ELSEVIER

Continuum solvent molecular dynamics study of flexibility in interleukin-8

Wendy Cornell, Roger Abseher, Michael Nilges, and David A. Case

Novartis Pharmaceutical Corporation, Summit, NJ 07901, USA
European Molecular Biology Laboratory, 69117 Heidelberg, Germany
Department of Molecular Biology, The Scripps Research Institute, La Jolla, California, USA

Generalized Born continuum solvent methods have been shown to provide a reasonable description of the equilibrium thermodynamics of aqueous solvation in a variety of applications to peptides, proteins, and nucleic acids. Here we study the performance of these methods in molecular dynamics simulations of interleukin-8, comparing nanosecond-length explicit solvent simulations with those using the generalized Born model. In general, the simulations show similar results, although movement away from the initial NMR-determined structure and average fluctuations about the mean are slightly higher for the continuum solvent results. In both simulations, the two helices that are packed on top of the core sheet move closer together, resulting in a structure that more closely resembles the X-ray structure. Principal-component (quasiharmonic) analysis is used to analyze the motions of these helices in both of the simulations and in the NMR ensemble of structures. Prospects for making more general use of continuum solvent models in protein dynamics simulations are discussed.

INTRODUCTION

Studies of proteins and nucleic acids using molecular dynamics simulations have shown considerable progress during the past five years.[1,2] With the use of Particle-Mesh Ewald (PME) to treat long-range electrostatics, we have obtained stable trajectories of macromolecules up to several nanoseconds with structural and dynamic features that are generally in good overall accord with experiment.[3–8] To accurately represent the environment, such simulations generally are carried out in the presence of explicit solvent molecules. The box of solvent molecules must be large enough to minimize electrostatic interactions between periodic images of the solute, generally leaving at least 10 Å between each edge of the box and the closest solute atom. This results in a system with approximately 10,000 water molecules for a protein like interleukin-8, considered here. In addition to increasing calculation time, complications also arise from the need to fully equilibrate these molecules and any counterions in the system. Even with today's powerful computational facilities, these simulations can be lengthy and costly. The electrostatic effects of a solvent of high dielectric, such as water, can be approximated by a continuum electrostatics model.[9–11] This model has been studied extensively, mostly through numerical solutions of electrostatic equations in a multiple dielectric model. The generalized Born (GB) model,[12–14] provides an approximate solution to the solute–solvent electrostatic polarization term (ΔG_{pol}). This model is often able to reproduce the solvation energies given by the Poisson-Boltzmann (PB) continuum solvent model for a variety of biomolecules, without the costly computations of the numerical solutions to Poisson's equation.[15–23] Furthermore, an extension to the GB model to account for salt effects in the linearized Debye-Huckel approximation has been shown to give good agreement with PB results.[18] Although many early applications of this model considered the energetics of a relatively small number of conformations (such as snapshots from explicit solvent simulations),[17,18,24–27] its use as an effective solvation model for MD simulations is a natural extension of this earlier work, and some studies for peptides, proteins, and nucleic acids have appeared.[19,28–30] Here we apply the GB model, using a pairwise descreening approximation introduced by Hawkins et al.,[31] to simulations of interleukin-8 in solution. Interleukin-8 is a noncovalently associated homodimer with a total molecular weight of 17 kDa, with each monomer consisting of 72 amino acid residues. Both X-ray[32] and NMR[33,34] structures have been determined for the wild type protein. A number of X-ray and NMR structures of mutant forms are also available,[35,36] and recently an NMR structure of the wild type protein complexed to a peptide analog of CXCR1 was published.[37] The topology of interleukin-8 is that of two α-helices oriented in anti-parallel fashion on top of a six-stranded β-sheet (see Figure 1). The α-helices are located at the C-terminal end of each monomer while the N-termini are fairly disordered. The dimerization interface consists of a β-sheet interaction between the two monomers and contacts between the helix of each monomer and the three-stranded β-sheet of the other monomer. Each monomer contains two disulfide bonds, Cys7-

Corresponding author: W. Cornell, Novartis Pharmaceutical Corporation, Summit, NJ 07901, USA

E-mail address: wendy.cornell@pharma.novartis.com

Journal of Molecular Graphics and Modelling 19, 136–145, 2001

655 Avenue of the Americas, New York, NY 10010

1093-3263/01/$–see front matter
PII S1093-3263(00)00124-8

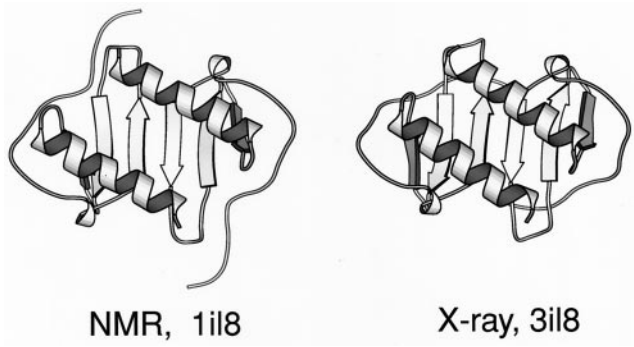

Figure 1. Secondary structures for interleukin-8 as determined by NMR and X-ray crystallography.

Cys-34 and Cys9-Cys50, which link the N-terminal segments to the central core. The N-terminal region has been shown to be important for receptor binding and activation, with residues Glu6-Leu7-Arg8 being especially critical.[38–40] This nonglobular topology is highly stable and native interleukin-8 is structured up to 80 C.[41] Interleukin-8 is a member of the chemokine (chemotactic cytokine) superfamily of proteins, which are involved in the recruitment and activation of immune cells in both normal and disease-related processes. The chemokine family has traditionally been further divided into CXC and CC subfamilies, with the designation referring to a characteristic amino acid pattern of two cysteines that occur at the N-terminal end of each monomer. Interleukin-8 is a member of the CXC subfamily, which has a single amino acid separating the two cysteines. Functionally, the two different subfamilies were distinguished by the target cells upon which they were observed to act, with CXC chemokines acting mainly upon neutrophils and CC chemokines upon monocytes, eosinophils, and lymphocytes. CXC chemokines are associated with acute inflammatory conditions and CC chemokines with chronic inflammatory diseases such as asthma, arthritis, and atherosclerosis. Later studies demonstrated the additional involvement of interleukin-8 in chronic conditions such as arthritis, inflammatory bowel disease, and psoriasis.[42] A third subfamily of chemokines, in which three amino acid residues are present in between the two cysteines, has recently been identified. Currently, only one chemokine belongs to this CX_3C grouping. A recent review summarized the variety of 36 known human chemokines, their associated receptors, and prospects for drug design against this class of targets.[42] All of the 17 mammalian chemokine receptors reported to date belong to the class of seven-transmembrane G-protein coupled receptors.[42] Interleukin-8 binds specifically and with affinities on the order of 1–2 nM to two such receptors on the surface of neutrophils: CXCR1 and CXCR2, also known as interleukin-8 receptors A and B, respectively.[43,44] Interleukin-8 also binds to the promiscuous receptor DARC, found on the surface of erythrocytes, with an affinity of about 20 nM. The NMR structure of the complex between interleukin-8 and the peptide analog of CXCR1 revealed the interaction surface to be primarily hydrophobic and side-chain mediated, involving amino acids Gln8, Ile10, Lys11, Tyr13, Phe17, Phe21, Ile40, Leu43, Arg47, and Leu49 in an N-terminal cleft of interleukin-8.[37] The hydrophobic nature of this cleft is maintained across both the CXC and CC subfamilies. While the specific role of the Glu6-Leu7-Arg8 sequence in IL8 is not evident from this structure of the complex, mutational studies of CXCR1 have identified Arg199 and Arg203 on loop 3 and Asp265 on loop 4 as being important for interleukin-8 binding. The basic residues along the helices of interleukin-8 are the components of the second known recognition site on the protein, associating with the sulfate groups of glycosaminoglycans (GAGs) on the surface of endothelial cells. Although no experimental structure of interleukin-8 in such a complex is currently available, a recent docking study suggested that His18, Lys20, Lys64, Lys67, and Arg68 are the key residues in interleukin-8 that interact with heparin.[45] This initial association serves to drive the localization of the immune response. Although interleukin-8 is a dimer under the conditions employed in the structural studies, the reported dimerization constant is actually four orders of magnitude greater than the association constant of 1–2 nM for the protein with CXCR1 or CXCR2 . The dimerization constant for wild type interleukin-8 has been reported to have a value of about 10–20 mM under physiological conditions of salt, pH, and temperature.[46,47] Recent studies have demonstrated that this value is highly sensitive to specific solution conditions.[48] A number of studies have been aimed at identifying the dimerization state of receptor-bound interleukin-8. Obligatory monomeric mutants of interleukin-8 were able to bind to CXCR1 and CXCR2 and elicit the expected biological response.[48] Similarly, obligatory dimeric structures, which had both chains expressed as a single chain with an intervening linker and disulfide bonds introduced between the two monomers, were also active on the receptors.[49] It is thus unclear whether interleukin-8 binds to its receptors as a monomer or dimer, and local conditions in vivo must play a major role in this equilibrium.

METHODS

Explicit Solvent Simulations

The explicit solvent MD simulations used the Cornell et al. force field[50] and the TIP3 model for water.[51] Long-range electrostatic interactions were handled by the PME method, using the recommended Amber accuracy parameters. This implies a cutoff for the real-space nonbonded interactions of 8 Å, and about a 1 Å grid spacing for the reciprocal space calculations. The calculations began with the NMR mean structure (1il8.pdb), and added solvent molecules such that the minimum distance of a protein atom to the edge of the rectangular box was at least 10 Å. This procedure added 9,697 water molecules. Eight chloride ions were added to yield a net neutral system. The final system had 31,437 atoms. The system was equilibrated over 70 ps, heating to 300K and adjusting the box dimensions to achieve a mean pressure of 1 atm. After this, 1 ns of production simulation in the NVE ensemble was carried out, using an MD time step of 1 fs. The mean total energy was about −80,000 kcal/mol, with an energy drift (decline) of about 40 kcal/mol over the nanosecond trajectory. Because of the large number of degrees of freedom, this corresponds to a drift in mean temperature of less than 1K. Snapshots were collected at 0.1 ps intervals for subsequent analysis.

Implicit Solvent Simulations

The GB model has been discussed extensively in earlier work, and the results presented here use the model previously applied

to MD simulations of nucleic acids.[30] The form used here contains a modification that incorporates a Debye-Huckel term to account for salt effects at low salt concentrations[18]:

$$\Delta G_{pol}^{GB} = -\frac{1}{2}\left(1 - \frac{e^{-\kappa f_{GB}}}{\varepsilon}\right)\sum_{ij}\frac{q_i q_j}{f_{GB}} \qquad (1)$$

where q_i and q_j are atomic partial charges, ε is the solvent dielectric constant, κ is the Debye-Huckel screening parameter, and the double sum runs over all pairs of atoms. f_{GB} depends on the effective Born radius α_i and the distance r_{ij} between atoms:

$$f_{GB} = [r_{ij}^2 + \alpha_{ij}^2 \exp(-D_{ij})]^{1/2}$$

where $\alpha_{ij} = (\alpha_i\alpha_j)^{1/2}$, and $D_{ij} = r_{ij}^2/(2\alpha_{ij})^2$. As $r_{ij} \rightarrow 0$, $f_{GB} \rightarrow \alpha_i$, the effective radius that establishes the self-energy of charges that arises from polarization of the surrounding dielectric medium. The f_{GB} function can be thought of as an interpolation formula that reduces to these self-energy terms at short distances, and to a Debye-Huckel screened Coulomb interaction at large distances. The effective Born radii roughly describe the average distance from a charge to the dielectric boundary, and depend on the positions and volumes of all other atoms in the solute. The original formulation[13] estimated this from a numerical integration procedure, and more recently, several analytical approximations to determining these self energies (and hence the α_i) have been proposed.[14,15,31] Here we adopt the method of Hawkins, Cramer, and Truhlar,[31,52] which uses a pairwise descreening approximation (PDA) to estimate α_i from a sum over atom pairs:

$$\alpha_i^{-1} = \rho_i^{-1} - \sum_{j\neq i} g(r_i, r_j, \rho_i, \rho_j) \qquad (3)$$

Here ρ_i is an intrinsic radius for atom i, and $g()$ is a positive function, so that the effective radius α_i is greater than the intrinsic radius ρ_i. The function g depends upon the positions and sizes of the atoms, but not on their charges. These general characteristics apply not only to the Hawkins et al. model used here,[31,52] but also to related approaches such as ACE[14] and the "fast analytical" method of Qiu et al.[15,19] For the PDA model used here, the explicit form of $g()$ is given by Equation 13 in in an article by Hawkins et al.[31]; this model, including exact derivatives of all terms, has been incorporated into version 6 of the Amber software package,[53] version 4 of Nucleic Acid Builder,[54] and version 3.8 of X-plor (T. Simonson and D.A. Case, unpublished). The final piece of this prescription for approximate electrostatics is to establish an algorithm for computing the intrinsic radii ρ_i. For each atomic sphere here characterized by its Bondi radius R_i, the contribution of all other spheres to dielectric screening is calculated from an analytical formula for two (possibly overlapping) spheres. One complication arises from the overestimation of the Born radius, and thus underestimation of G_{pol}, when these neighboring spheres overlap each other. A linear scaling of the R_i values can approximately account for this effect;[31] here the intrinsic radius ρ_i of atom i, used to compute effective Born radii, becomes

$$\rho_i = S_i(R_i + b_{offset}) \qquad (4)$$

where S_i and b_{offset} are parameters for this particular model. There are clearly many combinations of S_i, R_i, and b_{offset} that could be used. We began with the Bondi set of radii[55] for R_i (based on good experience in earlier numerical Poisson-Boltzmann calculations),[56–60] and adopted S_i values from the tinker software package.[61,62] The linearized Debye-Huckel approximation used here in conjunction with GB has been shown to successfully reproduce the salt dependence of Poisson-Boltzmann calculations. The Debye-Huckel screening parameter κ in Equation 1 is multiplied by 0.73 to account for the overestimation of salt effects, as described earlier.[18] The simulation started from the 1il8 NMR average structure and consisted of 100 ps of equilibration (heating slowly from zero to 300K), followed by 1 ns of production simulation. The production run used Newtonian dynamics, i.e., with no temperature regulation. The time step was 1 fs, and nonbonded interactions were truncated beyond 15 Å. For this fairly short cutoff, the ratio of fluctuations in the total energy to fluctuations in the kinetic energy averaged about 1%, and the mean drift of total energy was about 1 kcal/mol per 0.1 nsec of simulation. As a consequence of this lack of conservation of total energy, the mean temperature increased by about 5K (from 304K to 309K) during the simulation. For longer runs, such a drift would not be acceptable (one would need either longer cutoffs or some sort of temperature regulation), but for the short time period examined here this slight increase in mean temperature should be acceptable.

Structural Aspects of the Simulations

Interleukin-8 forms a tight, symmetric dimer in solution, with a central six-stranded β-sheet with three strands from each monomer. A cartoon of the secondary structure is shown in Figure 1. Two helices (one from the carboxy-terminal end of each monomer) pack onto the sheet at an angle of about 60° to the sheet direction. A primary distinction between the mean NMR structure and that from crystallography lies in the distance between the helices, as shown in Figure 1. The center-to-center distance is about 12 Å in the X-ray structure and about 14 Å in the NMR results.[32] Figure 2 (top) shows the closest Cα–Cα contact (between residues Leu 65 of one monomer and Glu 62 of the other), for both of the experimental structures and for the simulations discussed here. Both GB and explicit water simulations began from the NMR structure, but each moved within about 200 ps toward a structure with a helix–helix distance closer to that seen in the X-ray structure. The fluctuations in this distance are somewhat greater for the GB simulation (discussed below), but the average behavior over the second 500 ps of simulation is very similar in the two runs. This tendency for the two simulations to resemble the X-ray structure more than the NMR ensemble is also seen in Figure 3, which shows root-mean-square superpositions of both simulations onto the X-ray and NMR structures. The top panels show superpositions based on the backbone heavy atoms of the six-stranded β-sheet, and the bottom panels show results for the entire backbone, except for the relatively disordered N-terminal residues. Three principal results are apparent in this figure:

1. All RMS deviations are consistently smaller when comparison is made to the X-ray structure (dotted lines) than when comparisons are made to the NMR structure (solid lines). This is true even for just the residues in the sheet region, indicating that the differences in structures go beyond just the distance between the helices.
2. Fluctuations in the structures are larger for the GB simulations (left panels) than for the explicit solvent simulations (right panels), although the differences are not great.

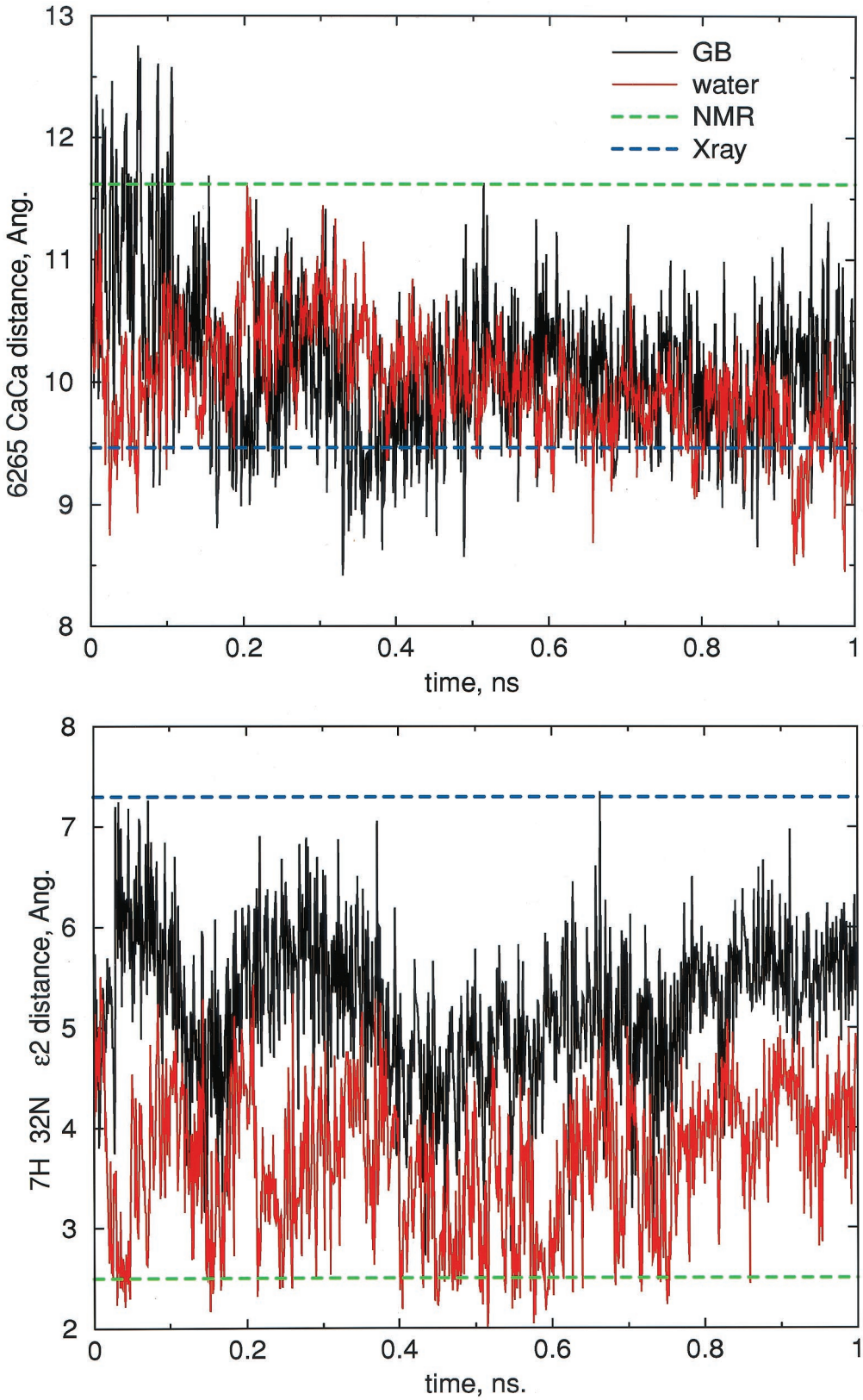

Figure 2. (top): Distances between the Cα atoms of Leu 62 on one monomer and Glu 65 on the other monomer. (bottom): distance from the H atom of Gln 7 and the Nϵ2 atom of His 32. Legend: black = Born; red = explicit solvent; green = the mean NMR structure; and blue = the X-ray result.

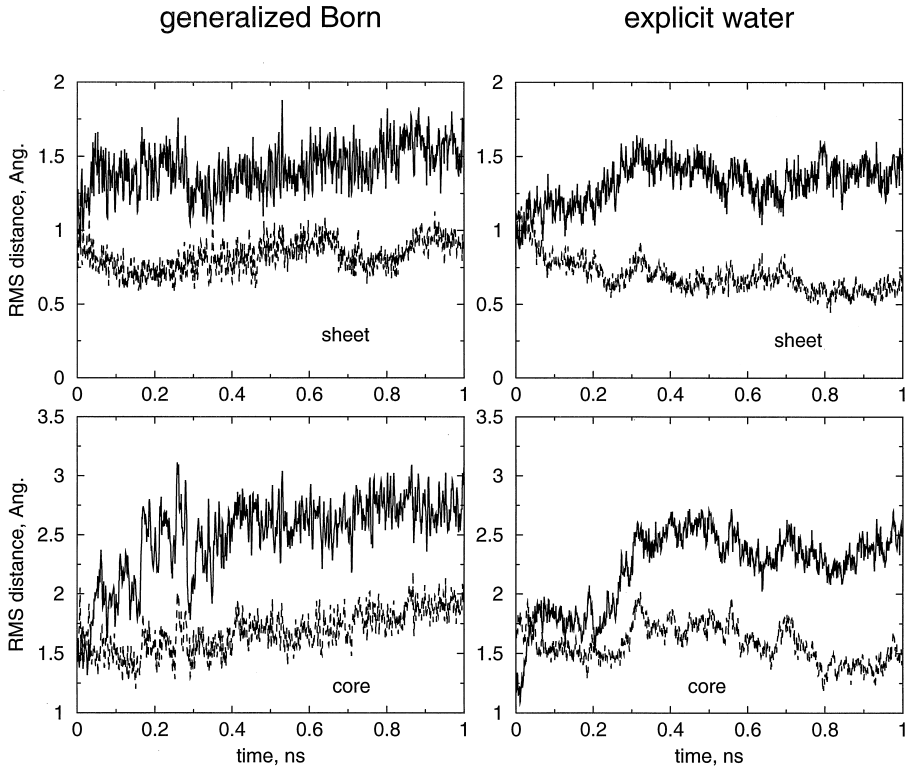

Figure 3. Root-mean-square deviations of calculated structures compared with the NMR mean (solid lines) or the X-ray structure (dashed lines). Top panels: superposition of the β-sheet region (residues 22–28, 36–42, and 46–51 of each monomer); bottom panels: superposition on the core region (residues 9–71 of each monomer), disregarding the disordered N-terminal residues. Left panels: generalized Born results; right panels: explicit water simulations.

3. In all cases the GB and explicit solvent results are quite similar.

A second place where there are significant differences between the NMR and X-ray structures is in the loop containing residues 28–35, and especially in the position of the side chain of His 32.[32] In the mean NMR structure, there is a hydrogen bond between this side chain and the backbone amide of Gln 7, whereas residues 32 and 7 are quite far apart in the X-ray structure. Figure 2 (bottom) shows that each of the simulations interpolates between these limits, leading to structures in which the hydrogen bond is broken, but 7 and 32 are not nearly as far apart as in the crystal structure. Again, that the two types of simulations move away from the starting configuration in similar fashions is encouraging from the point of view of using generalized Born simulations as a replacement for explicit water calculations. We emphasize that these results are probably more relevant to the quality of current force fields than to the question of which experimental structure is more relevant in solution. Indeed, Baldwin et al. point out that some NMR restraints are inconsistent with the X-ray conformation.[32] The key point for this study is that the mean structures obtained from explicit and implicit simulations are quite close to each other, providing initial support for the notion that the latter can replace the former for many purposes. The overall rms deviations of the simulations from either experimental structure are qualitatively in line with those typically seen in other recent simulations, i.e., rms backbone deviations of less than 1 Å for the elements of regular secondary structure, and deviations of

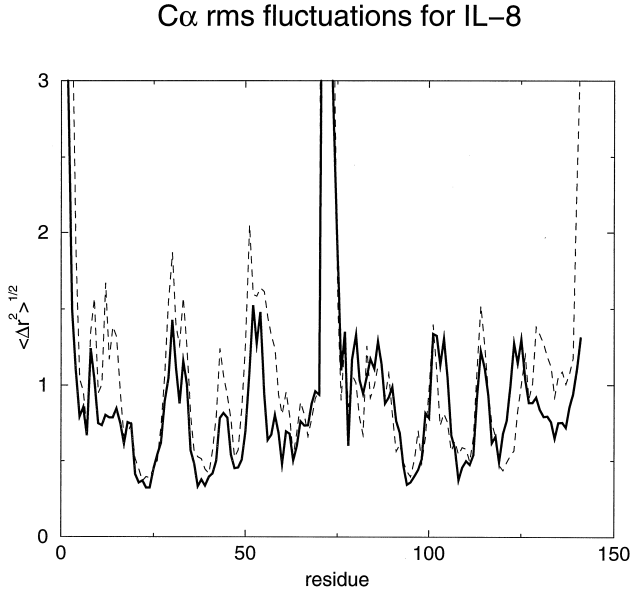

Figure 4. Atomic fluctuations of the Cα atoms of interleukin-8 about the mean. Solid line: explicit water simulation; dashed line: generalized Born simulation. Residues 1–71 are for the first monomer, 72–142 for the second.

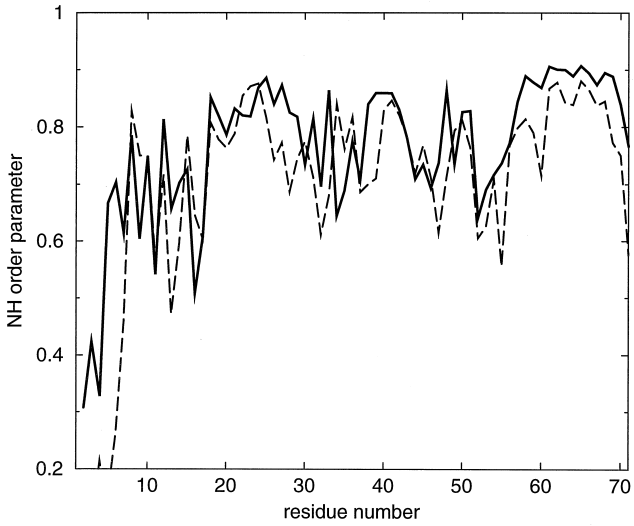

Figure 5. Order parameters (S2) for N-H relaxation. Values are averaged for equivalent positions in the two monomers. Solid line: explicit water simulation; dashed line: GB simulation.

1.5 to 2 Å for the entire backbone (omitting only the disordered N-terminal tail.)

DYNAMIC ASPECTS OF THE SIMULATIONS

Fluctuations about the mean

Figure 4 shows root-mean-square fluctuations about the mean for the Cα atoms in the two simulations. Generally speaking, the comparison between the explicit solvent simulation (solid lines) and the GB simulation (dashed lines) is quite good, although overall fluctuations are slightly larger in the implicit solvent results, which is in accord with results reported earlier for nucleic acids.[30] Similar behavior is seen for the order parameters for nitrogen relaxation shown in Figure 5, calculated by protocols discussed earlier.[63] These values reflect the amplitude of angular motions of the NH bond at picosecond time scales;[63,64] order parameters near unity indicate little motion on this time scale, with increasing motion as the order parameter is reduced. These can be approximately compared with experimental data extracted from NMR relaxation measurements. Figure 5 shows that the general order parameters are nearly the same for the GB and explicit water simulations, although individual residues vary somewhat. Both simulations show order parameters between 0.8 and 0.9 for regular regions of secondary structure, in agreement with experiment.[65] Order parameters near 0.4 for residues 2–4, and near 0.6 for residue 72 are also in agreement with experiment. The remaining observed order parameters are unusually high, with no values below 0.6 for residues 7–71. This is nearly (but not completely) consistent with computed values shown in Figure 5.

Principal Component Analysis

Fluctuations can also be compared with performing a quasiharmonic or principal component analysis, obtained by diagonalizing the second-moment matrix of fluctuations about the mean.[66,67] Figure 6 shows a close agreement between the quasiharmonic frequency distributions derived from GB and from explicit water simulations. In agreement with the rms fluctuations described above, the frequencies for the GB simulation are slightly lower (and hence have slightly larger amplitudes) than those from the explicit solvent simulation. It is well known that most of the fluctuation amplitude arises from a fairly small number of "essential" modes.[66–68] Figure 7

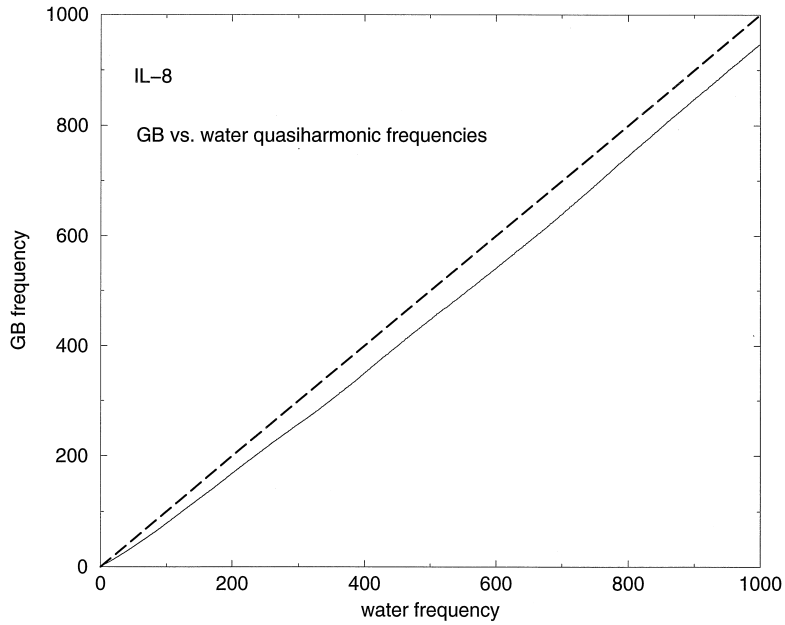

Figure 6. Plot of frequencies of quasiharmonic modes calculated from the GB simulation versus those calculated from the explicit solvent simulation, for modes with frequencies below 1000 cm-1 (solid line). The dashed line indicates the ideal case of identical frequencies.

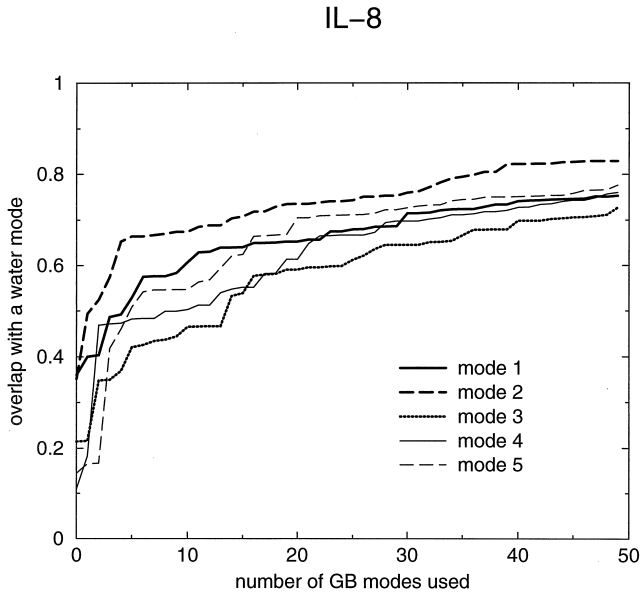

Figure 7. Plot of the overlap of each of the first five quasiharmonic modes calculated from the explicit solvent simulation with modes calculated from the GB simulation. The overlap is computed as the cumulative inner products of the specified water eigenvector with the GB eigenvectors, and it is plotted as a function of the number of GB modes used in the calculation.

shows the overlap of the five highest-amplitude modes from the explicit water simulation, with subspaces derived from various numbers of modes from the GB simulation. Each of the water modes overlaps to greater than 0.7 with a subspace formed from the first 40 GB modes. This is similar to the behavior that would be seen from independent simulations in the same force field,[69] and supports the notion that the nature of fluctuations seen for the GB simulation has a character very similar to that of the explicit water reference simulation.

TIMINGS

The simulations used a modified version of Amber release 6 on a Cray T3E-900 computer. Table 1 lists the performance results of a 1.0-ps simulation of myoglobin, a protein of comparable size to the interleukin-8 dimer (154 residues for myoglobin versus 142 for interleukin-8). These calculations used a multiple-time-step algorithm for molecular dynamics,[70] computing the bond, angle, and dihedral terms every 1 fs; the nonbonded terms with distances less than 8 Å every 2 fs; and the nonbonded terms with distances between 8 and 30 Å every 4 fs. Our implementation of the generalized Born model requires three cycles through nonbonded pairs. In the first cycle, the effective radii are computed from Equation 3; a second loop computes the energies from Equation 1 and the derivatives with respect to the explicit distance dependence; a final loop includes the contributions to the total force from derivatives of Equation 1 with respect to the effective radii. Timings for each of these loops is shown, as well as times for communication among nodes of the effective radii, the coordinates, and the total forces. As Table 1 shows, load balancing among the nodes (which is handled by a simple round-robin scheduling) is quite good. For example, the *loop1* timing on 64 nodes (34.4 s) is 31.9 times faster than on two nodes (1097.1 s), which is nearly perfect speedup. The limitation on scaling to larger numbers of processors comes from the communication steps, which are only 0.07% of the total time for two processors, but rise to 6.0% of the total time at 64 processors. These communication times are much less than are required for explicit water simulations, since the number of atoms is so much smaller here. The overall speedup of 58 on 64 processors is quite good for a molecular dynamics simulation and points the way to increased use of large parallel machines for such calculations. It is worth noting that the current codes are not yet highly optimized; in particular, the treatment of cutoffs and long-range electrostatic interactions could probably be improved. The generalized Born model interactions decay as $\exp(-\kappa r)/\varepsilon r$ (cf. Equation 1), rather than as a bare Coulomb $1/r$ interaction in the explicit simulations. This should allow cutoff procedures to be effective at shorter distances than for explicit simulations, perhaps still using fast-multipole or other acceleration methods for long-range interactions. The myoglobin calculations (using a conservative 30 Å cutoff, compared with 15 Å used for the IL-8 GB simulations) were concerned more with demonstrating the fundamental character of the GB model than with execution speed, but future work will address these issues as well. In this respect, it is worth noting that the IL-8 simulations (run without multiple time-step algorithms) are only about 25% faster than the explicit water simulations: on 16 T3E-900 processors, 105 steps required 59,665 s for the explicit simulations versus

Table 1. Timings for 154-residue protein simulation

N	total	speedup	loop1	gbcomm	loop2	loop3	FRCcoll	RDdist
2	2886.1	2.0	1097.1	0.48	756.4	999.9	1.32	0.24
4	1448.3	3.98	548.8	1.36	378.2	500.3	2.15	0.34
8	729.6	7.91	274.7	1.70	189.1	250.4	3.06	0.45
16	369.1	15.64	137.4	1.66	94.6	125.5	3.33	0.60
32	188.9	30.56	68.8	1.56	47.3	62.9	3.60	0.63
64	98.7	58.48	34.4	1.75	23.6	31.5	3.40	0.82

The *total* column give execution times; *speedup* is relative to a single processor; *loop1*, *loop2* and *loop3* are times for processes defined in the text; *gbcomm* is the time required to communicate the effective radii among nodes; *FRCcoll* is the communication time required to add partial forces together and redistribute them to all nodes; *RDdist* is the communication time required to distribute updated coordinates to all nodes after each MD step. All times are wallclock times, in seconds.

45,022 s for generalized Born. We expect the performance advantage of implicit models to increase with increasing system size and as the implicit computer codes are further optimized.

CONCLUSIONS

As noted in the introduction, implicit solvent models are for the first time becoming fast enough to be used on a routine basis for molecular dynamics simulations. We report our first experiences with protein simulations, complementing earlier molecular dynamics studies with this model for nucleic acids.[30] While it is clear that only tentative conlcusions should be drawn from studies on a single protein, the overall nature of the implicit solvent results for interleukin-8 strongly suggest that they could replace explicit simulations for many purposes. The timing results quoted above show only a modest gain (for a given number of MD steps) for running generalized Born versus explicit solvent simulations, but it should be noted that convergence of simulations should be accelerated with implicit solvation models, so that fewer simulation steps will be needed in many cases. Furthermore, improvements in efficiency (or further simplifications of the method) may make future implicit solvation model simulations less time-consuming. Overall, the results appear promising enough to warrant further investigation, which is currently in progress.

ACKNOWLEDGMENTS

This work was supported in part by NIH grant GM-57513 and by a joint Scripps-Novartis project, SFP-1248.

REFERENCES

1 Leontis, N.B., and SantaLucia, J., Eds. In: *Molecular Modeling of Nucleic Acids*, American Chemical Society, Washington, DC, 1998
2 Auffinger, P., and Westhof, E. Simulations of the molecular dynamics of nucleic acids. *Curr. Opin. Struct. Biol.* 1998, **8**, 227–236
3 York, D.M., Yang, W., Lee, H., Darden, T., and Pedersen, L.G. Toward the accurate modelling of DNA: The importance of long-range electrostatics. *J. Am. Chem. Soc.* 1995, **117**, 5001–5002
4 Cheatham, T.E., III, Miller, J.L., Spector, T.I., Cieplak, P., and Kollman, P.A. Molecular dynamics simulations on nucleic acid systems using the Cornell et al. Force Field and Particle Mesh Ewald Electrostatics. In: *Molecular Modeling of Nucleic Acids*, Leontis, N.B., and SantaLucia, J., Eds., American Chemical Society, Washington, DC,1998, pp. 285–303
5 Beveridge, D.L., Young, M.A., and Sprous, D. Modeling of DNA via molecular dynamics simulations: Structure, bending, and conformational transitions. In: *Molecular Modeling of Nucleic Acids*, Leontis, N.B., and SantaLucia, J., Eds., American Chemical Society, Washington, DC,1998, pp. 260–284
6 MacKerell, A.D., Jr. Observations on the A versus B equilibrium in molecular dynamics simulations of duplex DNA and RNA. In: *Molecular Modeling of Nucleic Acids*, Leontis, N.B., and SantaLucia, J., Eds., American Chemical Society, Washington, DC,1998, pp. 304–311
7 Feig, M., and Pettitt, B.M. Sodium and chlorine ions as part of the dna solvation shell. *Biophys. J.* 1999, **77**, 1769–1781
8 Fox, T., and Kollman, P.A. The application of different solvation and electrostatic models in molecular dynamics simulations of ubiquitin: How well is the X-ray structure "maintained"? *Proteins: Str. Func. Gen.* 1996, **25**, 315–334
9 Tomasi, J., and Persico, M. Molecular interactions in solution: an overview of methods based on continuous distributions of solvent. *Chem. Rev.* 1994, **94**, 2027–2094
10 Honig, B., Sharp, K., and Yang, A.-S. Macroscopic models of aqueous solutions: Biological and chemical applications. *J. Phys. Chem.* 1993, **97**, 1101–1109
11 Cramer, C.J., and Truhlar, D.G. Implicit solvation models: Equilibria, structure, spectra, and dynamics. *Chem. Rev.* 1999, **99**, 2161–2200
12 Constanciel, R., and Contreras, R. *Theor. Chim. Acta* 1984, **65**, 1
13 Still, W.C., Tempczyk, A., Hawley, R.C., and Hendrickson, T. Semianalytical treatement of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* 1990, **112**, 6127–6129
14. Schaefer, M., and Karplus, M. A comprehensive analytical treatment of continuum electrostatics. *J. Phys. Chem.* 1996, **100**, 1578–1599
15 Qiu, D., Shenkin, P.S., Hollinger, F.P., and Still, W.C. 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
16 Edinger, S.R., Cortis, C., Shenkin, P.S., and Friesner, R.A. Solvation free energies of peptides: Comparison of approximate continuum solvation models with accurate solution of the Poisson-Boltzmann equation. *J. Phys. Chem. B* 1997, **101**, 1190–1197
17 Srinivasan, J., Cheatham, T.E., III, Kollman, P., and Case, D.A. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate–DNA helices. *J. Am. Chem. Soc.* 1998, **120**, 9401–9409
18 Srinivasan, J., Trevathan, M.W., Beroza, P., and Case, D.A. Application of a pairwise generalized Born model to proteins and nucleic acids: Inclusion of salt effects. *Theor. Chem. Acc.* 1999, **101**, 426–434
19 Dominy, B.N., and Brooks, C.L., III. Development of a generalized born model parametrization for proteins and nucleic acids. *J. Phys. Chem. B* 1999, **103**, 3765–3773
20 Ghosh, A., Rapp, C.S., and Friesner, R.A. Generalized Born model based on a surface integral formulation. *J. Phys. Chem. B* 1998, **102**, 10983–10990
21 Rapp, C. S., and Friesner, R.A. Prediciton of loop geometries using a generalized Born model of solvation effects. *Proteins* 1999, **35**, 173–183
22 Jayaram, B., Liu, Y., and Beveridge, D.L. A modification of the generalized Born theory for improved estimates of solvation energies and pK shifts. *J. Phys. Chem.* 1998, **109**, 1465–1471
23 Jayaram, B., Sprous, D., and Beveridge, D.L. Solvation free energy of biomacromolecules: Parameters for a modified generalized Born model consistent with the AMBER force field. *J. Phys. Chem. B* 1998, **102**, 9571–9576
24 Jayaram, B., Sprous, D., Young, M.A., and Beveridge, D. Free energy analysis of the conformational prefer-

ences of A and B forms of DNA in solution. *J. Am. Chem. Soc.* 1998, **120**, 10629–10633
25 Srinivasan, J., Miller, J., Kollman, P.A., and Case, D.A. Continuum solvent studies of the stability of RNA hairpin loops and helices. *J. Biomol. Struct. Dyn.* 1998, **16**, 671–682
26 Wagner, F., and Simonson, T. Implicit solvent models: Combining an analytical formulation of continuum electrostatics with simple models of the hydrophobic effect. *J. Computat. Chem.* 1999, **20**, 322–335
27 Vorobjev, Y.N., Almagro, J.C., and Hermans, J. Discrimination between native and intentionally misfolded conformations of proteins: ES/IS, a new method for calculating conformational free energy that uses both dynamics simulations with an explicit solvent and an implicit solvent continuum model. *Proteins* 1998, **32**, 399–413
28 Schaefer, M., Bartels, C., and Karplus, M. Solution conformations and thermodynamics of structured peptides: Molecular dynamics simulation with an implicit solvation model. *J. Mol. Biol.* 1998, **284**, 835
29 Williams, D.J., and Hall, K.B. Unrestrained stochastic dynamics simulations of the UUCG tetraloop using an implicit solvation model. *Biophys. J.* 1999, **76**, 3192–3205
30 Tsui, V., and Case, D.A. Molecular dynamics simulations of nucleic acids using a generalized Born solvation model. *J. Am. Chem. Soc.* 2000, **122**, 2489–2498
31 Hawkins, G.D., Cramer, C.J., and Truhlar, D.G. Pairwise solute descreening of solute charges from a dielectric medium. *Chem. Phys. Lett.* 1995, **246**, 122–129
32 Baldwin, E.T., Weber, I.T., St. Charles, R., Xuan, J.-C., Appella, E., Yamada, M., Matsushima, K., Edwards, B.F.P., Clore, G.M., Gronenborn, A.M., and Wlodawer, A. Crystal structure of interleukin 8: Symbiosis of NMR and crystallography. *Proc. Natl. Acad. Sci. USA* 1991, **88**, 502–506
33 Clore, G.M., Appella, E., Yamada, M., Matsushima, K., and Gronenborn, A.M. Three-dimensional structure of interleukin-8 in solution. *Biochemistry* 1990, **29**, 1689–1696
34 Bonvin, A.M.J.J., and Brnger, A.T. Conformational variability of solution nuclear magnetic resonance structures. *J. Mol. Biol.* 1995, **250**, 80–93
35 Eigenbrot, C., Lowman, H.B., Chee, L., and Artis, D.R. Structural change and receptor binding in a chemokine mutant with rearranged disulfide: X-ray structure of E38C/C50A IL-8 at 2 resolution. *Proteins* 1997, **27**, 556–566
36 Gerber, N., Lowman, H., Artis, D.R., and Eigenbrot, C. Receptor-binding conformation of the "ELR" motif of IL-8: X-ray structure of the L5C/H33C variant at 2.35 resolution. *Proteins* 2000, **38**, 361–367
37 Skelton, N.J., Quan, C., Reilly, D., and Lowman, H. Structure of a CXC chemokine-receptor fragment in complex with interleukin-8. *Structure* 1999, **7**, 157–168
38 Clark-Lewis, I., Schumacher, C., Baggiolini, M., and Moser, B. Structure-activity relationships of interleukin-8 determined using chemically synthesized analogs. Critical role of NH2-terminal residuesand evidence for uncoupling of neutrophil chemotaxis, exocytosis, and receptor binding activities. *J. Biol. Chem.* 1991, **226**, 23128–23134
39 Clark-Lewis, I., DeWald, B., Geiser, T., Moser, B., and Baggiolini, M. Platelet factor 4 binds to interleukin-8 receptors and activates neutrophils when its N-terminus is modified with Glu-Leu-Arg. *Proc. Natl. Acad. Sci. USA* 1993, **90**, 3574–3577
40 Hebert, C.A., Vitangcol, R.V., and Baker, J.B. Scanning mutagenesis of interleukin-8 identifies a cluster of residues required for receptor binding. *J. Biol. Chem.* 1991, **266**, 18989–18994
41 Rajarathnam, K., Sykes, B.D., Dewald, B., Baggiolini, M., and Clark-Lewis, I. Disulfide bridges in interleukin-8 probes using non-natural disulfide analogues: dissociation of roles in structure from function. *Biochemistry* 1999, **38**, 7653–7658
42 Schwarz, M.K., and Wells, T.N.C. Interfering with chemokine networks. The hope for new therapeutics. *Curr. Opin. Chem. Biol.* 1999, **3**, 407–417
43 Holmes, W.E., Lee, J., Kuang, W.-J., Rice, G.C. and Wood, W.I. Structure and functional expression of a human interleukin-8 receptor. *Science* 1991, 1278–1280
44 Murphy, P.M., and Tiffany, H.L. Cloning of a complementary DNA encoding a functional human interleukin-8 receptor. *Science* 1991, **253**, 1280–1283
45 Bitomsky, W., and Wade, R.C. Docking of glycosaminoglycans to heparin-binding proteins: aalidation for AFGF, bFGF, and antithrombin and application to IL-8. *J. Am. Chem. Soc.* 1999, **121**, 3004–3013
46 Burrows, S., Doyle, M., Murphy, K., Franklin, S., White, J., Brooks, I., McNulty, D., Miller, S., Knutson, J., Porter, D., Young, P., and Hensley, P. Determination of the monomer-dimer equilibrium of interleukin-8 reveals it is a monomer at physiological concentrations. *Biochemistry* 1994, **33**, 12741–12745
47 Paolini, J.F., Willard, D., Consler, T., Luther, M., and Krangel, M.S. The chemokines IL-8, monocyte chemoattractant protein-1, and I-309 are monomers at physiologically relevant concentrations. *J. Immunology* 1994, **153**, 2704–2717
48 Lowman, H.B., Fairbrother, W.J., Slagle, P.H., Kabakoff, R., Lui, J., Shire, S., and Herbert, C. Monomeric variants of IL-8: Effects of side chain substitutions and solution conditions upon dimer formation. *Prot. Sci.* 1997, **6**, 598–608
49 Leong, S.R., Lowman, H.B., Liu, J., Shire, S., Deforge, L.E., Gilleve-Castro, B.I., McDowell, R., and Hebert, C.A. IL-8 single-chain homodimers and heterodimers: Interactions with the chemokine receptors CXCR1, CXCR2, and DARC. *Prot. Sci.* 1997, **6**, 609–617
50 Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Jr., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 1995, **117**, 5179–5197
51 Jorgensen, W.L. *J. Chem. Phys.* 1982, **77**, 4156
52 Hawkins, G.D., Cramer, C.J., and Truhlar, D.G. 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
53 Case, D.A., Pearlman, D.A., Caldwell, J.C., III, T.E. Cheatham, Ross, W.S., Simmerling, C.L., Darden, T.A., Merz, K.M., Stanton, R.V., Cheng, A.L., Vincent, J.J., Crowley, M., Tsui, V., Radmer, R.J., Duan, Y., Pitera, J., Massova, I., Seibel, G.L., Singh, U.C., Weiner, P.K.,

and Kollman, P.A. *AMBER 6*. University of California, San Francisco, 1999
54 Macke, T.J., and Case, D.A. Modeling unusual nucleic acid structures. In: *Molecular Modeling of Nucleic Acids*, Leontis, N.B., and SantaLucia, J., Eds. American Chemical Society, Washington, DC,1998. pp. 379–393
55 Bondi, A. *J. Chem. Phys.* 1964, **64**, 441
56 Ösapay, K., Young, W., Bashford, D., Brooks, C.L., III, and Case, D.A. Dielectric continuum models for hydration effects of peptide conformational transitions. *J. Phys. Chem.* 1996, **100**, 2698–2705
57 Bashford, D., Case, D.A., Dalvit, C., Tennant, L., and Wright, P.E. Electrostatic calculation of side-chain pK_a values in myoglobin and comparison with NMR data for histidines. *Biochemistry* 1993, **32**, 8045–8056
58 Chen, J.L., Noodleman, L., Case, D.A., and Bashford, D. Incorporating solvation effects into density functional electronic structure calculations. *J. Phys. Chem.* 1994, **98**, 11059–11068
59 Richardson, W.H., Peng, C., Bashford, D., Noodleman, L., and Case, D.A. Incorporating solvation effects into density functional theory: Calculation of absolute acidities. *Int. J. Quantum Chem.* 1997, **61**, 207–217
60 Demchuk, E., Bashford, D., Gippert, G., and Case, D.A. Thermodynamics of a reverse turn motif. Solvent effects and side-chain packing. *J. Mol. Biol.* 1997, **270**, 305–317
61 Dudek, M.J., and Ponder, J.W. Accurate modeling of the intramolecular electrostatic energy of proteins. *J. Computat. Chem.* 1995, **16**, 791–816
62 http://dasher.wustl.edu/tinker
63 Palmer, A.G., and Case, D.A. Molecular dynamics analysis of NMR relaxation in a zinc-finger peptide. *J. Am. Chem. Soc.* 1992, **114**, 9059–9067
64 Lipari, G., and Szabo, A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. I. Theory and range of validity. *J. Am. Chem. Soc.* 1982, **104**, 4546–4559
65 Grasberger, B.L., Gronenborn, A.M., and Clore, G.M. Analysis of the backbone dynamics of interleukin-8 by 15N relaxation measurements. *J. Mol. Biol.* 1993, **230**, 364–372
66 Karplus, M., and Kushick, J.N. Method for estimating the configurational entropy of macromolecules. *Macromolecules* 1981, **14**, 325–332
67 Amadei, A., Linssen, A.B.M., and Berendsen, H.J.C. Essential dynamics of proteins. *Proteins* 1993, **17**, 412–425
68 Case, D.A. Normal mode analysis of protein dynamics. *Curr. Opin. Struct. Biol.* 1994, **4**, 285–290
69 Amadei, A., Ceruso, M.A., and Di Nola, A. On the convergence of the conformational coordinates basis set obtained by the essential dynamics analysis of proteins' molecular dynamics simulations. *Proteins* 1999, **36**, 419–435
70 Humphreys, D.D., Friesner, R.A., and Berne, B.J. Multiple-time-step molecular dynamics algorithm for macromolecules. *J. Phys. Chem.* 1994, **98**, 6885–6892