Implicit Solvent Models for Flexible Protein–Protein Docking by Molecular Dynamics Simulation

Ting Wang[1*] **and Rebecca C. Wade**[1,2]
[1]*European Media Laboratory, Heidelberg, Germany*
[2]*European Molecular Biology Laboratory, Heidelberg, Germany*

ABSTRACT **The suitability of three implicit solvent models for flexible protein–protein docking by procedures using molecular dynamics simulation is investigated. The three models are (i) the generalized Born (GB) model implemented in the program AMBER6.0; (ii) a distance-dependent dielectric (DDD) model; and (iii) a surface area-dependent model that we have parameterized and call the NPSA model. This is a distance-dependent dielectric model modified by neutralizing the ionizable side-chains and adding a surface area-dependent solvation term. These solvent models were first tested in molecular dynamics simulations at 300 K of the native structures of barnase, barstar, segment B1 of protein G, and three WW domains. These protein structures display a range of secondary structure contents and stabilities. Then, to investigate the performance of the implicit solvent models in protein docking, molecular dynamics simulations of barnase/barstar complexation, as well as PIN1 WW domain/peptide complexation, were conducted, starting from separated unbound structures. The simulations show that the NPSA model has significant advantages over the DDD and GB models in maintaining the native structures of the proteins and providing more accurate docked complexes. Proteins 2003;50:158–169.** © 2002 Wiley-Liss, Inc.

Key words: molecular dynamics simulation; protein stability; implicit solvent model; generalized Born model; protein docking

INTRODUCTION

Mutual recognition of proteins and their binding partners is a critical event in biological regulation and control processes. In many cases, protein-ligand/protein binding is accompanied by conformational changes of both ligand and protein. Therefore, conformational flexibility is one of the important factors to be taken into account in computer docking simulations. One approach to this is to perform rigid body docking followed by refinement of side-chain conformations by searching rotamers.[1] However, side-chain motions are generally coupled to backbone motions, and the latter can be significant.[2] A natural way to explore relevant backbone and side-chain flexibility is by molecular dynamics (MD) simulations.[3,4] But, because of the large size of proteins and the computationally intensive nature of MD simulation, docking will be dramatically more expensive when including explicit solvent molecules into the calculation. On the other hand, simulation in vacuum gives rise to artifacts. This is not surprising when one considers that interfacial water molecules are present in many structures of protein–protein complexes.[5,6] To achieve a balance between accuracy and efficiency, implicit models of the solvent[7] that are suitable for protein docking problems are required.

The first implicit solvent model used in MD simulations of proteins[8] is a distance-dependent dielectric model, in which the solvent electrostatic screening effect is simply expressed by a function $\epsilon(r) = nr$ (usually $n = 1$, 2, or 4). We refer to this as the DDD model. This dielectric dependence has no direct physical basis but is computationally efficient, reducing the time spent on computing square roots for interatomic distances.[8] It provides an approximate model of the dielectric screening of the solvent. However, the DDD model tends to drive rather close oppositely charged groups together, leading to formation of surface salt-bridges. For protein–protein docking, this leads to the problem that incorrect intramolecular salt-link contacts can be formed, and these prevent the correct intermolecular salt-link contacts being made. For example, previous MD simulations on barnase-barstar docking[2] showed that the loop residue Arg59 of barnase failed to form the key contact with Glu76 of barstar because of the formation of the incorrect intramolecular salt link between Arg59 and Glu60 of barnase (see Fig. 1).

A physically rigorous implicit solvent model is the continuum electrostatic model obtained by solving the Poisson–Boltzmann equation. However, computing Poisson–Boltzmann forces during simulations of protein dynamics is very computationally demanding.[9] For flexible protein docking applications, we are interested in having an implicit solvent model that is more accurate than the DDD model but requires no significant increase of computational expense. Consequently, we tested two models: the generalized Born (GB) model[10,11] and a modified DDD

Grant sponsor: Human Frontiers in Science Program; Grant number: RG0234/2000-M; Grant sponsor: Klaus Tschira Stiftung; Grant sponsor: DAAD fellowship.

*Correspondence to: Dr. Ting Wang, European Media Laboratory, Schloss-Wolfsbrunnenweg 33, 69118 Heidelberg, Germany.
E-mail: Ting.wang@eml.villa-bosch.de

Received 16 April 2002; Accepted 22 July 2002

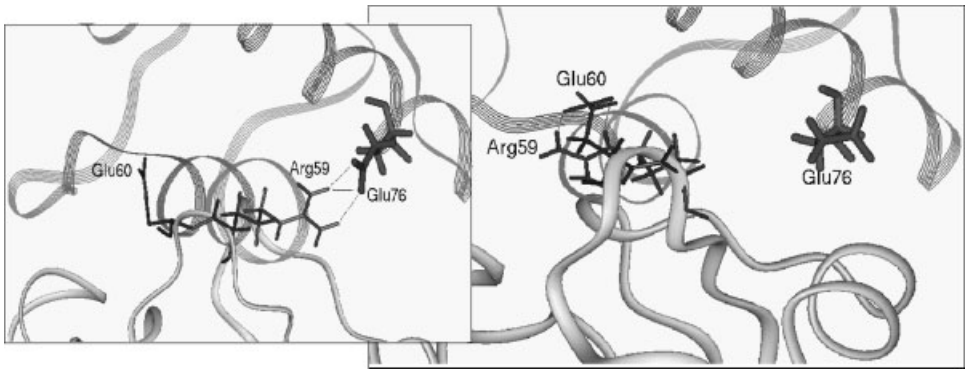

Fig. 1. **Left:** Crystal structure of the barnase (solid ribbon)-barstar (line ribbon) complex, focusing on the key contact between Arg59 of barnase and Glu76 of barstar. **Right:** Docked barnase-barstar structure obtained from simulations with the DDD model. Arg59 of barnase failed to form the key contact with Glu76 of barstar because of the formation of the intramolecular salt link between Arg59 and Glu60 of barnase.

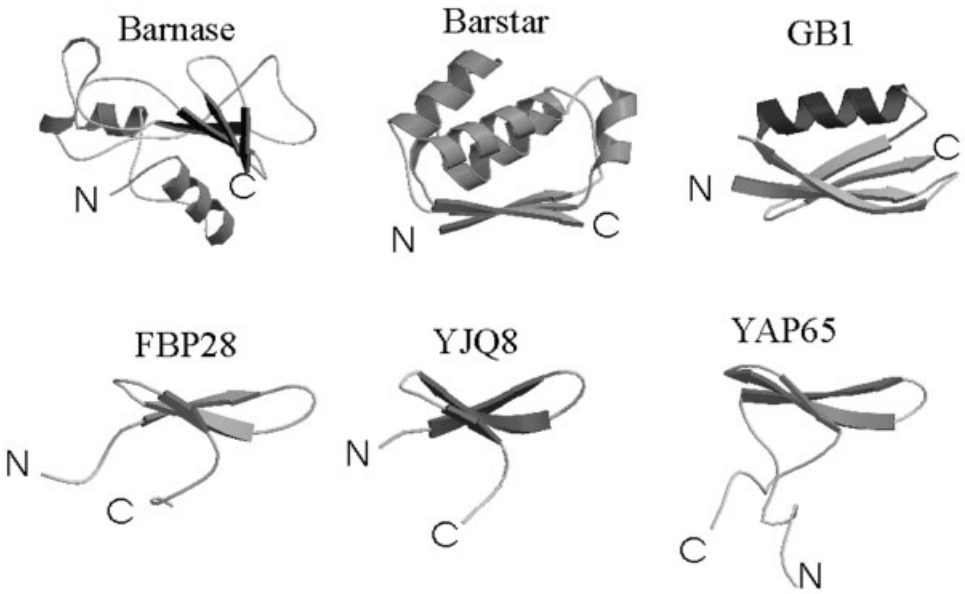

Fig. 2. Secondary structures of the six protein structures studied: barnase, barstar, segment B1 of protein G (G_{B1}), FBP28 WW domain, YJQ8 WW domain, and YAP65 WW domain.

model that we parameterized and call the NPSA model. In the latter model, the partial charges of the ionizable side-chains were neutralized (N) and polarized (P), and a surface area (SA)-dependent solvation term was added. Similar terms have been used to implicitly model the solvent by other authors.[12–15] Ionizable groups are neutralized to avoid the overestimation of charge interactions in vacuo, whereas, at the same time, charges are adjusted to maintain the strength of salt links. The surface-area dependent term accounts approximately for the energetically favorable solvation of hydrophilic groups and the energetically unfavorable solvation of hydrophobic groups.

First, to investigate the abilities of these three implicit solvent models to maintain the native structures of proteins at room temperature, MD simulations were conducted for each of the following six protein structures: barnase, barstar, segment B1 of protein G (G_{B1}), and three WW domains, FBP28, YJQ8, and YAP65. These six proteins exhibit a diverse range of secondary structures: one three-stranded β-sheet in the WW domains, one four-stranded β-sheet and one long α-helix in G_{B1}, and multiple β-sheets and α-helices in barnase and barstar (see Fig. 2). These proteins also vary in thermal stability.[16,17] For example, the melting temperature is 54.5°C for barnase,[18] 73.3°C for barstar,[18] 87.5°C for G_{B1},[19] 61.1°C for FBP28,[20] and 43.6°C for YAP65.[20] Then, to investigate the performance of the implicit solvent models in protein docking, MD simulations were conducted for the complexation of barnase with its inhibitor barstar and for the complexation of the PIN1 WW domain with a doubly phosphorylated peptide ligand (Tyr-P.Ser-Pro-Thr-P.Ser-Pro-Ser). The initial conformations were the separated unbound structures (except for the peptide for which only the bound structure was available). In both complexations, the electrostatic interactions are important driving forces for docking, and the proteins undergo modest changes in conformation

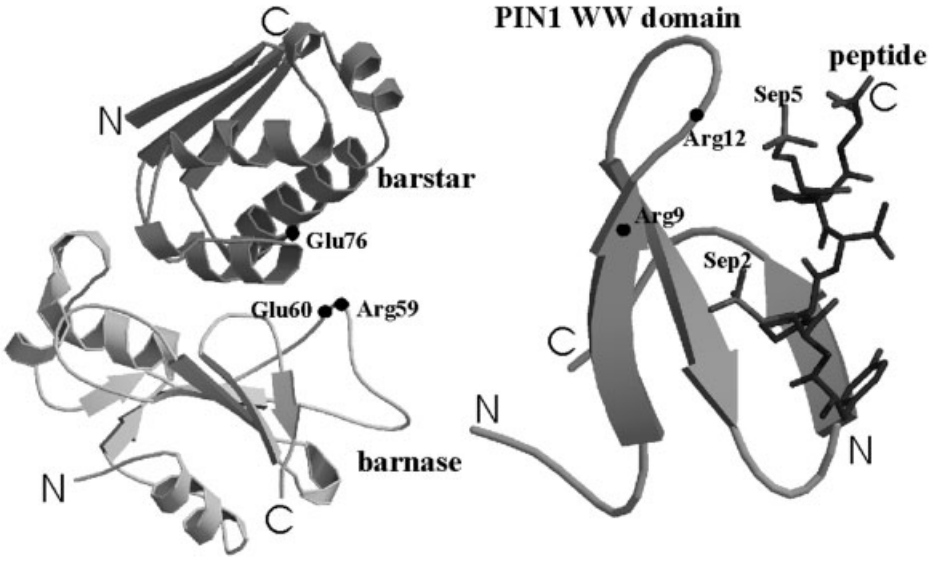

Fig. 3. Crystal structures of the barnase-barstar complex (**left**) and the PIN1 WW domain-phosphoserine peptide complex (**right**). The key residues are labeled.

upon binding. Figure 3 shows the structures of the barnase-barstar complex and the PIN1 WW domain-phosphopeptide complex.

MATERIALS AND METHODS

Implicit Solvent Models

The following three implicit solvent models were used.

Distance-dependent dielectric (DDD) model

In the DDD model, the solvent electrostatic screening effect is simply expressed by a function $\epsilon(r) = nr$ with $n = 1$. This modifies only the coulombic term in the force field. Nonlinear DDD models, such as that of Hassan et al.,[21] are not considered in this article.

Generalized Born (GB) model

The GB model has recently received considerable attention as a rapid approximation to the numerical solution of the Poisson equation. In the GB model, the electrostatic contribution, $\Delta G^{\mathrm{GB}}_{\mathrm{elec}}$, to the free energy of solvation of a solute consists of two terms (Eq. 1). One is the change in coulombic interaction energy on moving a set of N partial charges from vacuum to solvent (the first term in Eq. 1), and the other is the Born solvation self-energy change of the N partial charges (the second term in Eq. 1).

$$\Delta G^{\mathrm{GB}}_{\mathrm{elec}} = -\left(1-\frac{1}{\varepsilon}\right)\sum_{i=1}^{N}\sum_{j=i+1}^{N}\frac{q_i q_j}{r_{ij}} - \frac{1}{2}\left(1-\frac{1}{\varepsilon}\right)\sum_{i=1}^{N}\frac{q_i^2}{a_i} \quad (1)$$

where q_i and q_j are atomic partial charges, r_{ij} is the distance between partial charges i and j, ε is the solvent dielectric constant, and a_i is the effective atomic Born radius, which describes the distance from a charge to the dielectric boundary (the radius of the charge cavity) and depends on the positions and volumes of all the other atoms in the solute. Therefore, in the GB model, charge–charge interactions depend not only on the intercharge distance but also on the extent of solvent exposure of the charges. The two terms in Eq. 1 are usually merged into a single expression in Eq. 2:

$$\Delta G^{\mathrm{GB}}_{\mathrm{elec}} = -\frac{1}{2}\left(1-\frac{1}{\varepsilon}\right)\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{f_{\mathrm{GB}}} \quad (2)$$

Equation 1 is only valid for widely separated charged particles. To account for the cross-polarization effects that arise between close partially charged atoms in a molecule, Still et al.[10] introduced the following functional form of f_{GB} (Eq. 3):

$$f_{\mathrm{GB}} = \sqrt{(r_{ij}^2 + a_i a_j \exp(-r_{ij}/4a_i a_j))} \quad (3)$$

The effective Born radii are the key parameters in the GB model, which can be accurately computed by a numerical integration procedure as proposed by Still and coworkers.[10] But to be a computationally efficient model, approximations have to be introduced for the effective Born radius calculation. To date, four analytical approximation methods have been developed. These are the pairwise descreening approximation introduced by Hawkins et al.,[22,23] the pairwise approximation with empirical parameters (P1–P5) introduced by Qiu et al.,[24] the pairwise Gaussian atomic volume approximation introduced by Schaefer and Karplus,[25] and the surface integral formulation introduced by Ghosh et al.[26] The GB models with the different approximation methods perform differently, but in general they are able to reproduce solvation energies in good agreement with experiment for small molecules.[27,28] However, macromolecules appear challenging. Parameterizations and applications, particularly to proteins, are still undergoing testing.[28] Recently, the GB model was used for MD simulations of a 10-base pair DNA/RNA[29] and five

TABLE I. Modified Partial Charges (e) for the Ionizable Side-Chains, N-, C-termini, and Phosphoserine (SEP) Used in the NPSA Model

	ASP		GLU		LYS		ARG		SEP[a]	
Ionizable Side-chains	CG	1.3994	CD	1.4054	NZ	−1.5854	$NH_{1,2}$	−0.9627	OG	−0.2946
	$OD_{1,2}$	−0.6514	$OE_{1,2}$	−0.6688	$HZ_{1,2,3}$	0.4400	$HH_{1,2,3,4}$	0.3978	P	1.8200
	CB	0.0697	CG	0.1136	CE	−0.1143	CZ	0.3076	$OP_{1,2,3}$	−0.8400
							HE	0.2456		

	NGLY		NPRO		NMET		NLYS		NTYR	
NTER	H	0.2642	H	0.3120	H	0.2984	H	0.3165	H	0.2873
	N	−1.0057	N	−1.2020	N	−1.1408	N	−1.2034	N	−1.1060

	CLYS		CPRO		CMET		CARG		CSER		CGLU	
CTER	O	−0.5252	O	−0.5697	O	−0.6105	O	−0.5266	O	−0.5632	O	−0.5930
	C	1.2488	C	1.2631	C	1.4013	C	1.2550	C	1.3113	C	1.3420

[a]A net side-chain charge of −2e was used in the DDD and GB models. The atomic partial charges were assigned as the following, based on the AMBER potential in InsightII: OG = −0.6546e, P = 1.49e, $OP_{1,2,3}$ = −0.91e. A net side-chain charge of 0e was tried for the NPSA model but gave worse results than when a net charge of −1e was used.) Therefore, the atomic partial charges were assigned as the following on the basis of the CHARMM22 potential in InsightII: OG = −0.2946e, P = 2.82e, $OP_{1,2,3}$ = −0.84e.

proteins[28,30–34] by the first three approximation methods, respectively. Here, we use the GB model incorporated in the program AMBER6.0,[35] which adopts the approximation by Hawkins et al.[22,23] to conduct the MD simulations of proteins and, for the first time, protein docking.

Surface-area dependent (NPSA) model

The NPSA model is similar to surface-area dependent models of Wesson and Eisenberg,[12] Fraternali and van Gunsteren,[13] Lazaridis et al.,[14] and Ferrara et al.[15] In the NPSA model, the DDD function $\epsilon(r) = nr$ was used as in the DDD model. The value of the coefficient n depends on force fields. In this study, it was set to 1 after testing different values. To further mimic the solvent-screening effect on the electrostatic interactions of charged groups, the atomic partial charges on the side-chains of the ionic amino acids were neutralized, but the polarity was significantly increased to stabilize salt links and to prevent the burial of charged groups. In addition, to account for nonpolar solvation effects, a solvent-accessible surface area term was incorporated as in Eq. 4. The atomic solvation parameter σ_i is set to 0.012 kcal/mol · Å^2 for nonpolar carbon and sulfur atoms, −0.060 kcal/mol · Å^2 for polar nitrogen and oxygen atoms and zero for other atoms. The same values of the atomic solvation parameter σ_i were used by Ferrara et al.[14,15] together with the CHARMM19 force field. The solvent-accessible surface area A_i was computed by the approximate analytical method proposed by Weiser et al.[36] and implemented in AMBER6.0. To use this surface term with the NPSA model in AMBER6.0, we modified the source code of the Sander module because the surface calculation is bound to the GB/SA model and only a single atomic solvation parameter is used in the GB/SA model.

$$\Delta G^{\mathrm{NPSA}} = -\left(1 - \frac{1}{r_{ij}}\right)\sum_{i=1}^{N}\sum_{j=i+1}^{N}\frac{q_i' q_j'}{r_{ij}} + \sum_{i}^{N}\sigma_i A_i \qquad (4)$$

The modified partial atomic charges, q_i', of the ionizable side-chains (Asp, Glu, Lys, and Arg) and N- and C-terminii occurring in the proteins were derived in two steps. First, the original AMBER charges of terminal atoms were changed to make the whole residue neutral. That was to make negative charges less negative and positive charges more positive for negatively charged groups and the opposite for positively charged groups. Then the neutralized charges of oppositely charged terminal atoms (H, N of Lys, Arg, and N-termini; O, C of Asp, Glu, and C-termini) were changed to increase the polarity. As a result, compared to the original AMBER partial atomic charges, each of the terminal oxygen atoms of Asp, Glu, and C-terminii was less negative by 0.2 e, and the terminal carbon atoms were more positive by 0.6 e. Each of the terminal hydrogen atoms of Lys and N-terminii was more positive by 0.1 e, the atom NZ of Lys more negative by 1.2 e, the atom CE of Lys more negative by 0.1 e, and the terminal nitrogen atom N of N-terminii more negative by 1.3 e. For Arg, each of the hydrogen atoms $HH_{1,2,3,4}$ was less positive by 0.05 e, each of atoms $NH_{1,2}$ less negative by 0.1 e, CZ less positive by 0.5 e, and HE less positive by 0.1 e. Table I lists the atoms with modified partial charges for each residue.

Preparation of Protein Structures

Protein structures solved by either X-ray crystallography or NMR experiments were used as the initial structures for the simulations.

The unbound barnase structure was taken from chain A of the crystal structure solved at 2.1 Å resolution (PDB entry 1bnj), and the unbound barstar structure was the minimized average structure from NMR experiments (PDB entry 1bta). The chains A and D of the crystal structure of the barnase-barstar complex solved at 2.0 Å resolution (PDB entry 1brs) were used for the bound barnase and barstar coordinates, respectively

The structure of segment B1 of protein G (G_{B1}) was taken from the crystal structure solved at 1.92 Å (PDB entry 1pgb).

The first models of the NMR structures (PDB entry 1e0l and 1e0n) were used for the WW domains FBP28 and YJQ8, respectively. The NMR structure of the YAP65 WW

domain was kindly provided by Dr. Maria Macias (European Molecular Biology Laboratory, Heidelberg, Germany).

For docking simulations, separated unbound structure pairs were prepared as the starting structures. This was done by superimposing the unbound structures onto the bound structures in a complex, which had been manually separated by 5 Å along the line between their mass centers. The unbound structures and the structure of the complex between barnase and barstar were as described above. For PIN1 WW domain-peptide docking, the unbound PIN1 WW domain was extracted from the full-length PIN1 crystal structure solved at 1.3 Å (PDB entry 1pin), and the bound structure was taken from the full-length PIN1-peptide crystal structure solved at 1.84 Å (PDB entry 1f8a). In addition, the orientation of barnase and barstar in a diffusional encounter complex produced by Brownian dynamics simulations[37] of their diffusional association was used to provide alternative starting structures for barnase-barstar docking simulations. This "two-contact" encounter complex has, with the bound protein conformations, a 4.6 Å barstar interface Cα root-mean-square deviation (RMSD) from the crystallographic complex.

The all-atom AMBER94 force field was used for all the molecules except the phosphoserine peptide. The atomic partial charges of the phosphate group in the peptide were assigned according to the AMBER potential in the InsightII software package,[38] and the phosphoserine residue (three letter symbol: Sep) was assigned a net double negative charge of −2e in the DDD and GB models but only a single negative charge of −1e in the NPSA model (see Table I). The prepared structure of each protein–protein pair was loaded into the Xleap module of AMBER6.0 to obtain the topology and the coordinate files for the subsequent energy minimizations and MD simulations.

Simulation Protocols

The prepared starting structures were first energy minimized by 200 steps of conjugate gradient minimization in AMBER6.0. The relaxed structures were then subjected to MD simulations. First, the system was gradually heated from 0 to 300 K in 40 ps, and then it was equilibrated for 260 ps with a time constant for heat bath coupling of 0.5 ps at 300 K. After this equilibration, a 1-ns production run was performed with a heat bath coupling constant of 1.0 ps at 300 K. In all of the simulations, the timestep was 1 fs, and the nonbonded interactions were updated every 10 timesteps with a cutoff of 10 Å. Overall translational and rotational motion was removed every 0.1 ps. The bonds involving hydrogen atoms were constrained by the SHAKE algorithm. For simulations with the GB model, the effective Born radius offset was set to −0.13 Å. Coordinates were saved every 1 ps during the 1.3-ns simulation.

For the docking simulations, only 300-ps simulations (40-ps heating and 260-ps equilibrium) were conducted initially because the complexes usually formed within 100 ps. Selected simulations were then extended for longer times.

All the simulations were performed with AMBER6.0. For each single protein simulation and each docking simulation, three runs were performed, each with a different implicit solvent model (DDD, NPSA, and GB, respectively).

RESULTS AND DISCUSSION

Simulations of Single Native Structures

Trajectory analysis

The behavior of the proteins during MD simulation with the different solvent models was compared by evaluating the following structural and dynamic parameters.

1. RMSD from the starting structure. Because the MD protocol is expected to maintain the starting structures in native conformations, a good solvent model should give small RMSD values that become constant during the trajectory.
2. Root-mean-square fluctuation (RMSF). The thermal mobility of the protein should be manifested by reasonable atomic fluctuations during the simulations.
3. Radius of gyration. This measure of compactness of the molecule is expected to match the experimentally measured value.

The time evolutions of these parameters are plotted in Figure 4 and the average values are listed in Table II.

The left-hand column in Figure 4 shows the time evolution of the RMSD of the backbone atoms in regions of α/β secondary structure during the 1-ns production runs. The largest RMSDs were obtained either for the GB model or for the DDD model, and the RMSDs showed the most unstable behavior for the GB model. On the other hand, the NPSA model yielded trajectories with stable low RMSDs in all cases. The average α/β backbone atom RMSD values in NPSA simulations are 1.0 Å for barnase, 1.4 Å for barstar, 0.7 Å for G_{B1}, 1.0 Å for FBP28, 2.0 Å for YJQ8, and 1.0 Å for YAP65, indicating complete preservation of the protein secondary structures in all cases except for YJQ8. YJQ8 is expected to be less stable than the other WW domains because it lacks one of the conserved signature residues of this domain.[16] In general, FBP28 is more stable than YAP65 and less sensitive to variation in the implicit solvent models. This is consistent with FBP28's higher thermal stability[16,17] and its greater stability in MD simulations with explicit solvents.[17] YAP65 is particularly sensitive to the choice of implicit solvent model, and this is not surprising when one considers that it was previously observed to be highly sensitive to modeling of the ionic strength in MD simulations with explicit solvent.[39]

The percentage of residues participating in α-helices and/or β-sheets is 53% in barnase, 61% in barstar, 68% in G_{B1}, 35% in FBP28, 48% in YJQ8, and 30% in YAP65. In general, the greater the proportion of α/β residues, the more the stability of the secondary structure represents the stability of the whole structure. This can been seen by

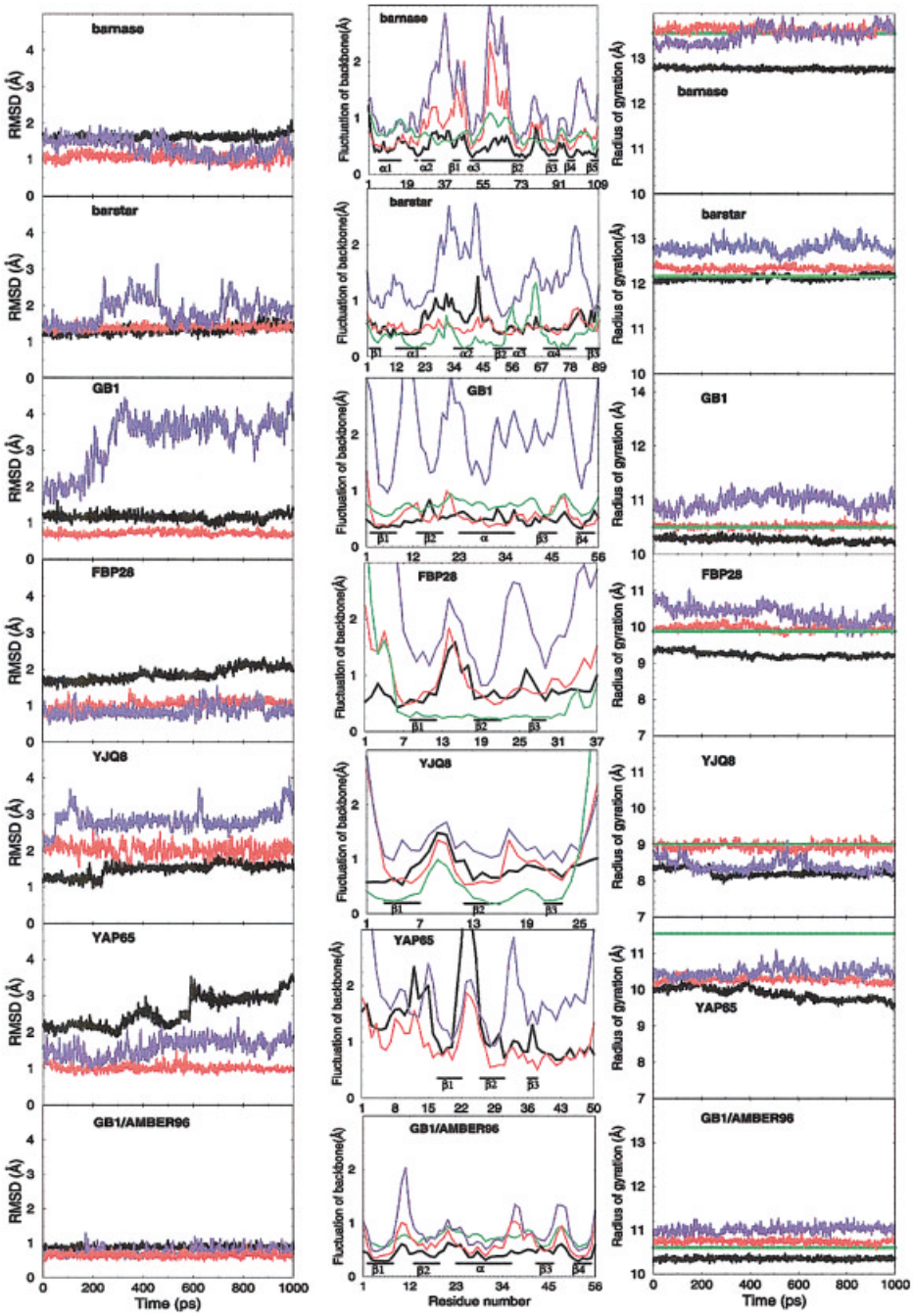

Fig. 4. Time evolution of the RMSD of α-helix and β-sheet backbone atoms (**left** column), fluctuation of backbone atoms averaged over residues (**middle** column), and radius of gyration (Rg) (**right** column), in the production runs of 1-ns duration with different solvent models. Red: NPSA model; blue: GB model; black: DDD model; green: X-ray or NMR experimental data. The bottom row shows the results with the AMBER96 force field for G_{B1}.

the difference between the RMSD values of α/β backbone atoms and of all backbone atoms listed in Table II. For example, the differences are small in barstar and G_{B1} but large for the WW domains with long disordered N- and C-termini.

The large secondary structure backbone RMSDs of the three WW domains in some of the simulations (see Table II) are the result of extensive unfolding of YJQ8 with the GB and the NPSA models and of both YAP65 and FBP28 with the DDD model. The analysis of the trajectories showed that the unfolding of YJQ8 began with the breaking of the hydrogen bonds between β1 and β2 strands with the GB model and the breaking of the hydrogen bonds near turn 2 with the NPSA model. The unfolding of YAP65 and

TABLE II. Average Values of Backbone RMSD, α/β Backbone RMSD, Residue Fluctuation, and Radius of Gyration in the Production Run of 1 ns with Different Solvent Models

	Backbone RMSD (Å)			α/β Backbone RMSD (Å)			Residue fluctuation (Å)				Radius of gyration (Å)			
	NPSA	GB	DDD	NPSA	GB	DDD	NPSA	GB	DDD	Expt	NPSA	GB	DDD	Expt
Barnase	1.62	2.93	2.43	1.06	1.31	1.63	0.81	1.28	0.49	0.54	13.56	13.49	12.78	13.56
Barstar	1.60	2.18	1.70	1.40	1.83	1.34	0.54	1.40	0.62	0.40	12.34	12.81	12.13	12.16
GB1	1.11	4.56	1.39	0.72	3.30	1.17	0.55	2.02	0.49	0.54	10.64	11.15	10.33	10.60
GB1[a]	0.93	1.23	1.02	0.65	0.76	0.88	0.62	0.84	0.44	0.54	10.74	11.01	10.35	10.60
FBP28	2.39	3.40	2.97	1.03	0.85	1.86	0.96	2.15	0.76	0.56	9.98	10.37	9.24	9.88
YJQ8	3.96	4.13	2.57	2.03	2.84	1.49	1.01	1.35	0.84	0.67	8.95	8.41	8.23	9.02
YAP65	6.38	5.82	5.46	1.02	1.59	2.56	0.97	1.89	1.31	nd.	10.29	10.49	9.90	11.55

[a]Results with AMBER96 force field; all other values were obtained by using the AMBER94 force field.

FBP28 with the DDD model began with the breaking of the hydrogen bonds near turn 2, which resulted in the simultaneous loss of the three β-strands. These unfolding mechanisms observed for YAP65 and FBP28 with the DDD model are not consistent either with experiments[20,40,41] or with MD simulations with explicit solvents,[17] which revealed the loss of the β3 strand as the initial stage of unfolding of these WW domains. Furthermore, in the simulation of FBP28 with the DDD model, the side-chain conformation of Trp30 changed to another rotamer. This rotamer transition did not occur with either the GB model or the NPSA model. Because Trp 30 is located on the peptide-binding face of the FBP28 WW domain, the side-chain conformation of Trp30 is important for binding of proline-rich peptides too, and consequently, the DDD model for FBP28 would not be not suitable for peptide-docking simulations.

The middle column in Figure 4 shows the fluctuations of backbone atoms averaged over residues during the production runs of 1-ns duration. For each protein, the secondary structure regions are marked by α for α-helices and β for β-sheets. It can be seen that the GB model yielded fluctuations with much larger magnitude than the NPSA and DDD models in each case. For comparison with experimental data, the B-factor values of backbone atoms averaged over each residue of the crystal structures of barnase and G_{B1} (PDB entries: 1bnj and 1pgb) were plotted in Figure 4 after scaling by $8\pi^2/3$ and taking the square root.

For barnase, the fluctuation with the NPSA model is consistent with the B-factor plot in most regions, with the main discrepancy being in the loop region between β2 and β3, where the B-factors indicate low mobility but the NPSA model gives a fluctuation peak. Considering the fact that the low B-factors in this region are due to crystal contacts,[42] the large fluctuation in the noncrystalline environment of the simulations is reasonable. But the sharp fluctuation peak in the simulations, which is particularly strong with the GB model, in the low B-factor region between α2 and β1 cannot be accounted for by the same reason. For G_{B1}, a marked feature of the B-factor plot is a long region of low values along the α-helix region (13–26). This was reproduced well by the NPSA model, to some extent by the DDD model, but not by the GB model.

For barstar, FBP28, and YJQ8, crystallographic B-factor values are not available, but the protein fluctuations can be characterized by the ensemble of structures obtained in the NMR structure determinations. The average fluctuations of backbone atoms per residue in the NMR structural ensembles were calculated and are plotted in Figure 4. For barstar, it can be seen that the fluctuations with the GB model are much larger than in the NMR ensemble (see Table II). The NPSA and DDD models show much lower fluctuations. The plot for the NPSA model is similar to the NMR fluctuations in most regions. The exception is the region between β2 and α4, which shows the greatest mobility in the NMR ensemble but is small for the NPSA and DDD models. For the three WW domains, the fluctuation curves for the simulations are rather similar in shape, showing troughs in the β-strands and peaks in the loop regions and terminii. For FBP28, the 10 structures in the NMR ensemble have very similar conformations and, consequently, do not show variations in the magnitude of fluctuations over the core region. However, the fluctuation in the NMR ensemble for YJQ8 has a similar distribution to that for the simulations.

The right-hand column in Figure 4 shows the time evolution of the radius of gyration (Rg) during the 1-ns production run performed for each protein. It can be seen that the NPSA model yielded much more stable, constant Rg values than the GB and DDD models. To compare with the experimental starting structure, the Rg value of the starting structure (X-ray or NMR structure) of each protein is marked by a horizontal line in Figure 4. The DDD model always adopts the lowest Rg values, which are mostly below the Rg of the experimental structure. This shows that the DDD model leads to artifactual compaction of protein structures. Rg increased above the experimental value in simulations with both GB and NPSA models for all the protein structures except YJQ8 and YAP65. The Rg values with the GB model tend to be larger than those with the NPSA model. A slight increase of Rg is assumed to be realistic for crystal structures due to the absence of crystal packing in the simulations, but a large increase may imply unfolding of the protein structures. Hence, in this respect, the NPSA model performed better than the GB and DDD models, most often producing the closest Rg values to experimental values. For YJQ8, the NPSA model almost kept the Rg of the starting structure, whereas the GB and DDD models yielded reduced Rg values. For YAP65, both GB and NPSA models yielded about a 1 Å decrease in Rg

compared to the starting structure. This might arise from overall compaction of the long unstructured N- and C-terminii during the simulations.

Computational time

The simulations were implemented on a multiprocessor Origin2000 computer and were run on two to six processors with the parallel version of AMBER6.0. To compare the computational times of different models, single processor runs were performed for barnase and FBP28 specifically for this purpose. For barnase (109 amino acids), the 1-ns production run took 82.7 h (3.4 days) with the DDD model, 273.3 h (11.4 days) with the NPSA model, and 557.1 h (23.2 days) with the GB model. For FBP28 (37 amino acids), the 1-ns production run took 16.7 h (0.7 days) with the DDD model, 67.6 h (2.8 days) with the NPSA model, and 109.2 h (4.5 days) with the GB model. Accordingly, the NPSA model is about 3.4–4 times slower than the DDD model, and the GB model is about 6.5–6.7 times slower than the DDD model. The time distribution showed that about 76% of computational time was spent on the calculation of the surface term in the NPSA model and about 77% on the Born radii calculation in the GB model. Therefore, an improved NPSA timing can be expected by implementing a faster surface term computation.

Amber 96 force field

It is surprising that the highly thermostable protein G_{B1} showed dramatic deviations from the native structure with the GB model. The secondary structure of G_{B1} was deformed during the simulation, and an obvious structural change was that the four-stranded β-sheet was reduced to less than half its original length. Destabilization of β-structure in simulations with the AMBER94 force field has been noted before, and a reparameterization to improve the stability of β-structure was introduced into the AMBER96 force field.[43] Therefore, we conducted simulations with the three solvent models and AMBER96 for G_{B1}. The results are shown in Figure 4 and Table II to compare to those with AMBER94. It can be seen that the results for the GB model were significantly improved with AMBER96 in RMSD and fluctuation. The changes in AMBER96 prevented loss of secondary structure. There were also slight improvements for the NPSA and DDD models with the AMBER96 force field. Despite these considerable improvements in the simulations with the GB model, the NPSA model still provided the best agreement with experimental structural data.

To further investigate the effect of the AMBER96 force field on the performance of the GB model, we simulated barstar, in which the main secondary structures are α-helices. The RMSD evolution of the secondary structure is plotted in Figure 5 to compare with the previous result. It can be seen that the remarkable improvement for G_{B1} with the GB model is absent for barstar.

Comparison with Prior Studies

DDD model

The characteristics of our simulations with the DDD model are consistent with observations from other simulations with this model. The DDD model tends to overcompact proteins resulting in too small a radius of gyration and reduced fluctuation. Deviations in structure occur because of artifacts in the electrostatic model resulting in features such as overformation of surface salt bridges.[2,13,30]

GB model

For MD simulations of proteins with GB models, the reports in the literature[28,30–34] can be classified into three groups according to the approximation methods used to calculate the Born radii. The first group used the Born radii of Qiu with the CHARMM force field and simulated the protein G_{B1},[31] a designed 20-residue, three-stranded β-sheet protein Betanova,[32] and the active site of HIV protease.[28] Results comparable to explicit water simulations or Poisson continuum models were obtained. The second group used the Born radii of Hawkins with the AMBER94 force field and simulated a homodimer protein, interleukin-8.[34] The results were in general similar to explicit water simulations but with larger fluctuations. The third group used the Born radii of Schaefer implemented in the ACE potential and simulated the proteins G_{B1} and Raf-RBD.[30] Trajectories were found to be unstable on a long timescale (>1 ns), and even in the early stable stage, the backbone RMSD was in the range of 1.5–2.5 Å. Considering the fact that most residues of protein G_{B1} participate in the secondary structure, this RMSD value is too large for complete maintenance of the secondary structure. Methodologically, our GB simulations belong to the second group. They differ in the apparent instability and large fluctuations observed with the AMBER94 force field, although these were partially remedied by switching to the AMBER96 force field. It appears that the GB parameterization used in the first group outperformed others. However, it is also worth noting that in the first group's simulations of G_{B1} and Betanova, velocity reassignment was used throughout to maintain the temperature, which may have limited deviations from experimental structures. The sensitivity of GB performance to parameters is also shown in a recent article,[33] in which results of the GB model were improved by increasing the atomic radii of hydrogen atoms bonded to nitrogen atoms.

NPSA model

A number of studies have been conducted with surface area-dependent terms and altered charge distributions. The most similar model to our NPSA model is probably that used by Ferrara et al.[15] who applied it to folding studies on chymotrypsin inhibitor 2 (CI2), a 64-residue α/β protein. The CHARMM force field with the function, $\epsilon(r) = nr, n = 2$, was used, and the folding success rates were 50% from random initial conformations and 83% from initial conformations unfolded by high-temperature MD simulation. More recently, the same protocol was used to perform 1-ns MD simulations at 300 K for six small proteins[44] (with a maximum number of residues of 85), and reasonable trajectories were obtained. But large deviations from the native conformations were observed in longer simula-

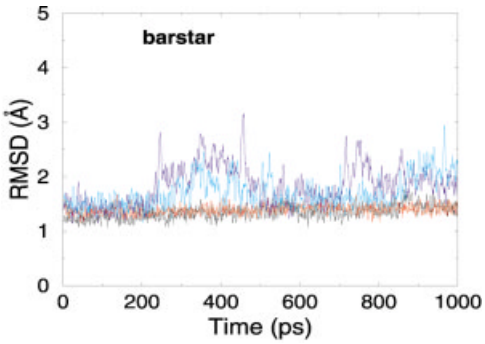

Fig. 5. The light blue line shows the α/β backbone RMSD of barstar obtained in the GB simulation with the AMBER96 force field. The other three lines (dark blue, red, and black) show the results with AMBER94 force field, as in Figure 4.

tions as well as in simulations of three extra proteins, one of which was barnase (PDB entry 1a2p) with a Cα RMSD of 3.8 Å over the 0.5- to 1.0-ns time interval. There are several differences between the protocol of Ferrara et al. and our NPSA protocol. First, the force fields and parameters, in particular partial atomic charges, are different. Our NPSA protocol is based on the AMBER all-atom force field and uses a smaller DDD coefficient n of 1. Second, the algorithms to calculate solvent-accessible surface areas are different. The algorithm used in the studies of Ferrara et al. is apparently more rapid, but probably less accurate than that used in our studies. As shown in this article, our NPSA protocol performed well in maintaining the protein native structures with reasonable atomic fluctuations, indicating the ability to provide a realistic approximation to solvent effects.

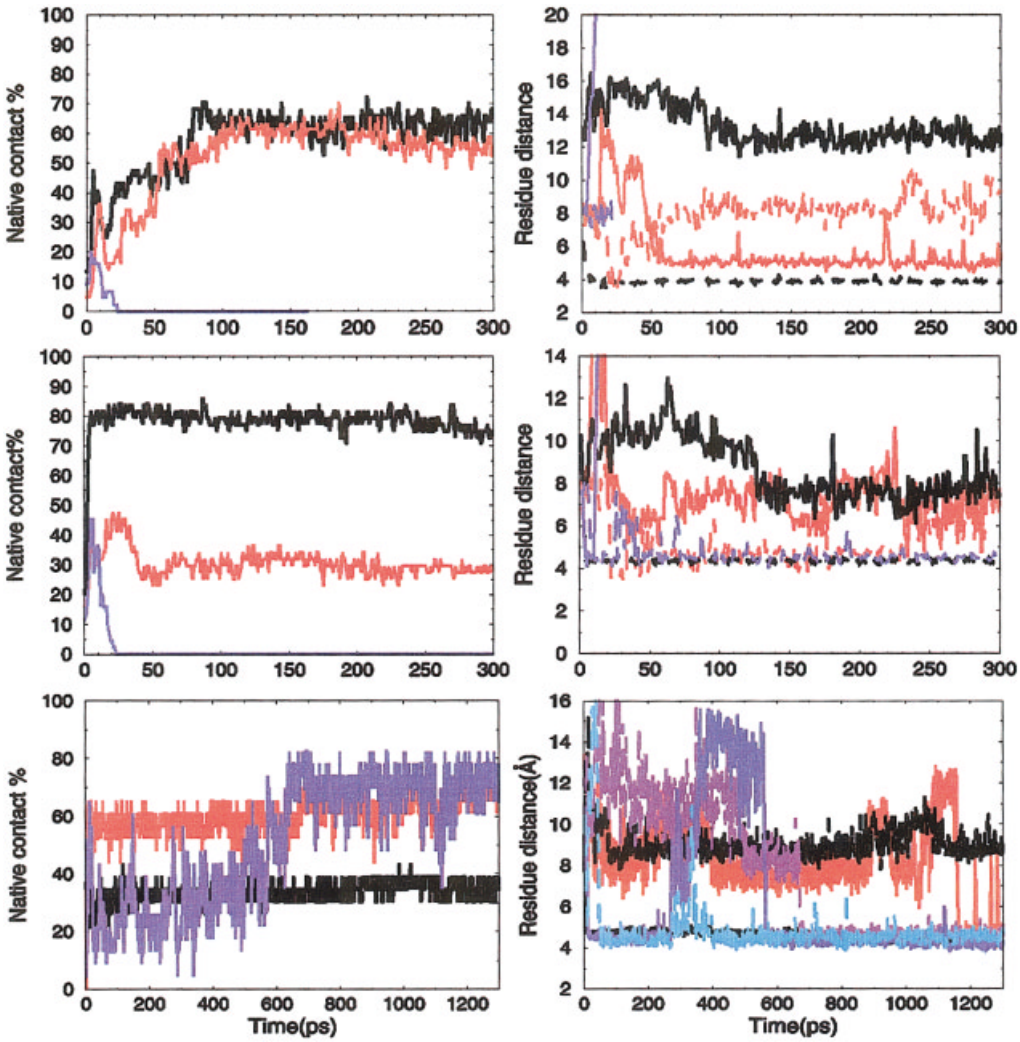

Fig. 6. Time evolution of native contact percentage (**left** column) and distances between key residues (**right** column) during the docking simulations with the different solvent models. Red, violet: NPSA model; blue, cyan: GB model; black: DDD model. 1st row: barnase-barstar docking, starting from 5 Å separated unbound structures; 2nd row: barnase-barstar docking, starting from unbound structures in a two-contact encounter complex; 3rd row: PIN1 WW domain-phosphoserine peptide docking, starting from 5 Å separated structures with the unbound WW domain structure. **Solid line:** distance between Arg59 of barnase and Glu76 of barstar and distance between Arg12 of the WW domain and Sep5 of the peptide; **dashed line:** distance between Arg59 of barnase and Glu60 of barnase and distance between Arg9 of the WW domain and Sep2 of the peptide.

Docking Simulations

The docking results were first analyzed by the percentage of native residue contacts existing in the bound complex that are formed in the docking simulation. Two residues are defined as in contact if the van der Waals spheres of any atoms (excluding nonpolar hydrogens) of these residues are within 1 Å of each other. This was one of the criteria for assessing docking in CASP2.[45] According to this definition, the number of native residue contacts in the barnase-barstar complex is 44, and in the PIN1 WW domain-peptide complex is 23.

Furthermore, the distances between the interface residues involved in conformational changes on docking was monitored with the aim of investigating the abilities of the different solvent models to address the protein flexibility problem. The main conformational change on the binding of barnase to barstar occurs in the loop containing Arg59 of barnase. Although the overall Cα RMSD of barnase is only 0.48 Å, the Cα atom of Arg59 shows 2 Å deviation between the unbound and bound conformations, and the side-chain of Arg59 rotates to form a salt link to the Glu76 of barstar. In addition, as mentioned in the Introduction, docking with the DDD model did not show formation of the contact between Arg59/barnase and Glu76/barstar but, instead, an intramolecular salt link between Arg59/barnase and Glu60/barnase. Therefore, the distances between Arg59/barnase and Glu76/barstar as well as between Arg59/barnase and Glu60/barstar were monitored for each barnase-barstar docking simulation.

On PIN1 WW domain-peptide binding, although the WW domain undergoes a slight global twist resulting in an overall Cα RMSD of 1.34 Å, the main conformational change in the WW domain occurs in the loop around Arg12. The loop backbone moves up toward the C-terminal of the peptide, and the side-chain of Arg12 forms a salt link to the phosphoserine Sep5 of the peptide on binding. The other phosphoserine, Sep2, comes within hydrogen-bonding distance of Arg9/PIN1. Experiments[46,47] showed that both phosphate groups contribute to binding and together result in the highest binding affinity of this peptide to full-length PIN1. Therefore, the distances between Arg12/PIN1 and Sep5/peptide and between Arg9/PIN1 and Sep2/peptide were monitored for each docking simulation. Note that the PIN1 WW domain residue numbers refer to those in the free PIN1 WW domain NMR structure with PDB entry 1i6c. The inter-residue distances were actually evaluated as interatomic distances between Arg:CZ and Glu:CD atoms for Arg-Glu distances and between Arg:CZ and Sep:P atoms for Arg-Sep distances. A hydrogen bond can be considered to form if these distances are shorter than 5 Å.

Barnase-barstar docking

The top left plot in Figure 6 shows the time evolution of the formation of native contacts in 300-ps simulations, starting from 5 Å separated unbound structures of barnase and barstar. It can be seen that complexation occurs with both DDD and NPSA models but not with the GB model. In fact, in the simulation with the GB solvent model, barnase and barstar appear to "run away" from each other shortly after the beginning of the simulation and barstar later became highly deformed. In the simulation with the NPSA solvent model, 50% of the native contacts were formed after about 60 ps, and from then on about 60% of the native contacts were present for the remainder of the simulation. The simulation with the DDD model showed a similar contact percentage formation, and the fraction of native contacts stabilized at a slightly higher percentage of 63%. However, as mentioned in the Introduction, the DDD model resulted in formation of additional non-native intramolecular hydrogen bonds by forcing the oppositely charged residues close to each other. For example, the interface residues Arg59/barnase and Glu76/barstar formed intramolecular salt links with Glu60/barnase and Arg75/barstar, respectively, leaving a "binding gap" without intermolecular salt links in this region. (See the top right plot in Figure 6 for the evolution of inter-residue distances during the simulations.) Such a binding gap would be misleading in studies in which docking might be used to guide mutagenesis because it could lead to energetically important residues being missed. In barnase and barstar, mutation of Arg59 to alanine causes a drop in binding free energy of about 5 kcal/mol, and this residue is considered a hot-spot residue.[48] This binding gap was partially remedied in the NPSA model, in which Arg59/barnase formed a hydrogen bond with Glu76/barstar, although the further movement of Arg59/barnase to Glu76/barstar to form more hydrogen bonds was hindered by the hydrogen bond of Arg59/barnase with Asp35/barstar. Nevertheless, most native interface hydrogen bonds were formed in the simulation with the NPSA model with maintenance of internal hydrogen bonds.

The middle left plot in Figure 6 shows the time evolution of the native contacts in 300-ps simulations, starting from the two-contact encounter complex with unbound structures. The distances between the key residues involved in the conformational change are shown in the middle right plot in Figure 6. It can be seen that there is again no complexation with the GB solvent model, whereas the DDD solvent model yielded considerably more native contacts than the NPSA solvent model (80% vs 30%). The two-contact encounter complex from the BD diffusion simulations has a center-of-mass to center-of-mass distance between barnase and barstar that is about 4 Å longer than in the crystal structure of the complex, and the line between the centers of mass deviates about 2° from that of the crystallographic complex. The overall Cα RMSD is 1.3 Å between the two-contact encounter complex starting structure and the 5 Å separation starting structure. Longer simulations (>1 ns) with both NPSA and DDD models did not show improvement of formation of native contacts or the formation of the interface hydrogen bond between Arg59/barnase and Glu76/barstar. However, docking simulations starting from the bound two-contact encounter complex gave much better results with all three solvent models yielding >80% native contacts and Arg59/barnase forming a salt link to Glu76/barstar as in the crystal structure of complex.

"Run away" problem of GB docking

For barnase-barstar docking, we conducted simulations starting from four different conformations and orientations (one 5 Å separated starting structure and one encounter complex, each with bound and unbound conformations) for each solvent model. With the GB model, barnase and barstar docked with formation of most of the native contacts only in the simulations starting with the bound conformations. In the other two simulations with the GB model, a few intermolecular contacts were formed during the first 10 ps, and then the two molecules moved away from each other quickly. It appears that in these simulations, intermolecular attraction was not strong enough to overcome molecular desolvation energies. On the other hand, we noticed that the 10 Å nonbonded cutoff used in the simulation protocol might be too short for docking. So we increased the cutoff to 30 Å and redid the docking simulations with the GB model starting with the unbound structures pulled 5 Å apart. Nevertheless, the run away problem remained. However, the run away problem never occurred in DDD or NPSA simulations.

PIN1 WW domain-peptide docking

As occurred in the barnase-barstar docking simulations, the PIN1 WW domain and the peptide ran away from each other, and no complex was formed in the simulation with the GB model starting from the 5 Å separated unbound structures. But we found by chance that another simulation with the GB model yielded complexation by starting from a slightly different PIN1 WW domain orientation that was generated by superimposing only a fraction of WW domain residues of the free form onto the bound form. Therefore, the latter docking result was used to evaluate the GB model. In addition, dramatic fluctuations of native contact percentage and key residue distance were observed in the 300-ps simulations with both NPSA and GB models. Therefore, the simulations were extended to 1.3 ns, and the native contact percentages reached stable values of about 65% with the NPSA model and about 75% with the GB model (see the bottom left plot in Fig. 6). The DDD model resulted in much less fluctuation of the fraction of native contacts formed, but this fraction leveled out at ~35%. Furthermore, no intermolecular hydrogen bonds to the peptide phosphate groups were formed during the simulation. A hydrogen-bonding distance between Arg9/PIN1 and Sep2/peptide was observed with the NPSA model from 700 ps onward and for most of the time with the GB model (see the bottom right plot in Fig. 6). A hydrogen-bonding distance between Arg12/PIN1 and Sep5/peptide was observed only in the last 200 ps with the NPSA model but most of the time with the GB model. Simulation for a further 1 ns showed that the hydrogen bond distance between Arg9/PIN1 and Sep2/peptide was stable but that between Arg12/PIN1 and Sep5/peptide was unstable with both NPSA and GB models. The structural analysis showed that besides the two Arg-Sep hydrogen bonds, Arg12/PIN1 also formed a hydrogen bond with the Sep2/peptide in the GB simulation. The last snapshot at 1.3 ns showed backbone RMSDs of 2.35 Å and 3.08 Å from the bound PIN1 WW domain with the NPSA and the GB models, respectively.

Docking simulations starting from the bound structures gave native contact percentages of 90% with NPSA, 80% with GB, and 70% with DDD models, respectively. But the close contact between Arg9/PIN1 and Sep2/peptide was not formed with the NPSA model when both Sep5 and Sep2 are within hydrogen-bonding distance of Arg12/PIN1. A further 1-ns simulation still did not show the formation of hydrogen bond between Arg9/PIN1 and Sep2/peptide with the NPSA model.

CONCLUSION

In this study, we compared three different implicit solvent models, the DDD, GB, and NPSA models by MD simulations. The simulations of single protein structures showed the clear advantages of the NPSA model over the GB and DDD models in maintaining the native structures with residue fluctuations consistent with the experimental results. The NPSA model is about 2 times faster than the GB model. For protein–protein docking, the GB model was found to be very sensitive to parameterization and simulation protocols, resulting in frequent failure to form a docked complex. When docked complexes were obtained, the three implicit solvent models produced complexes of similar accuracy, with the NPSA model showing reliable behavior. We conclude that the NPSA model provides a reasonably fast robust approximation of solvent effects that is suitable for applications of MD simulations to explore protein structure and protein–protein docking.

ACKNOWLEDGMENTS

This work was supported in part by the Human Frontiers in Science Program, the Klaus Tschira Stiftung, and a DAAD fellowship to T.W. We thank David Case for helpful advice on using AMBER and Alexey Onufriev for valuable discussion of generalized Born models. We thank Maria Macias for providing coordinates of the YAP WW domain. We thank Lutz Ehrlich, Isabel Tomas Oliveira, and Razif Gabdoulline for assistance with choice and preparation of protein structures for simulations. We thank Peter Winn for critical reading of the manuscript.

REFERENCES

1. Jackson RM, Gabb HA, Sternberg MJ. Rapid refinement of protein interfaces incorporating solvation: application to the docking problem. J Mol Biol 1998;276:265–285.
2. Ehrlich LP, Wade RC. Protein-protein docking. Rev Comp Chem 2001;17:61–97.
3. Mangoni M, Roccatano D, Di Nola A. Docking of flexible ligands to flexible receptors in solution by molecular dynamics simulation. Proteins 1999;35:153–162.
4. Pak Y, Enyedy IJ, Varady J, Kung JW, Lorenzo PS, Blumberg PM, Wang S. Structural basis of binding of high-affinity ligands to protein kinase C: prediction of the binding modes through a new molecular dynamics method and evaluation by site-directed mutagenesis. J Med Chem 2001;44:1690–1701.
5. Buckle AM, Schreiber G, Fersht AR. Protein-protein recognition: crystal structural analysis of a barnase-barstar complex at 2.0-A resolution. Biochemistry 1994;33:8878–8889.
6. Bhat TN, Bentley GA, Boulot G, Green MI, Tello D, Dall'Acqua W, Souchon H, Schwarz FP, Mariuzza RA, Poljak RJ. Bound water

molecules and conformational stabilization help mediate an antigen-antibody association. Proc Natl Acad Sci 1994;91:1089–1093.
7. Cramer CJ, Truhlar DG. Implicit solvatin models: equilibria, structure, spectra, and dynamics. Chem Rev 1999;99:2161–2200.
8. Gelin BR, Karplus M. Side-chain torsional potentials: effect of dipeptide, protein and solvent. Biochemistry 1979;18:1256–1268.
9. Sharp K, Honig B. Electrostatic interactions in macromolecules: theory and applications. Annu Rev Biophys Biophys Chem 1990;19: 301–332.
10. Still WC, Tempczyk A, Hawley RC, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. J Am Chem Soc 1990;112:6127–6129.
11. Bashford D, Case DA. Generalized Born models of macromolecular solvation effects. Annu Rev Phys Chem 2000;51:129–152.
12. Wesson L, Eisenberg D. Atomic solvation parameters applied to molecular dynamics of proteins in solution. Protein Sci 1992;1:227–235.
13. Fraternali F, van Gunsteren WF. An effecient mean solvation force model for use in molecular dynamics simulations of proteins in aqueous solution. J Mol Biol 1996;256:939–948.
14. Lazaridis T, Karplus M. Effective energy function for proteins in solution. Proteins 1999;35:133–152.
15. Ferrara P, Apostokakis J, Caflisch A. Computer simulations of protein folding by targeted molecular dynamics. Proteins 2000;39: 252–260.
16. Macias MJ, Gervais V, Civera C, Oschkinat H. Structural analysis of WW domains and design of a WW prototype. Nat Struct Biol 2000;7:375–379.
17. Ibragimova GT, Wade RC. Stability of the beta-sheet of the WW domain: a molecular dynamics simulation study. Biophys J 1999; 77:2191–2198.
18. Protasevich II, Schulga AA, Vasilieva LI, Polyakov KM, Lobachov VM, Hartley RW, Kirpichnikov MP, Makarov AA. Key role of barstar Cys-40 residue in the mechanism of heat denaturation of bacterial ribonuclease complexes with barstar. FEBS Lett 1999; 445:384–388.
19. Alexander P, Fahnestock S, Lee T, Orban J, Bryan P. Thermodynamic analysis of the folding of the streptococcal protein G IgG-binding domains B1 and B2: why small proteins tend to have high denaturation temperatures. Biochemistry 1992;14:3597–3603.
20. Ferguson N, Johnson CM, Macias M, Oschkinat H, Fersht A. Ultrafast folding of WW domains without structured aromatic clusters in the denatured state. Proc Natl Acad Sci USA 2001;98: 13002–13007.
21. Hassan SA, Guarnieri F, Mehler EL. A general treatment of solvent effects based on screened coulomb potentials. J Phys Chem B 2000;104:6478–6489.
22. Hawkins GD, Cramer C, Truhlar DG. Pairwise solute descreening of solute charges from a dielectric medium. Chem Phys Lett 1995;246:122–129.
23. Hawkins GD, Cramer CJ, Truhlar DG. Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. J Phys Chem 1996;100:19824–19839.
24. Qiu D, Shenkin PS, Hollinger FP, Still WC. The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate Born radii. J Phys Chem A 1997;101:3005–3014.
25. Schaefer M, Karplus M. A comprehensive analytical treatment of continuum electrostatics. J Phys Chem 1996;100:1578–1599.
26. Ghosh A, Rapp CS, Friesner RA. Generalized Born model based on a surface integral formulation. J Phys Chem 1998;102:10983–10990.
27. Best SA, Merz KM, Reynolds CH. Free energy perturbation study of octanol/water partition coefficients: comparison with continuum GB/SA calculations. J Phys Chem 1999;103:714–726.
28. David L, Luo R, Gilson MK. Comparison of generalized Born and Poisson models: energetics and dynamics of HIV protease. J Comput Chem 2000;21:295–309.
29. Tsui V, Case DA. Molecular dynamics simulations of nucleic acids with a generalized Born solvation model. J Am Chem Soc 2000;122: 2489–2498.
30. Calimet N, Schaefer M, Simonson T. Protein molecular dynamics with the generalized Born/ACE solvent model. Proteins 2001;45: 144–158.
31. Dominy BN, Brooks CLI. Development of a generalized Born model parametrization for proteins and nucleic acids. J Phys Chem B 1999;103:3765–3773.
32. Bursulaya BD, Brooks CLI. Comparative study of the folding free energy landscape of a three-stranded β-sheet protein with explicit and implicit solvent models. J Phys Chem B 2000;104:12378–12383.
33. Tsui V, Case DA. Theory and applications of the generalized born solvation model in macromolecular simulations. Biopolymers 2001; 56:275–291.
34. Cornell W, Abseher R, Nilges M, Case DA. Continuum solvent molecular dynamics study of flexibility in interleukin-8. J Mol Graph Model 2001;19:139–145.
35. AMBER: Assisted model building with energy refinement. Version 6.0. San Francisco: Department of Pharmaceutical Chemistry, University of California; 1999.
36. Weiser J, Shenkin PS, Still WC. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). J Comput Chem 1999;20:217–230.
37. Gabdoulline RR, Wade RC. Simulation of the diffusional association of barnase and barstar. Biophys J 1997;72:1917–1929.
38. InsightII. Version 2000. San Diego: Molecular Simulations Inc.; 2000.
39. Ibragimova G, Wade RC. Importance of explicit salt ions for protein stability in molecular dynamics simulation. Biophys J 1998;74:2906–2911.
40. Ferguson N, Pires JR, Toepert F, Johson CM, Pan YP, Volkmer-Engert R, Schneider-Mergener J, Daggettt V, Oschkinat H, Fersht A. Using flexible loop mimetics to extend phi-value analysis to secondary structure interactions. Proc Natl Acad Sci 2001;98: 13008–13013.
41. Crane JC, Koepf EK, Kelly JW, Gruebele M. Mapping the transition state of the WW domain beta-sheet. J Mol Biol 2000;298: 283–292.
42. Buckle AM, Henrick K, Fersht AR. Crystal structural analysis of mutations in the hydrophobic cores of barnase. J Mol Biol 1993;234:847–860.
43. Kollman P, Dixon W, Cornell T, Fox T, Chipot C, Polhoille A. In: Willkinson A, Weiner P, van Gunsteren W, editors. Computer simulation of biomolecular systems. Vol. 3. Dordrecht: Kluwer/Escom; 1997. p 83–90.
44. Ferrara P, Apostolakis J, Caflisch A. Evaluation of a fast implicit solvent model for molecular dynamics simulations. Proteins 2002; 46:24–33.
45. Dixon JS. Evaluation of the CASP2 docking section. Proteins 1997;Suppl.1:198–204.
46. Verdecia MA, Bowman ME, Lu KP, Hunter T, Noel JP. Structural basis of phosphoserine-proline recognition by group IV WW domains. Nat Struct Biol 2000;7:639–643.
47. Wintjens R, Wieruszesk J-M, Drobecq H, Rousselot-Pailley P, Buee L, Lippens G, Landrieu I. 1H NMR study on the binding of Pin1 Trp-Trp domain with phosphothreonine peptides. J Biol Chem 2001;276:25150–25156.
48. Schreiber G, Fersht AR. Energetics of protein-protein interactions: analysis of the barnase-barstar interface by single mutations and double mutant cycles. J Mol Biol 1995;248:478–486.