strain Energy Induced in Ligand on Binding

The energetics of binding of trimethoprim to DHFR are summarized in Table IV. A detailed summary of the procedure for obtaining these quantities is given in the footnote to this table. The energy has been partitioned into components to reflect the distribution of strain in the isolated and bound enzyme. The total intramolecular energy (valence and nonbond) of trimethoprim in the isolated molecule is 37.5 kcal mol^{-1}, while the corresponding energy of trimethoprim in the protein complex is 42.7 kcal mol^{-1}. The total strain energy on binding is therefore 5.2 kcal mol^{-1}, reflecting the conformational constraints imposed by the enzyme on the ligand. We would expect the strain energy to be stored in the low-energy distortion modes of a molecule. Table IV shows that this is the case; most of the strain induced on binding is stored in the torsional and steric (VDW) degrees of freedom. For example, the conformational requirements of binding force C6 of the pyrimidine ring to within 3.1 Å of C2′ of the trimethoxyphenyl group, as compared with the isolated molecule where these groups are 3.3 Å apart. This results in an increase of repulsive interaction in the bound ligand of 1.5 kcal mol^{-1}.

Changes in Ligand Entropy on Binding

An assumption often made in theoretical calculations of the free energy of binding is that the entropic contributions to binding are approximately equal for the binding of similar inhibitors to similar enzymes.[42] Studies on small peptides have indicated that vibrational entropies are of crucial importance in conformational equilibria.[26] The contribution of the entropy to the energy of binding may be of similar importance. A comparison of the energies and entropies of the free and bound ligand is given in Table IV. These contributions have been calculated from the Einstein equations, which express the vibrational partition function in terms of the vibrational frequencies. More recently, the Kirkwood integration method has been used to attain relative free energies.[43,44] The total change in free energy amounts to ≈21 kcal mol^{-1}, which is substantially more than the strain energy induced in the ligand by binding. This change in entropy is due to the restriction in ligand motion by the protein, which is reflected in the low-frequency modes that are shifted to higher frequencies in the bound ligand then in the free ligand. The normal modes corresponding to the lowest 14 frequencies show components of translational and rotational motion for the bound ligand.

CONCLUSIONS

This is the first protein simulation that uses the complete force field developed for small molecules,

which includes cross terms and all hydrogen atoms. The convergence criterion used ensured that the complex would undergo no further conformational changes with further minimization, i.e., a local minimum close to the experimental structure had been reached.

Our comparison of these simulations with the observed structures has shown that the calculated structure with all hydrogens included explicitly fits well, especially in the active site. Large deviations between the experimental and this minimized structure occur in surface regions of the protein. The united-atom model, on the other hand, shows a generally worse fit to the experimental structure, with the area around the active site having significantly larger deviations than the all-atom simulation has. Thus, it may be necessary to include all hydrogen atoms to understand the intermolecular forces and structural consequences of ligand binding.

The analysis of the energetics of the all-atom minimized structure shows that the total strain that the enzyme induces on the ligand trimethoprim is about 5 kcal mol^{-1}. Interactions of the pyrimidine ring of the ligand with neighboring residues of the protein are mainly electrostatic. In contrast, the predominant contribution to the interaction energy of neighboring residues of the protein with the trimethoxyphenyl moiety is from van der Waals interactions. Solvent molecules such as water 33 have been found to be conserved in all DHFR crystals, and their inclusion in the simulation is essential. The magnitude of the energy of interaction of nearby waters that hydrogen bond to the ligand is similar to the size of interactions of nearby protein residues with the ligand. Even some waters that are not involved in direct interactions have sizable interactions with the ligand. Thus, a shell of waters, rather than just a single layer, may be very important for accurate simulations of the active site of enzyme complexes. Finally, we have examined the entropy of both the bound and isolated ligand, the first step in the investigation of the importance of entropy in ligand binding for different ligands or for the same ligand in DHFRs from various species.

ACKNOWLEDGMENTS

This work was supported by Grant GM 30564 from the National Institutes of Health. D.J.O. was the recipient of a Control Data Corporation Pacer Postdoctoral Fellowship, Number 84PCR12A/B. Grants of Supercomputer time were provided by the National Science Foundation (No. DMB 8421273) at Purdue University and by Control Data Corporation at Colorado State University. The INSIGHT simulation software package (Biosym Technologies, Inc., 10065 Barnes Canyon Road, Suite A, San Diego, CA 92121) was used for part of the analysis described here.

REFERENCES

1. Struthers, R.S., Rivier, J., Hagler, A.T., Design of peptide analogs: Theoretical simulation of conformation, energetics, and dynamics. In "Conformationally Directed Drug Design: Peptides and Nucleic Acids As Templates or Targets." Vida, J.A., Gordon, M., eds. Washington, DC: American Chemical Society. 1984:239–261.
2. Hagler, A.T., Osguthorpe, D.J., Dauber-Osguthorpe, P., Hempel, J.C., Dynamics and conformational energetics of a peptide hormone: Vasopressin. Science 227:1309–1315, 1985.
3. Blakley, R.L., "The Biochemistry of Folic Acid and Related Pteridines." Amsterdam: North Holland, 1969.
4. Friedkin, M., Thymidylate synthetase. Adv. Enzymol. 38:235–293, 1973.
5. Gready, J.E., Dihydrofolate reductase: Binding of substrates and inhibitors and catalytic mechanism. Adv. Pharmacol. Chemother. 17:37–102, 1980.
6. Burchall, J., Hitchings, G.H., Inhibitor binding analysis of dihydrofolate reductases from various species. Mol. Pharmacol. 1:126–136, 1965.
7. Baker, B.R., Metabolite antagonism by enzyme inhibition. "Medicinal Chemistry", 3rd ed. Burger, A., ed. New York: Wiley-Interscience, 1970:196–228.
8. Huennekens, F.M., Vitols, K.S., Whitely, J.M., Neef, V.G., Dihydrofolate reductase. Methods Cancer Res. 13:199–225, 1976.
9. McCormack, J.J., Dihydrofolate reductase inhibitors as potential drugs. Med. Res. Rev. 1:303–331, 1980.
10. Roth, B., Cheng, C.C., Recent progress in the medicinal chemistry of 2,4-diaminopyrimidines. Prog. Med. Chem. 19:269–331, 1982.
11. Matthews, D.A., Bolin, J.T., Burridge, J.M., Filman, D.J., Volz, K.W., Kaufman, B.T., Beddell, C.R., Champness, J.N., Stammers, D.K., Kraut, J., Refined crystal structures of *Escherichia coli* and chicken liver dihydrofolate reducatase containing bound trimethoprim. J. Biol. Chem. 260:381–391, 1985.
12. Matthews, D.A., Alden, R., Bolin, J.T., Freer, S.T., Hamlin, R., Xuong, N., Kraut, J., Poe, M., Williams, M., Hoogstein, K., Dihydrofolate reductase: X-ray structure of the binary complex with methotrexate. Science 197:452–455, 1977.
13. Matthews, D.A., Alden, R.A., Bolin, J.T., Filman, D.J., Freer, S.T., Hamlin, R., Hol, W.G.J., Kisliuk, R.L., Pastore, E.J., Plante, L.T., Xuong, N., Kraut, J., Dihydrofolate reductase from *Lactobacillus casei:* X-ray structure of the enzyme methotrexate NADPH complex. J. Biol. Chem. 253:6946–6954, 1978.
14. Matthews, D.A., Alden, R.A., Freer, S.T., Xuong, N., Kraut, J., Dihydrofolate reductase from *Lactobacillus casei:* Stereochemistry of NADPH binding. J. Biol. Chem. 254:4144–4151, 1979.
15. Volz, K.W., Matthews, D.A., Alden, R.A., Freer, S.T., Hansch, C., Kaufman, B.T., Kraut, J., Crystal structure of avian dihydrofolate reductase containing phenyltriazine and NADPH. J. Biol. Chem. 257:2528–2536, 1982.
16. Bolin, J.T., Filman, D.J., Matthews, D.A., Hamlin, R.C., Kraut, J., Crystal structure of *E. coli* and *L. casei* dihydrofolate reductase refined at 1.7 Å resolution. I. General features and binding of methotrexate. J. Biol. Chem. 257:13650–13662, 1982.
17. Filman, D.J., Bolin, J.T., Matthews, D.A., Kraut, J., Crystal structure of *E. coli* and *L. casei* dihydrofolate reductase refined at 1.7 Å resolution. II. Environment of bound NADPH and implication for catalysis. J. Biol. Chem. 257:13663–13672, 1982.
18. Hagler, A.T., Theoretical simulation of conformation, energetics, and dynamics of peptides in conformation in biology & drug design. "The Peptides". Vol. 7. Hruby, V.J., Meienhofer, J., eds. 1985:213–299.
19. Hagler, A.T., 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. 96:5319–5327, 1974.
20. Hagler, A.T., Lifson, S., Energy functions for peptides and proteins. II. The amide hydrogen bond and calculation of amide crystal properties. J. Am. Chem. Soc. 96:5327–5335, 1974.
21. Lifson, S., Hagler, A.T., Dauber, P., Consistent force field studies of intermolecular forces in hydrogen bonded crys-

tals. 1. Carboxylic acids, amides, and the C=O· · ·H-O hydrogen bonds. J. Am. Chem. Soc. 101:5111–5121 (1979).
22. Hagler, A.T., Lifson, S., Dauber, P., Consistent force field studies of intermolecular forces in hydrogen bonded crystals. 2. A benchmark for the objective comparison of alternative force fields. J. Am. Chem. Soc. 101:5122–5130, 1979.
23. Hagler, A.T., Dauber, P., Lifson, S., Consistent force field studies of intermolecular forces in hydrogen bonded crystals. 3. The C=O· · ·H-O hydrogen bond and the analysis of the energetics and packing of carboxylic acids. J. Am. Chem. Soc. 101:5131–5141, 1979.
24. Dauber-Osguthorpe, P., Osguthorpe, D.J., Wolff, J., Hagler, A.T., Derivation of a valence force field for peptides and proteins. in preparation.
25. Hill, T.L., An Introduction to Statistical Thermodynamics. Reading, MA: Addison-Wesley, 1960.
26. Hagler, A.T., Stern, P.S., Sharon, R., Becker, J.M., Naider, F., Computer simulation of the conformational properties of oligopeptides. Comparison of theoretical methods and analysis of experimental results. J. Am. Chem. Soc. 101:6842–6852, 1979.
27. Roberts, V.A., Dauber-Osguthorpe, P., Osguthorpe, D.J., Levin, E., Hagler, A.T., A comparison of the binding of the ligand trimethoprim to bacterial and vertebrate dihydrofolate reductases. Isr. J. Chem. 27:198–210, 1986.
28. Brooks, C.L., III, Brunger, A., Karplus, M., Active site dynamics in protein molecules: A stochastic boundary. Biopolymers 24:843–865, 1985.
29. Wipff, G., Dearing, A., Weiner, P.K., Blaney, J.M., Kollman, P.A., Interactions: The interaction of L- and D-N-Acetyltryptophanamide with alpha-chymotrypsin. J. Am. Chem. Soc. 105:997–1005, 1983.
30. Fehlhammer, H., Bode, W., Huber, R., Crystal structure of bovine trypsinogen at 1.8 Å resolution, II. Crystallographic refinement, refined crystal structure and comparison with bovine trypsin. J. Mol. Biol. 111:415–438, 1977.
31. Kossiakoff, A.A., Chambers, J.L., Kay, L.M., Stroud, R.M., Structure of bovine trypsinogen at 1.9 Å resolution. Biochemistry 16:654–664, 1977.
32. Takano, T., Structure of myoglobin refined at 2.0 Å resolution, II. Structure of deoxymyoglobin from sperm whale. J. Mol. Biol. 110:569–584, 1977.
33. Phillips, S.E.V., Structure and refinement of oxymyoglobin at 1.6 Å resolution. J. Mol. Biol. 142:531–554, 1980.
34. Wlodawer, A., Bott, R., Sjolin, L., The refined crystal structure of ribonuclease A at 2.0 Å resolution. J. Biol Chem. 257:1325–1332, 1982.
35. Borkakoti, N., Moss, D.S., Palmer, R.A., Ribonuclease-A: Least squares refinement of the structure at 1.45 Å resolution. Acta Crystallogr. B38:2210–2217, 1982.
36. Marquart, M., Walter, J., Deisenhofer, J., Bode, W., Huber, R., The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. Acta Crystallogr. B39:480–490, 1983.
37. Walter, J., Steigemann, W., Singh, T.P., Bartunik, H., Bode, W., Huber, R., On the disordered activation domain in trypsinogen: Chemical labelling and low-temperature crystallography. Acta Crystallogr. B38:1462–1472, 1982.
38. Wang, D., Bode, W., Huber, R., Bovine chymotrypsinogen A, X-ray crystal structure analysis and refinement of a new crystal form at 1.8 Å resolution. J. Mol. Biol. 185:595–624, 1985.
39. Schlegel, H.B., Poe, M., Hoogsteen, K., Models for the binding of methotrexate to *Escherichia coli* dihydrofolate reductase. Direct effect of carboxylate of ASP 27 Upon UV spectrum of methotrexate. Mol Pharmacol. 20:154–158, 1981.
40. Koetzle, T.F., Williams, G.J.B., The crystal and molecular structure of the antifolate drug trimethoprim, a neutron diffraction study. J. Am. Chem. Soc. 98:2074–2078, 1976.
41. Dauber-Osguthorpe, P., Roberts, V.A., Osguthorpe, D.J., Wolff, J., Hagler, A.T., Unpublished results.
42. Blaney, J.M., Weiner, P.K., Dearing, A., Kollman, P.A., Jorgensen, E.C., Oatley, S.J., Burridge, J.M., Blake, C.C.F., Molecular mechanics simulation of protein-ligand interactions: Binding of thyroid hormone analogues to prealbumin. J. Am. Chem. Soc. 104:6424–6434, 1982.
43. Wong, C.F., McCammon, J.A., Dynamics and design of enzymes and inhibitors. J. Am. Chem. Soc. 108:3830–3832, 1986.
44. Bash, P.A., Singh, U.C., Brown, F.K., Langridge, R., Kollman, P.A., Calculation of the relative change in binding free energy of a protein inhibitor complex. Science 235:574–575, 1987.

APPENDIX

Potential Energy Surface

The potential energy, V, is expressed as a function of the internal degrees of freedom and the nonbonded distances as given in equation 1:

$$
\begin{aligned}
V = {} & \sum\left\{D_b\left[1 - e^{-a(b-b_0)}\right]^2 - D_b\right\} + \frac{1}{2}\sum H_0(\theta - \theta_0)^2 \\
& + \frac{1}{2}\sum H_\phi(1 + s\cos n\phi) + \frac{1}{2}\sum H_\chi \chi^2 \\
& + \sum\sum F_{bb'}(b - b_0)(b' - b_0') \\
& + \sum\sum F_{\theta\theta'}(\theta - \theta_0)(\theta' - \theta_0') + \sum\sum F_{b\theta}(b - b_0)(\theta - \theta_0) \\
& + \sum F_{\phi\theta\theta'}\cos\phi(\theta - \theta_0)(\theta' - \theta_0') + \sum\sum F_{\chi\chi'}\chi\chi' \\
& + \sum\frac{A}{r^{12}} - \frac{B}{r^6} + \sum\frac{q_i q_j}{r}
\end{aligned}
\tag{1}
$$

Basically, the terms in equation 1 represent the energies required to deform the internal coordinates (which have the values b, 0, ϕ, and χ) from their unstrained standard values, denoted by the subscript "0", or the energies of nonbonded interactions. The terms given represent the energies for (in the order used in equation 1) bond stretching or compression; angle bending; torsion angle twisting; out-of-plane deformation of a planar system; coupling between bond deformations; coupling between angle bendings; coupling between a bond deformation and an angle bending; coupling between a torsion angle twisting and the two associated angle bendings; coupling between out-of-plane deformations; repulsive and dispersive interactions between nonbonded atoms; and coulombic interactions between nonbonded atoms. The parameters D_b, H_0, H_ϕ, H_χ, F_{ij} are the force constants for the corresponding intramolecular deformations, A and B are the parameters for the nonbond repulsive and dispersive interactions, and q_i are the partial charges carried by each atom.

TABLE I. Potential Function Parameters: Bond Parameters*

Atom types	cb(1)	cb(2)	cb(3)
h c	108.60	1.105	1.771
c c	88.00	1.526	1.915
c c'	76.00	1.520	1.930
o' c'	145.00	1.230	2.060
n hn	93.00	1.026	2.280
n c'	97.00	1.320	2.000
n c	72.00	1.460	2.290
ho oh	104.00	0.960	2.280
c oh	96.00	1.420	2.000
c o	96.00	1.420	2.000
ho o	95.00	0.960	2.280
h c'	108.60	1.105	1.771
c' o	100.00	1.370	2.000
c' o-	135.00	1.250	2.000
h cp	116.00	1.080	1.770
cp cp	120.00	1.340	2.000
c cp	76.00	1.510	1.930
cp oh	96.00	1.370	2.000
hs sh	87.50	1.330	1.770
c sh	57.00	1.820	2.000
c s	57.00	1.800	2.000
s s	45.00	2.000	2.000
hn n3	88.00	1.026	2.280
c n3	68.00	1.470	2.290
cp np	140.00	1.260	2.000
hn np	93.00	1.026	2.280
cp c5	70.00	1.340	2.000
h c5	116.00	1.080	1.770
c5 c5	70.00	1.390	2.000
c c5	76.00	1.510	1.930
np c5	80.00	1.380	2.000
n cp	70.00	1.340	2.000
o* h*	104.00	0.960	2.280
p oh	75.00	1.570	2.000
o' p	140.00	1.530	2.000
o- p	120.00	1.530	2.000
s o-	120.00	1.530	2.000
h p	56.00	1.500	2.000
c np	84.20	1.475	2.000
np c=	140.00	1.260	2.000
c' c=	80.70	1.500	2.000
c c=	80.70	1.500	2.000
c= c=	163.80	1.330	2.000
h c	90.40	1.090	2.000
c oe	83.00	1.470	2.000
c' oe	100.00	1.370	2.000

b Is the bond length in the trial structure. f(b)=cb(1)(1-exp(-cb(3)*(b-cb(2))))**2

TABLE II. Potential Function Parameters: Angle-Bending Parameters*

Atom types			ct(1)	ct(2)	ct(3)	ct(4)	ct(5)
h	c	h	39.50	106.40	0.00	0.00	0.00
h	c	c	44.40	110.00	0.00	12.60	38.40
c	c	c	46.60	110.50	28.50	60.20	60.20
o'	c'	c	68.00	120.00	25.00	2.00	30.00
h	c	c'	45.00	109.50	2.00	2.00	38.40
c	c	c'	46.60	110.50	28.50	60.20	60.20
c'	c	c'	46.60	110.50	28.50	60.20	60.20
c'	n	hn	37.50	115.00	2.00	23.30	2.00
c	n	c'	111.00	118.00	25.30	31.50	30.00
c	n	hn	35.00	122.00	2.00	23.30	2.00
n	c	c	50.00	109.50	25.00	35.00	35.00
h	c	n	51.50	109.50	2.00	2.00	40.00
n	c	c'	50.00	109.50	25.00	35.00	35.00
n	c'	c	53.50	114.10	45.90	2.00	31.50
o'	c'	n	68.00	120.00	10.00	2.00	2.00
c	o	ho	58.50	106.00	0.00	57.00	0.00
h	c	o	57.00	109.50	0.00	0.00	64.40
c	c	o	70.00	109.50	25.00	0.00	0.00
hn	n	hn	33.00	125.00	10.00	2.00	2.00
o'	c'	o	145.00	123.00	0.00	0.00	0.00
c	c'	o	122.80	110.00	0.00	0.00	0.00
c'	o	ho	50.00	112.00	0.00	0.00	0.00
h	c'	o	55.00	110.00	0.00	0.00	0.00
h	c'	o'	55.00	120.00	25.00	2.00	30.00
o−	c'	o−	145.00	123.00	50.00	0.00	0.00
c	c'	o−	68.00	120.00	50.00	0.00	30.00
h	c'	o−	55.00	120.00	50.00	0.00	30.00
c	n	c	37.00	120.00	50.00	50.00	50.00
h	c'	n	45.00	120.00	0.00	0.00	0.00
h	cp	cp	37.00	120.00	−16.00	0.00	50.00
cp	cp	cp	90.00	120.00	45.00	−50.00	−50.00
h	c	cp	44.40	110.00	0.00	12.60	38.40
c	cp	cp	44.20	120.00	0.00	−50.00	−50.00
c	c	cp	46.60	110.50	28.50	60.20	60.20
ho	o	cp	50.00	109.00	0.00	0.00	0.00
o	cp	cp	60.00	120.00	50.00	0.00	0.00
c	sh	hs	48.00	96.00	0.00	0.00	0.00
h	c	sh	45.00	109.00	0.00	0.00	0.00
h	c	s	45.00	109.00	0.00	0.00	0.00
c	c	sh	62.00	109.00	2.00	50.00	50.00
c	c	s	62.00	109.00	2.00	50.00	50.00
c	s	c	58.00	99.00	0.00	0.00	0.00
c	s	s	75.00	109.50	0.00	0.00	0.00
h	c	n3	57.30	109.50	0.00	0.00	0.00
hn	n3	hn	36.00	105.50	0.00	0.00	0.00
c	n3	c	86.30	112.00	0.00	0.00	0.00
c	n3	hn	41.60	110.00	0.00	0.00	0.00
c	c	n3	50.00	109.50	0.00	0.00	0.00
np	cp	np	102.00	134.00	0.00	0.00	0.00
cp	cp	np	90.00	120.00	0.00	0.00	0.00
h	cp	np	50.00	120.00	0.00	0.00	50.00
cp	np	cp	75.10	114.00	0.00	0.00	0.00
hn	np	cp	27.50	120.00	0.00	0.00	0.00
h	cp	c5	37.00	120.00	−16.00	0.00	50.00
cp	cp	c5	90.00	120.00	45.00	−50.00	−50.00
cp	c5	c5	90.00	120.00	0.00	0.00	0.00
h	c5	c5	37.00	120.00	−16.00	0.00	50.00
c5	c5	c5	90.00	120.00	0.00	0.00	0.00
h	c	c5	44.40	110.00	0.00	12.60	38.40
c	c5	c5	44.20	120.00	0.00	−50.00	−50.00
c	c	c5	46.60	110.50	28.50	60.20	60.20
np	c5	np	102.00	134.00	0.00	0.00	0.00
np	c5	c5	90.00	120.00	0.00	0.00	0.00
cp	c5	np	90.00	120.00	0.00	0.00	0.00
h	c5	np	40.00	120.00	0.00	0.00	50.00
c5	np	c5	75.10	114.00	0.00	0.00	0.00
hn	np	c5	27.50	120.00	0.00	0.00	0.00
c	c5	np	50.00	120.00	0.00	0.00	0.00
n	c'	n	102.00	120.00	0.00	0.00	0.00
c	o	cp	50.00	109.50	0.00	0.00	0.00
n	cp	np	102.00	120.00	0.00	0.00	0.00
n	cp	cp	102.00	120.00	0.00	0.00	0.00
hn	n	cp	37.50	115.00	2.00	2.00	23.30
c'	c	n3	50.00	109.50	25.00	35.00	35.00
cp	c	cp	46.60	110.50	28.50	60.20	60.20
c	o	c	60.00	109.50	0.00	57.00	57.00
c	o	c'	60.00	109.50	0.00	57.00	57.00
h*	o*	h*	50.00	104.50	−14.50	31.30	31.30
c	o	p	72.00	120.00	0.00	0.00	0.00
p	o	p	72.00	120.00	0.00	0.00	0.00
o	p	o	110.00	109.50	80.00	0.00	0.00
o'	p	o	110.00	109.50	40.00	0.00	0.00
o'	p	o'	110.00	109.50	40.00	0.00	0.00
o−	p	o−	110.00	109.50	40.00	0.00	0.00
o	p	o−	110.00	109.50	40.00	0.00	0.00
h	p	o−	80.00	109.50	0.00	0.00	60.00
o−	s	o−	110.00	109.50	40.00	0.00	0.00
c	np	c5	70.00	128.80	0.00	0.00	0.00
o	c	np	80.00	109.50	0.00	0.00	0.00
h	c	np	80.00	109.50	0.00	0.00	0.00
c	c	np	80.00	109.50	0.00	0.00	0.00
n	c'	c=	53.50	114.10	45.90	2.00	31.50
c'	c=	c=	36.20	122.30	28.50	60.20	60.20
c	c=	c=	36.20	122.30	28.50	60.20	60.20
h	c=	c	37.50	120.00	0.00	38.40	0.00
c=	c	c=	46.60	110.50	28.50	60.20	60.20
h	c	c=	44.40	110.00	0.00	12.60	38.40
h	c=	c=	33.80	121.20	0.00	0.00	38.40
np	c=	c=	90.00	120.00	0.00	0.00	0.00
c=	np	c=	75.10	114.00	0.00	0.00	0.00
o'	c'	c=	50.00	120.00	0.00	0.00	0.00
c	c=	c'	50.00	120.00	0.00	0.00	0.00
h	c=	np	40.00	120.00	0.00	0.00	50.00
c	np	c=	50.00	120.00	0.00	0.00	0.00
cp	np	c5	75.10	114.00	0.00	0.00	0.00
np	cp	c5	90.00	120.00	0.00	0.00	0.00
n	cp	c5	102.00	120.00	0.00	0.00	0.00

f(th) = ct(1)*(th-ct(2))**2, angle bending energy;
+ f(b1,b2) = ct(3)*(b1 - cb1)*(b2-cb2), bond-bond cross term energy;
+ f(th,b) = ct(4)*(b1-cb1)*(th-ct(2))+ct(5)*(b2-cb2)*(th-ct(2)), bond-angle cross term energy.
*th, the valence angle in the trial structure; b1 and b2, the bond lengths, in the trial structure, for the two bonds involved in the angle; cb1 and cb2, the "standard" bond lengths for these two bonds, obtained from Table I.

TABLE III. Potential Function Parameters: Torsional Parameters†

Atom types				cp(1)	cp(2)	cp(3)	cp(4)
*	c	c	*	1.4225	3	1	−10.5
*	c	c′	*	0.0000	0	0	0
*	n	c	*	0.0000	0	0	0
*	c	o	*	0.3900	3	1	0
*	c	of	*	0.3900	3	1	0
*	c′	n2	*	6.0000	2	−1	0
*	n2	cr	*	6.8000	2	−1	0
*	c′	o	*	4.5000	2	−1	0
*	cp	cp	*	12.0000	2	−1	−8.5
*	c	cp	*	0.0000	0	0	0
*	o	cp	*	1.5000	2	−1	0
*	cp	of	*	3.5000	2	−1	0
*	c	sh	*	0.4200	3	1	−10.5
*	c	s	*	0.4200	3	1	−10.5
*	s	s	*	5.5000	2	1	0
*	c	n3	*	0.8000	3	1	−10.5
*	cp	np	*	4.0000	2	−1	0
*	n2	cp	*	10.0000	2	−1	0
*	o	p	*	0.7500	3	1	0
*	np	c=	*	4.0000	2	−1	0
*	c=	c=	*	16.3000	2	−1	0
*	c	c=	*	1.2660	3	−1	0
*	c	np	*	0.0000	0	0	0
*	c′	c=	*	1.8000	2	1	0
c	n	c′	c	3.2000	2	−1	−8
hn	n	c′	c	1.2000	2	−1	−12
c	n	c′	o′	3.8000	2	−1	−8
hn	n	c′	o′	1.8000	2	−1	−12
hn	n	c′	h	1.2000	2	−1	−12
c	n	c′	h	3.2000	2	−1	−8
hn	n	cr	n2	3.4000	2	−1	0
c	n	cr	n2	3.4000	2	−1	0

†f(ph)=cp(1)*(1+cp(3)*cos(cp(2)*ph)), torsional energy; +f(th1,th2,ph)=cp(4)*(th1−ct1)*(th2−ct2)*cos(ph), angle-angle-torsion cross term energy. pH, the torsion angle in the trial structure.

A "*" for an atom type indicates that this can be any atom type; th1 and th2, the valence angles, in the trial structure, for the two angles involved in the torsion angle; ct1 and ct2, the "standard" valence angles for these two bonds, obtained from Table II.

TABLE IV. Potential Function Parameters: Out-of-Plane Parameters*

Atom types				cp(1)	cp(2)
c	c′	n	o′	10.000000	0.010000
c′	n	c	hn	0.050000	0.010000
c	c′	n2	o′	10.000000	0.010000
h	c′	n2	o′	10.000000	0.010000
hn	n2	c′	hn	0.050000	0.010000
c	c′	o′	o′	11.600000	0.000000
h	c′	o′	o′	11.600000	0.000000
c′	n	c	c	0.050000	0.000000
h	c′	o′	n	0.050000	0.000000
cp	cp	cp	h	0.370000	0.000000
cp	cp	cp	c	0.370000	0.000000
cp	cp	cp	cp	0.370000	0.000000
cp	cp	cp	o′	0.000000	0.000000
cp	cp	h	np	0.370000	0.000000
h	cp	np	np	0.370000	0.000000
cp	cp	cp	np	0.370000	0.000000
n2	cp	np	np	0.370000	0.000000
cp	n2	hn	hn	0.000000	0.000000
n	c′	n2	n2	10.000000	0.010000
o′	c′	n2	n2	10.000000	0.010000
hn	np	cp	cp	0.370000	0.000000
c	cp	cp	np	0.370000	0.000000
n2	cp	cp	np	0.370000	0.000000
c=	c′	n2	o′	10.000000	0.010000
c	c=	c′	c=	11.100000	0.000000
h	c=	np	c=	11.100000	0.000000
h	c=	c	c=	11.100000	0.000000
c	np	cp	cp	0.370000	0.000000
cp	cp	np	np	0.370000	0.000000

*f(oph) = cp(1)*oph**2, out of plane energy; +f(oph1,oph2)=cp(2)*oph1*oph2, out of plane cross term energy. oph, the improper torsion angle in the trial structure (defined in such a way that a planar system has a value of 0); oph1 and oph2, two adjacent out of plane interactions. The central angle in the out of plane interaction is the second of the atoms in the list.

TABLE V. Potential Function Parameters: Angle-Angle Interaction Parameters*

Atom types			ctf(1)
c	(c	c	)c −7.900000
h	(c	c	)h −7.900000
h	(c	c	)c −7.900000
h	(c	c′	)h −7.500000
c′	(n	c	)hn −7.500000
h	(c	n	)c −7.900000
n	(c	h	)c −7.900000
h	(c	c	)n −7.900000
h	(c	c′	)c −7.900000
h	(c	c	)c′ −7.900000
n	(c	h	)c′ −7.900000
h	(c	c′	)n −7.900000
h	(c	n	)c′ −7.900000
n	(c	c	)c′ −7.900000
n	(c	c′	)c −7.900000
c	(c	n	)c′ −7.900000
o′	(c′	n	)c −7.500000
h	(c	n	)h −7.500000
c	(n	c	)c′ 10.000000
c	(n	c′	)c 10.000000
cp	(cp	h	)cp 14.000000
h	(cp	cp	)cp 10.000000
h	(c	cp	)h −7.900000
h	(c	c	)cp −7.900000
h	(c	cp	)c −7.900000
h	(c	s	)h −10.000000
h	(c	sh	)h −10.000000
np	(cp	n	)np −8.000000
hn	(n	cp	)hn −8.000000
c	(c	n3	)c′ −7.900000
c	(c	c′	)n3 7.900000
c′	(c	c	)n3 −7.900000
h	(c	n3	)c′ −7.900000
h	(c	c′	)n3 −7.900000
c′	(c	h	)n3 −7.900000
h	(p	o−	)o− 30.000000
o−	(p	h	)o− 30.000000
c	(c	n	)c −7.900000
n	(c	c	)c −7.900000
c	(c	c′	)c −7.900000
c	(c	c	)c′ −7.900000

f(t1,t2)−ctf(1)(t1−ct1,t1)(t2−ct2,t2)) t1 and t2, two angles in the trial structure; ct1,t1) and ct(2,t2), the "standard" values for the angles, obtained from Table II.

The second atom in the list is the common vertex of the two angles, and the bond between the second and the third atoms is the bond that is common to the two angles.

The parameter for any interaction which is not included in this list is 0.0.

TABLE VI. Potential Function Parameters: Nonbond Parameters[a]

Atom types		A	C_6
h	h	7108	32
cg	cg	1790340	528
o′	o′	272894	498
n	n	2266872	1230
c′	c′	2968753	1325
c	c	1981049	1125
hn	hn	0.000000	0.000000
s	s	365906	250
o*	o*	629358	625
h*	h*	0.002500	0.005000
p	p	10530818	2902
na	na	30746	74

[a] $v_{NB} = A_{ij}/r_{ij}^{12} - C_{6ij}/r_{ij}^{6}$; $A_{ij} = (A_{ii}A_{jj})^{1/2}$ $C_{6ij} = (C_{6ii}C_{6jj})^{1/2}$ [e.g. $A_{CH} = (A_{CC}A_{HH})^{1/2}$]

TABLE VII. Potential Function Parameters: Charges*

Atom type (1)	Charge
Water	
o*	−.82
h*	.41
h*	.41
Pyrimidine moiety of trimethoprim	
h	.15
cp	.15
np	−.50
hn	.36
cp	.48
n2	−.50
hn	.36
hn	.36
np	−.50
cp	.36
n2	−.50
hn	.36
hn	.36
cp	.00
CH2 moiety of trimethoprim	
c2	−.20
h	.13
h	.13
Trimethoxyphenyl moiety of trimethoprim	
cp	.00
cp	−.10
h	.10
cp	.15
of	−.30
c3	−.15
h	.10
h	.10
h	.10
cp	.15
of	−.30
c3	−.15
h	.10
h	.10
h	.10
cp	−.10
h	.10
Gly	
n	−.50
hn	.28
cg	.02
h	.10
h	.10
c′	.38
o′	−.38
Ala	
n	−.50
hn	.28
ca	.12
h	.10
c3	−.30
h	.10
h	.10
h	.10
c′	.38
o′	−.38
Val	
n	−.50
hn	.28
ca	.12
h	.10
c1	−.10
h	.10
c3	−.30
h	.10
h	.10
h	.10
c3	−.30
h	.10
h	.10
h	.10
c′	.38
o′	−.38
Leu	
n	−.50
hn	.28
ca	.12
h	.10
c2	−.20
h	.10
h	.10
c1	−.10
h	.10
c3	−.30
h	.10
h	.10
h	.10
c3	−.30
h	.10
h	.10
h	.10
c′	.38
o′	−.38

TABLE VII. Potential Function Parameters: Charges* (Continued)

Atom type (1)	Charge
Ile	
n	−.50
hn	.28
ca	.12
h	.10
c1	−.10
h	.10
c3	−.30
h	.10
h	.10
h	.10
h	.10
c2	−.20
h	.10
h	.10
c3	−.30
h	.10
h	.10
h	.10
c′	.38
o′	−.38
Phe	
n	−.50
hn	.28
ca	.12
h	.10
c2	−.20
h	.10
h	.10
cp	.00
cp	−.10
h	.10
cp	−.10
h	.10
cp	−.10
h	.10
cp	−.10
h	.10
c′	.38
o′	−.38
Tyr	
n	−.50
hn	.28
ca	.12
h	.10
c2	−.20
h	.10
h	.10
cp	.00
cp	−.10
h	.10
cp	−.10
h	.10
cp	−.10
h	.10
c′	.38
o′	−.38
Met	
n	−.50
hn	.28
ca	.12
h	.10
c2	−.20
h	.10
h	.10
c2	−.30
h	.10
h	.10
s	.12
c3	−.32
h	.10
h	.10
h	.10
c′	.38
o′	−.38
Met-N terminal-positive charge	
n3	−.50
hn	.36
hn	.36
hn	.36
ca	.12
h	.17
c2	−.20
h	.13
h	.13
c2	−.30
h	.12
h	.12
s	.12
c3	−.32
h	.10
h	.10
h	.10
c′	.38
o′	−.35
Ser	
n	−.50
hn	.28
ca	.12

TABLE VII. Potential Function Parameters: Charges* (Continued)

Atom type (1)	Charge
h	.10
c2	−.17
h	.10
h	.10
oh	−.38
ho	.35
c′	.38
o′	−.38
Thr	
n	−.50
hn	.28
ca	.12
h	.10
c1	−.07
h	.10
c3	−.30
h	.10
h	.10
h	.10
oh	−.38
ho	.35
c′	.38
o′	−.38
Asn	
n	−.50
hn	.28
ca	.12
h	.10
c2	−.20
h	.10
h	.10
c′	.38
o′	−.38
n2	−.56
hn	28
hn	.28
c′	.38
o′	−.38
Gln	
n	−.50
hn	.28
ca	.12
h	.10
c2	−.20
h	.10
h	.10
c2	−.20
h	.10
h	.10
c′	.38
o′	−.38
n2	−.56
hn	.28
hn	.28
c′	.38
o′	−.38
Cys-reduced	
n	−.50
hn	.28
ca	.12
h	.10
c2	−.30
h	.10
h	.10
sh	.00
hs	.10
c′	.38
o′	−.38
Glu-negative charge	
n	−.50
hn	.26
ca	.12
h	.07
c2	−.22
h	.09
h	.09
c2	−.38
h	.15
h	.15
c−	.34
o−	−.57
o−	−.57
c′	.38
o′	−.41
Trp	
n	−.50
hn	.28
ca	.12
h	.10
c2	−.20
h	.10
h	.10
c5	.00
c5	.01
h	.10
np	−.50
hn	.28
c5	.11
c5	.00
cp	−.10
h	.10
cp	−.10
h	.10
cp	−.10
h	.10
cp	−.10

TABLE VII. Potential Function Parameters: Charges* (Continued)

Atom type (1)	Charge
h	.10
c′	.38
o′	−.38
Pro	
n	−.42
ca	.06
h	.10
c2	−.20
h	.10
h	.10
c2	−.20
h	.10
h	.10
c2	.06
h	.10
h	.10
c′	.38
o′	−.38
Lys-positive charge	
n	−.50
hn	.28
ca	.12
h	.10
c2	−.20
h	.12
h	.12
c2	−.20
h	.12
h	.12
c2	−.20
h	.12
h	.12
c2	−.08
h	.19
h	.19
n3	−.50
hn	.36
hn	.36
hn	.36
c′	.38
o′	−.38
Arg-C terminal-negative charge, positive charge on sidechain	
n	−.50
hn	.28
ca	.08
h	.02
c2	−.20
h	.07
h	.07
c2	−.20
h	.13
h	.13
c2	−.09
h	.13
h	.13
n	−.50
hn	.36
cr	.45
n2	−.50
hn	.36
hn	.36
n2	−.50
hn	.36
hn	.36
c−	.34
o−	−.57
o−	−.57
Arg-positive charge	
n	−.50
hn	.28
ca	.12
h	.10
c2	−.20
h	.11
h	.11
c2	−.20
h	.13
h	.13
c2	−.09
h	.13
h	.13
n	−.50
hn	.36
cr	.45
n2	−.50
hn	.36
hn	.36
n2	−.50
hn	.36
hn	.36
c′	.38
o′	−.38
Asp-negative charge	
n	−.50
hn	.26
ca	.12
h	.07
c2	−.26
h	.07
h	.07
c−	.34
o−	−.57
o−	−.57
c′	.38
o′	−.41

(continued)

TABLE VII. Potential Function Parameters: Charges* (Continued)

Atom type (1)	Charge
His	
n	−.50
hn	.28
ca	.12
h	.10
c2	−.20
h	.10
h	.10
c5	.10
np	−.42
c5	.27
h	.13
np	−.50
hn	.28
c5	.01
h	.13
c'	.38
o'	−.38
His-positive charge	
n	−.50
hn	.30
ca	.12
h	.12
c2	−.20
h	.15
h	.15
c5	.16
np	−.50
hn	.36
c5	.39
h	.25
np	−.50
hn	.36
c5	.10
h	.22
c'	.38
o'	−.36

*f(q)=q(i)*q(j)/r. See listing of Residue Library.
r, interatomic distance in the trial structure.
1: This gives the atom type that is used to select the appropriate potential parameters for this atom.
2: This gives the No. of the atom within the residue to which this atom is connected, e.g., a 5 means that this atom is connected to atom No. 5 in this residue. A 0 means that this atom is connected to the last atom of the previous residue, or to no atom if this is the first residue in a molecule. A −1 (or −2) means that this atom is connected to the second (or third) last atom of the previous residue.
3: If this column is non-zero, then this gives the No. of an additional atom within the residue to which this atom is connected.
4: This indicates that this atom is at the center of an out-of-plane interaction.

TABLE VIII. Atom Type Equivalence Table*

Atom type	h	o'	n	c	cg	s	c'	d	hn	hc	o	n2	cp	cr	c−	o−	p	ca	c3	oh
Bond	h	o'	n	c	c	s	c'	h	hn	ho	o	n	cp	c'	c'	o−	p	c	c	oh
Valence	h	o'	n	c	c	s	c'	h	hn	ho	o	n	cp	c'	c'	o−	p	c	c	o
Torsion	h	o'	n	c	c	s	c'	h	hn	ho	o	n2	cp	cr	c'	o'	p	c	c	o
o-o-plane	h	o'	n	c	c	s	c'	h	hn	ho	o'	n2	cp	c'	c'	o'	p	c	c	o'
Non-bond	h	o'	n	c	cg	s	c'	h	hn	hn	o'	n	c'	c'	c'	o'	p	cg	cg	o'
Mass	1.01	16.00	14.01	12.01	12.01	32.06	12.01	2.01	1.01	1.01	16.00	14.01	12.01	12.01	12.01	16.00	30.97	12.01	12.01	16.00

Atom type	cn	s1	np	hp	c2	hs	c1	sh	n3	c5	o*	h*	of	c=	oe	c+	na	op
Bond	c	s	np	h	c	hs	c	sh	n3	c5	o*	h*	oh	c=	oe	c+	na	oh
Valence	c	s	np	h	c	hs	c	sh	n3	c5	o*	h*	o	c=	o	c+	na	o
Torsion	c	s	np	h	c	hs	c	sh	n3	cp	o*	h*	of	c=	o	c+	na	o
o-o-plane	c	s	np	h	c	hs	c	sh	n3	cp	o*	h*	o'	c=	o'	c+	na	o'
Non-bond	cg	s	n	h	cg	h	cg	s	n	c'	o*	h*	o'	c'	o'	c+	na	op
Mass	12.01	32.06	14.01	1.01	12.01	1.01	12.01	32.06	14.01	12.01	16.00	1.01	16.00	12.01	16.00	40.08	22.99	16.00

*The atom type equivalence table is used as a means of reducing the No. of distinct parameters that need to be specified in the potential file. This is done by assigning, for each atom type, a set of atom types, which the first atom type will be considered equivalent to for the purpose of finding potential parameters. For example, if we consider atom type ca, which is used for the alpha carbon atom in most residues, in order to find a bond stretching parameter for ca, one would look in the potential file for the corresponding bond involving atom type c (which is listed in the bond row for atom type ca); e.g., the bond stretching parameter for a ca-h bond would be found under c-h in the potential file. Similarly, the entries in the valence, torsion, o-o-plane, and non-bond rows give the equivalent atom types for the purpose of finding valence angle, torsion angle, out-of-plane and nonbond parameters, respectively. The mass row gives the mass for each atom type.