A Comparison of Structural and Dynamic Properties of Different Simulation Methods Applied to SH3

D. M. F. van Aalten,* A. Amadei,‡ R. Bywater,§ J. B. C. Findlay,* H. J. C. Berendsen,‡ C. Sander,¶ and P. F. W. Stouten‖

*Department of Biochemistry and Molecular Biology, University of Leeds, Leeds LS2 9JT, England; ‡Department of Biophysical Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, the Netherlands; §Biostructure Department, Novo Nordisk A/S, DK-2880, Bagsvaerd, Denmark; ¶Biocomputing, EMBL, 69117 Heidelberg, Germany; and ‖DuPont Merck Pharmaceutical Company, Experimental Station E500/3212 Wilmington, Delaware 19880-0353 USA

ABSTRACT The dynamic and static properties of molecular dynamics simulations using various methods for treating solvent were compared. The SH3 protein domain was chosen as a test case because of its small size and high surface-to-volume ratio. The simulations were analyzed in structural terms by examining crystal packing, distribution of polar residues, and conservation of secondary structure. In addition, the "essential dynamics" method was applied to compare each of the molecular dynamics trajectories with a full solvent simulation. This method proved to be a powerful tool for the comparison of large concerted atomic motions in SH3. It identified methods of simulation that yielded significantly different dynamic properties compared to the full solvent simulation. Simulating SH3 using the stochastic dynamics algorithm with a vacuum (reduced charge) force field produced properties close to those of the full solvent simulation. The application of a recently described solvation term did not improve the dynamic properties. The large concerted atomic motions in the full solvent simulation as revealed by the essential dynamics method were analyzed for possible biological implications. Two loops, which have been shown to be involved in ligand binding, were seen to move in concert to open and close the ligand-binding site.

INTRODUCTION

A large effort is being made to develop simulation methods that produce physically realistic results but do not require much computer time. Traditional methods for simulating proteins in an aqueous environment employ explicit water, which is costly, as a large part of the energy evaluations involve solvent-solvent interactions that are of little interest. To account for the effects of solvent, several methods have been developed that use potentials of mean force, also known as solvation terms. Recently a fast method based on excluded volumes rather than the more traditional accessible surface area approaches was reported and has been tested by stochastic dynamics (SD) (van Gunsteren and Berendsen, 1988) simulations of bovine pancreatic trypsin inhibitor (BPTI) (Stouten et al., 1993). Addition of this term to the standard GROMOS vacuum force field significantly improved physical properties that can be derived from a molecular dynamics (MD) trajectory, as compared to the BPTI crystal structure and a full solvent simulation (Stouten et al., 1993). However, the performance of this solvation term was assessed on the basis of 1) static, structural properties of the final and averaged MD structures, such as secondary structure, hydrogen bonds, and distribution of polar residues; 2) the behavior of one, relatively rigid, protein; and 3) a comparison with only two other simulation methods, vacuum SD and explicit solvent MD. In this work we address all three issues.

1) In addition to the conventional analysis of the final MD structures, we apply the essential dynamics analysis technique (Amadei et al., 1993). This method can reveal large concerted atomic motions that characterize the overall motion of the protein. Analyses of lysozyme (Amadei et al., 1993) and thermolysin (van Aalten et al., 1995) trajectories have shown that these motions can be linked to the biological funtion of these proteins. Simulation methods can only be regarded as equivalent if they produce comparable essential motions.

2) Simulations of the SH3 protein domain (Fig. 1) were carried out. This protein was chosen because due to its high surface-to-volume ratio, its behavior in simulations is expected to be strongly influenced by the solvent models used. In addition, SH3 does not contain any disulfide bridges, whereas BPTI has three, making it more rigid. SH3 possesses a ligand-binding site, probably constructed by loops that undergo conformational changes upon binding.

3) To carry out a more rigorous assessment of the performance of the solvation term as well as the effect of varying other simulation conditions, we have carried out simulations using a large set of distinct protocols: several combinations of full/reduced charges; SD and MD; presence or absence of crystallographic water; solvation term versus explicit water versus no solvent treatment. Comparison of two thermolysin trajectories generated with molecular dynamics, using a reduced-charge force field (all residues having zero net charge but considerable dipoles) for one simulation and a full-charge force field for the other, has revealed that the essential motions in the two trajectories were comparable (van Aalten et al., 1995). Stouten et al.

Received for publication 8 August 1995 and in final form 6 November 1995.

Address reprint requests to Dr. D. M. F. van Aalten, Department of Biochemistry and Molecular Biology, University of Leeds, Leeds LS2 9JT, England. Tel.: ++44-113-2333117; Fax: ++44-113-2333167; E-mail: aalten@bmbpcu01.leeds.ac.uk.

0006-3495/96/02/684/00 $2.00

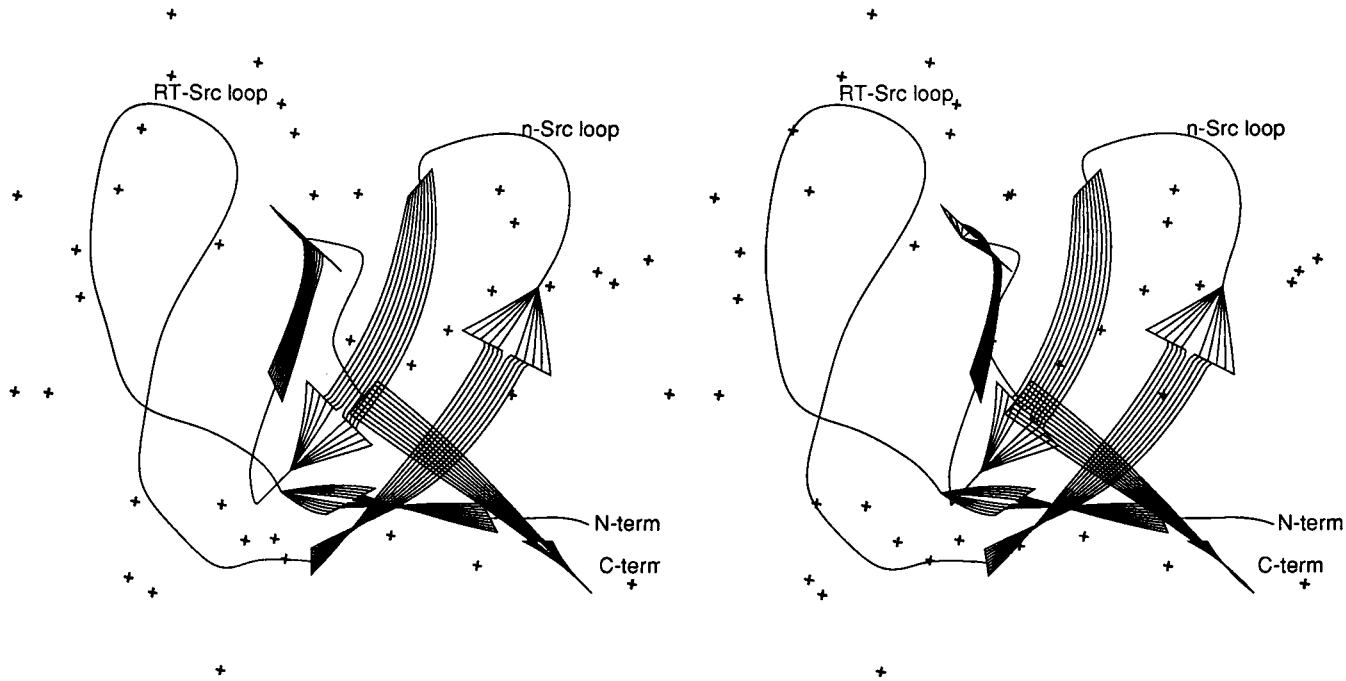

FIGURE 1 Stereo drawing of the SH3 crystal structure. The termini and the RT-Src and n-Src loops are labeled. The crosses represent crystallographic water molecules.

(1993) found that SD with reduced charges and the solvation term performed well as assessed by static structural criteria. Although in some respects it yielded more realistic physical properties than explicit water simulations, the latter had been carried out with standard GROMOS water-carbon van der Waals parameters that were subsequently found to be incorrect (van Buuren et al., 1993). We compare our results with comparison studies described earlier (Stouten et al., 1993; van Aalten et al., 1995; Yun-yu et al., 1988).

The main aim of the work described here is to identify the method or methods that are best suited to replacing explicit water simulations as measured by the static and dynamic properties resulting from simulations of different proteins.

MATERIALS AND METHODS

The essential dynamics method has been extensively described elsewhere (Amadei et al., 1993; van Aalten et al., 1995) and is similar to the multidimensional trajectory fitting first described by Garcia (1992). In short, it involves the construction of a covariance matrix from atomic displacements. A set of eigenvalues and eigenvectors is calculated from this covariance matrix. These eigenvectors correspond to directions in configurational space representing concerted motions of atoms in Cartesian space. For convenient analysis, the eigenvectors are ordered by decreasing eigenvalue. The central hypothesis of the essential dynamics method is that only the directions indicated by eigenvectors with sufficiently high eigenvalues are important for the description of the overall biological motion of the protein. In fact, these eigenvectors span a subspace of the hyperspace in which all "essential" motion takes place. The motion along the rest of the eigenvectors, which are orthogonal to this plane (subspace) are merely small, nonessential, gaussian fluctuations ("near constraints") (Amadei et al., 1993).

If two trajectories of a protein are concatenated and subsequently analyzed by the essential dynamics method, the resulting eigenvectors reveal differences and similarities in equilibrium structure and dynamics between two trajectories (van Aalten et al., 1995). The eigenvectors describe motions that are present in both trajectories. By studying the fluctuation and average structure along these eigenvectors in the separate trajectories, differences in dynamic and structural behavior can be revealed. We use this approach to demonstrate the differences in structure and dynamics caused by the application of different solvent models.

The implementation and application of the solvation term has been described in detail elsewhere (Stouten et al., 1993). It involves the addition of a solvation free energy function to the GROMOS (van Gunsteren and Berendsen, 1987) force field:

$$E_{\text{solv}}(i) = \text{Solpar}(i)\sum_{i\neq j}\text{Vol}(j)\exp(-r_{ij}^2/2\sigma^2),$$

where

$$\sum_{i\neq j}\text{Vol}(j)\exp(-r_{ij}^2/2\sigma^2)$$

represents the occupied volume around atom i, Solpar is an atom-specific solvation parameter, Vol(j) the fragmental volume of atom j, r_{ij} the distance between atoms i and j, and σ is a constant with a value of 3.5 Å. This approximates the effect of solvent without including explicit water molecules and has been shown to yield significantly better physical properties than vacuum simulations for BPTI (Stouten et al., 1993).

All simulations were carried out using the GROMOS program suite (van Gunsteren and Berendsen, 1987) running on Silicon Graphics R8000 CPUs. Structure manipulations and protein trajectory visualizations were carried out using the WHAT IF (Vriend, 1990) modeling program. Secondary structures and hydrogen bonds were analyzed with DSSP (Kabsch and Sander, 1983). Analysis of exposed and buried polar surfaces was performed with the POL DIAG88 program suite (Baumann et al., 1989). Residue packing properties were investigated using the WHAT IF Quality Control tool (Vriend and Sander, 1993).

The simulations were started from the SH3 crystal structure (Musacchio et al., 1992). Eight different methods of simulation were tested (Table 1).

TABLE 1 List of abbreviations used to indicate the different simulation methods

MDVACR	MD, no crystal waters, reduced-charge force field
MDCRYR	MD with crystal waters, reduced-charge force field
SDVACR	SD, no crystal waters, reduced-charge force field
SDCRYR	SD with crystal waters, reduced-charge force field
SDVACRS	SD, no crystal waters, reduced-charge force field, solvation term
SDCRYRS	SD with crystal waters, reduced-charge force field, solvation term
SDCRYFS	SD with crystal waters, full-charge force field, solvation term
MDWATF	MD with crystal waters, full-charge force field, additional solvent molecules and periodic boundary conditions

In the case of MDWATF, the protein was placed in a truncated octahedron filled with water molecules taken from an equilibrated liquid configuration (Berendsen et al., 1981). The net protein charge of +1.0 was compensated for by replacing the water molecule at the highest electrostatic potential by a chloride ion.

All simulations were preceded by a steepest descents energy minimization (EM) until no significant energy change could be detected. Relevant EM parameters are listed in Table 2.

The energy-minimized structures were subsequently subjected to a start-up MD/SD simulation using a "heat-up" protocol detailed in Table 2. Several parameters were slowly changed over a period of 25 ps. Velocities were taken from a Maxwellian distribution at the beginning of each of the 25 ps. The final parameters of the "heat-up" protocol were used for a further 75-ps equilibration. Subsequently, 400-ps data collection runs were started, on which all the analyses described below were performed.

TABLE 2 Parameters used for EM/MD/SD

EM parameters	
Updating of pair list	Every 10 steps
Nonbonded cut-off radius	8 Å
Long range coulomb cut-off radius	10 Å
Bond lengths	Restrained by SHAKE
SHAKE relative geometrical tolerance	0.001
MD/SD parameters for "heatup" and continuation run	
Updating of pair list	Every 10 steps
Bond lengths	Restrained by SHAKE
SHAKE relative geometrical tolerance	0.001
Time step	1–2.0 fs
Protein atom restraining force	9000–0 kJ mol^{-1} nm^{-2}
Nonbonded cut-off radius	7–8 Å
Long-range coulomb cut-off radius	7–10 Å
Temperature	0–300 K
Temperature coupling constant (solute)	0.01–0.1 ps
Temperature coupling constant (solvent)	0.01–0.1 ps
Pressure coupling constant (only with MDWATF)	0.05–0.5 ps
SD friction constant	91 ps^{-1}
MDWATF isothermal compressibility	$4.6 * 10^{-10}$ m^2 N^{-1}
MDWATF pressure	1 atm

Ranges apply to the heat-up protocol; final parameters of the heat-up protocol were used during the continuation MD/SD runs.

RESULTS AND DISCUSSION

Structural analysis

The root mean square (rms) deviation of the C-α atom positions of the MD structures as compared to the SH3 crystal structure (Musacchio et al., 1992) is shown in Fig. 2. Most simulations stabilized during the equilibration period, with the exception of SDVACRS, SDCRYRS, and SDCRYFS. Analyses of equilibrium properties such as positional rms fluctuations and hydrogen bond occurrence will be unreliable for these simulations. Table 3 contains values for the average rms deviation and the mean square fluctuation of this rms deviation in the trajectories.

Hydrogen bonds were analyzed in the eight trajectories. After each picosecond, the total number of hydrogen bonds was calculated using DSSP (Kabsch and Sander, 1983). The averages taken over the whole trajectories are listed in Table 3. The relatively low number of hydrogen bonds in the crystal structure may be caused by a few hydrogen bonds falling just outside the hydrogen bond criteria of DSSP. Calculating the number of hydrogen bonds over an ensemble of structures seems to lead to a small increase. The simulations in vacuum without the solvation term result in a relatively high number of hydrogen bonds. This is probably caused by side chains folding back onto the protein because of the lack of solvation.

The secondary structure assignments calculated as averages over the trajectories were compared to that of the crystal structure. Table 3 lists secondary structure indentity percentages. Application of a stochastic term and friction in vacuum molecular dynamics seems to result in proper conservation of the secondary structure (SDVACR/SDCRYR). Application of the solvation term results in loss of secondary structure.

The WHAT IF option QUALTY generates a quality control index. This number indicates how well the packing of residue types matches packing arrangements in well-refined protein crystal structures (Vriend and Sander, 1993). The higher the quality control index, the better the packing statistics match the average database packing properties. Indices lower than approximately −1.2 indicate a probable error in the structure. Table 3 lists the quality control indices calculated as averages over the trajectories. The SH3 crystal structure gives the highest score. The vacuum simulations (MDVACR, MDCRYR, SDVACR, SDCRYR) and MDWATF also give relatively high scores. Inclusion of the solvation term seems to generate deviations from normal crystal packing properties.

POL DIAG88 (Baumann et al., 1989) was used to analyze polar surface statistics calculated as averages over the trajectories (Table 3). The evaluated properties for all eight simulations fell within the boundaries calculated from a database containing 150 proteins (Baumann et al., 1989). MDWATF is close to the mean for all properties analyzed. SDCRYFS shows the largest deviations.

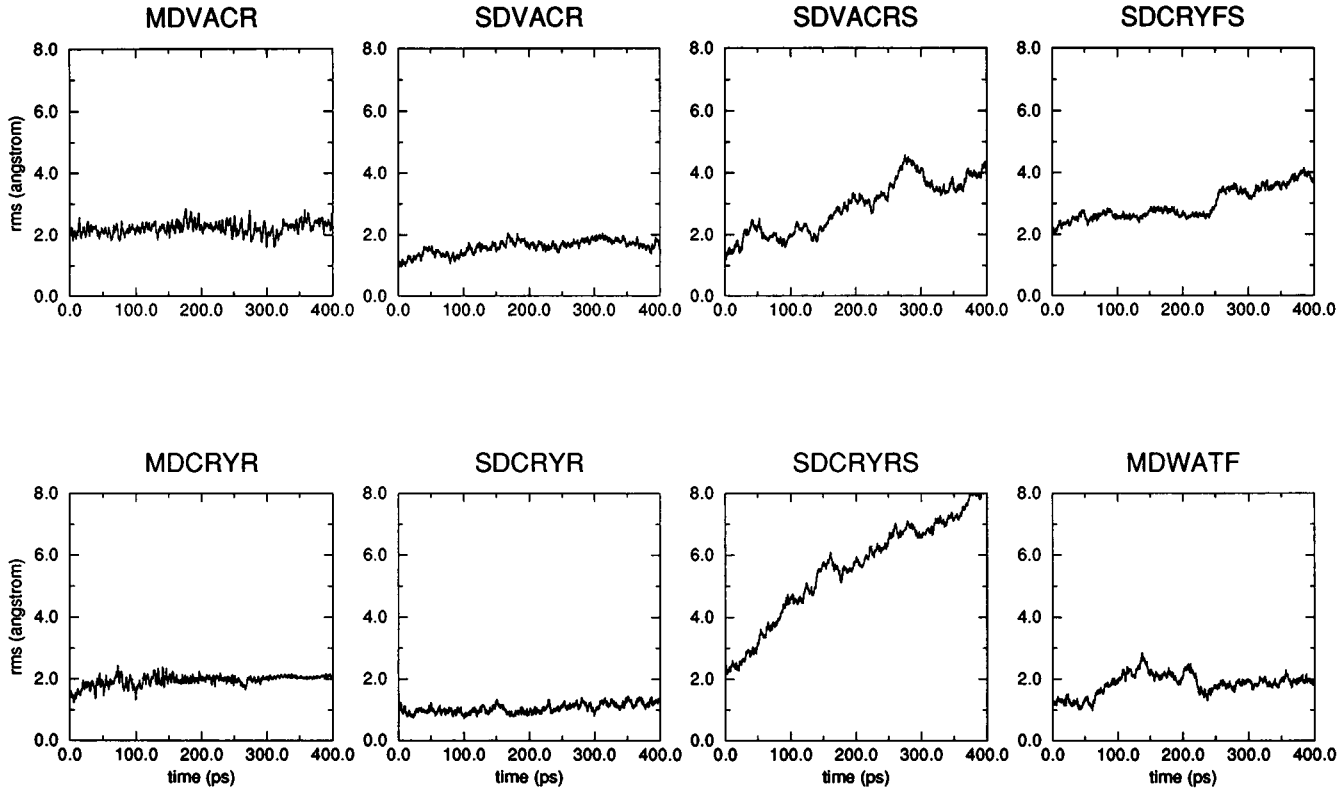

FIGURE 2 2Rms deviation of the C-α atoms in all simulations with respect to the SH3 crystal structure.

The crystallographic B-factors can be compared to the positional atomic fluctuations as calculated from the trajectories. The correlation coefficients resulting from these comparisons for the backbone atoms in the protein core are listed in Table 3. The large deviation from crystallographic B-factors observed for SDVACRS, SDCRYRS, and SDCRYFS is probably due to their increasing deviation from the starting structure (see also Fig. 2).

The protein-accessible surface properties were analyzed in two ways. The total accessible surface area and the radius of gyration were calculated as averages over each trajectory (Table 3). Simulating with the solvation term leads to a significant increase in total accessible surface area and protein size. This may imply that the solvation term is too strong, that is, atoms have too large a preference for being in vacuum space. To determine whether there is a special type of atom/residue with such a preference, the residue accessibilities calculated from the SH3 crystal structure and those averaged over the trajectories were compared (correlation coefficients listed in Table 3). All simulations, except those with the solvation term, show good correlation coefficients. Taken together with the POL DIAG88 analysis, which showed no large deviation of internal/external polar (side chain) fractions for SDVACRS and SDCRYRS, this means that the deviation in total accessible surface area is probably due to an error in the solvation term that acts on all atoms.

Essential dynamics of MDWATF

A covariance matrix was constructed using the 400-ps MDWATF trajectory. Because we are mainly interested in backbone conformational changes, the covariance matrix was constructed using only the C-α coordinates, resulting in a matrix of dimension 171 (3 $\times$ 57 C-α atoms). This has been shown to be a reasonable approximation of the essential space in the all-atom space (Amadei et al., 1993; van Aalten et al., 1995). Fig. 3 shows a plot of the first 50 eigenvalues obtained from the MDWATF covariance matrix, i.e., the 50 largest eigenvalues. The figure shows that there are only a few eigenvectors with large eigenvalues. The central essential dynamics hypothesis is that only these eigenvectors are "essential," i.e., the motions along such eigenvectors are responsible for the large, biologically relevant conformational changes in proteins. The eigenvector components (not shown) show that a few regions of the protein appear to be involved in the motion along the essential eigenvectors: the N- and C-termini, a region around residue 15, and a region around residue 33. These residues are in the tips of the RT-Src and n-Src loops (Noble et al., 1993), respectively (Fig. 1). Recently these loops were identified as part of the ligand-binding site of SH3 (Musacchio et al., 1994; Yu et al., 1994a,b). The eigenvectors show that these regions move in concert. This motion can be visualized by projecting all trajectory frames onto a specific eigenvector (Amadei et al., 1993; van Aalten et al., 1995). A new trajectory is obtained that shows the motion in the direction defined by

TABLE 3 Data from the conventional analyses, averaged over each trajectory

	MD VACR	MD CRYR	SD VACR	SD CRYR	SD VACRS	SD CRYRS	SD CRYFS	MD WATF	PDB SH3
C-α positional deviations									
Average rms (Å)	2.22	1.94	1.61	1.05	2.90	5.56	2.96	1.84	—
Fluctuation in rms (Å^2)	0.20	0.17	0.22	0.15	0.87	1.63	0.50	0.35	—
Hydrogen bonds									
No. of H-bonds	47	45	47	45	44	41	38	42	38
Secondary structure									
% identity with SH3 PDB	85	91	91	92	82	77	62	90	—
Crystal packing									
WHAT IF QUALITY index	−1.20	−1.27	−1.22	−1.18	−1.69	−2.04	−1.95	−1.29	−0.68
POL_DIAG88 analysis									
INTPOL (0.158, 0.172, 0.190)	0.182	0.178	0.178	0.174	0.168	0.170	0.182	0.169	0.173
OUTPOL (0.147, 0.172, 0.204)	0.150	0.155	0.155	0.162	0.170	0.168	0.152	0.169	0.168
SCINT (0.074, 0.098, 0.118)	0.112	0.109	0.112	0.105	0.088	0.089	0.118	0.091	0.102
SCOUT (0.115, 0.143, 0.183)	0.122	0.125	0.122	0.131	0.140	0.136	0.115	0.144	0.140
Atomic fluctuations									
Corcof bft/fluc	0.79	0.29	0.34	0.52	0.24	−0.01	0.04	0.49	—
Accessibility									
Final tot. acc (Å^2)	3833	3876	3864	3886	4849	5299	4229	4185	3931
Corcof res. acc.	0.91	0.93	0.93	0.97	0.86	0.84	0.91	0.95	—
Protein size									
Radius of gyration (Å)	1.02	1.02	1.02	1.01	1.12	1.25	1.09	1.04	1.01

Simulations are indicated by abbreviations listed in Table 1. SH3 PDB, SH3 crystal structure. *Hydrogen bonds:* total number of hydrogen bonds. *POL_DIAG88 analysis:* the lowest, mean and highest values derived from a database of proteins (see text) are listed between brackets for all parameters. INTPOL, total buried polar fraction; OUTPOL, total exposed polar fraction; SCINT, side chain buried polar fraction; SCOUT, side chain exposed fraction. *Automic Fluctuations:* correlation coefficients between crystallographic B-factors and atomic postional rms fluctuations of backbone atoms in the core of the protein. *Accessibility and protein size:* the final total accessible surface, the radius of gyration and the correlation coefficients between residue accessibilities in the MD structures and the SH3 crystal structure are listed.

the eigenvector. We have studied the motions along the first four eigenvectors. The motion along the first eigenvector (the motion with the largest fluctuation) is displayed in Fig. 4. All four eigenvector motions were comparable: the RT-Src and n-Src loops move in concert with the termini. In some respects this is similar to the eigenvector motions observed in lysozyme (Amadei et al., 1993) and thermolysin (van Aalten et al., 1995) in that the residues involved in binding of the substrate (the RT-Src and n-Src loops) are highly flexible and move in concert.

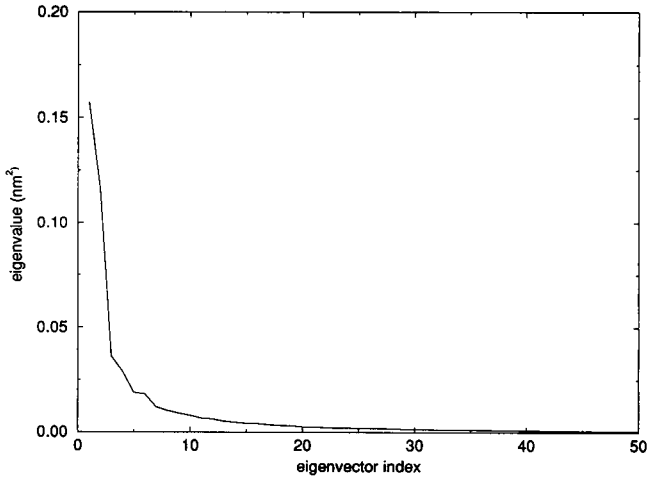

FIGURE 3 Plot of the 50 largest eigenvalues calculated from the MDWATF C-α covariance matrix.

It has been shown that the eigenvectors for which the motion has a gaussian positional distribution in combination with a small eigenvalue are not essential for describing the overall motion of the protein and are considered to be "near constraints" (Amadei et al., 1993; van Aalten et al., 1995). The positional distributions of the displacements along the eigenvectors were calculated and compared to ideal gaussians, derived from the eigenvalues. The correlation coefficients of the comparisons (Fig. 5) show that after approximately the eighth eigenvector the position distributions of the motions along the eigenvectors start to become gaussian. From the eighth vector onward these correlation coefficients are higher than 0.95, indicating an approximately 7-dimensional essential space. Interestingly, the dimension of the essential

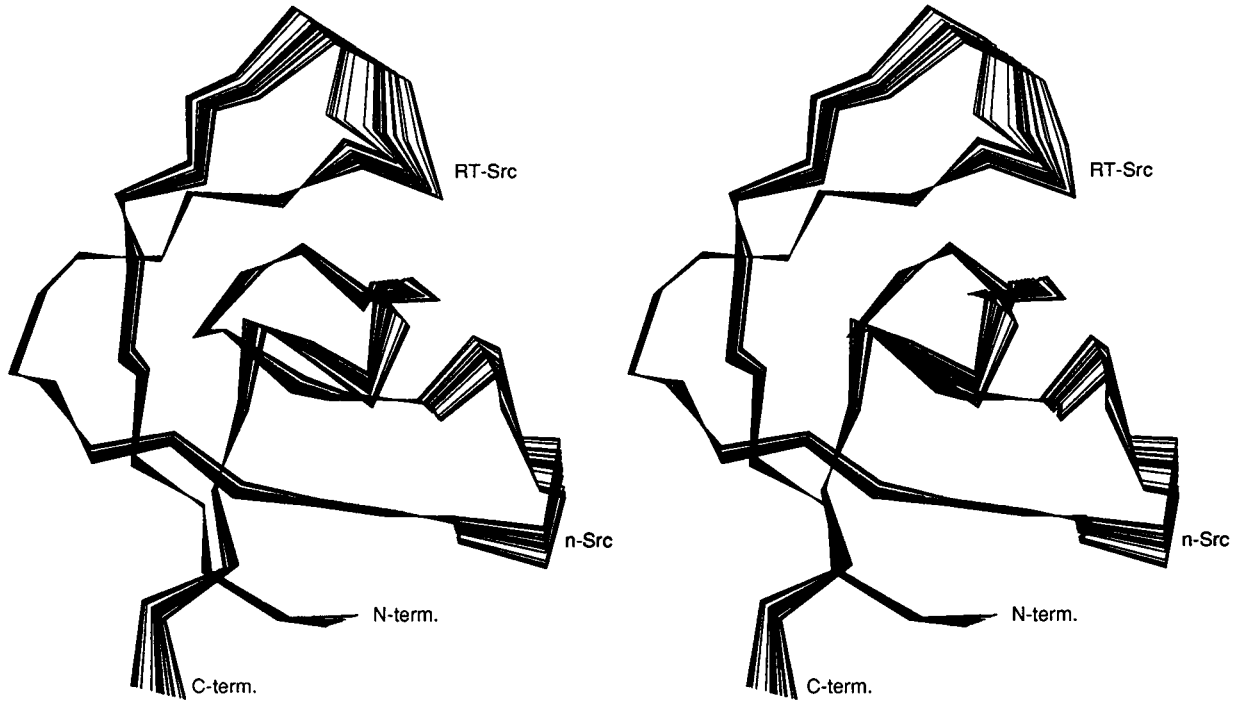

FIGURE 4 Tenwty-five frames taken at equally spaced intervals from the motion along the first eigenvector of the MDWATF C-α covariance matrix.

space is smaller than the dimensions found for lysozyme (Amadei et al., 1993) and thermolysin (van Aalten et al., 1995) (each approximately 10–15 essential eigenvectors). Although SH3 is a considerably smaller protein, this difference in essential space dimension may also be explained by considering the function of the proteins. Lysozyme and thermolysin are both enzymes, that is, they bind substrate, catalyze a reaction, and subsequently release the products. SH3, in contrast, is a binding protein and for this function only a few essential eigenvectors may be enough.

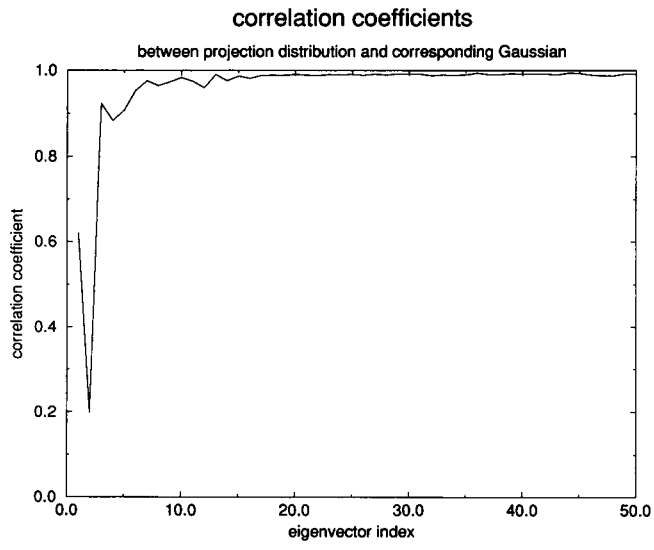

FIGURE 5 Correlation coefficients between the sampled distributions and ideal Gaussian distributions for the first 50 MDWATF eigenvectors.

Essential dynamics analysis of the combined trajectories

As mentioned in Materials and Methods, the essential dynamics method can reveal differences in dynamics and equilibrium structures between two trajectories (van Aalten et al., 1995). Because we consider MDWATF to be the most realistic simulation (Table 3), we compared the dynamics of MDWATF to the other seven trajectories. Each of the trajectories was concatenated with the MDWATF trajectory, and we applied the essential dynamics analysis to the resulting seven trajectories. These trajectories will be indicated as MDVACR+MDWATF through SDCRYFS+MDWATF. The eigenvectors with large eigenvalues calculated from the covariance matrices constructed from these trajectories indicate directions that may have a large eigenvalue because of differences in average projections (equilibrium structures), i.e., at the point from passing from one trajectory to the other a large displacement is observed. On the other hand, large eigenvalues may be caused by the fact that fluctuations in the directions described by the eigenvector are large in both trajectories. The average projections of the single (MDVACR through SDCRYFS) trajectories onto the eigenvector sets calculated from the combined trajectories are shown in Fig. 6. If there are differences in equilibrium structure between the single (MDVACR through SDCRYFS) trajectories and the MDWATF trajectory, these average projections will show a

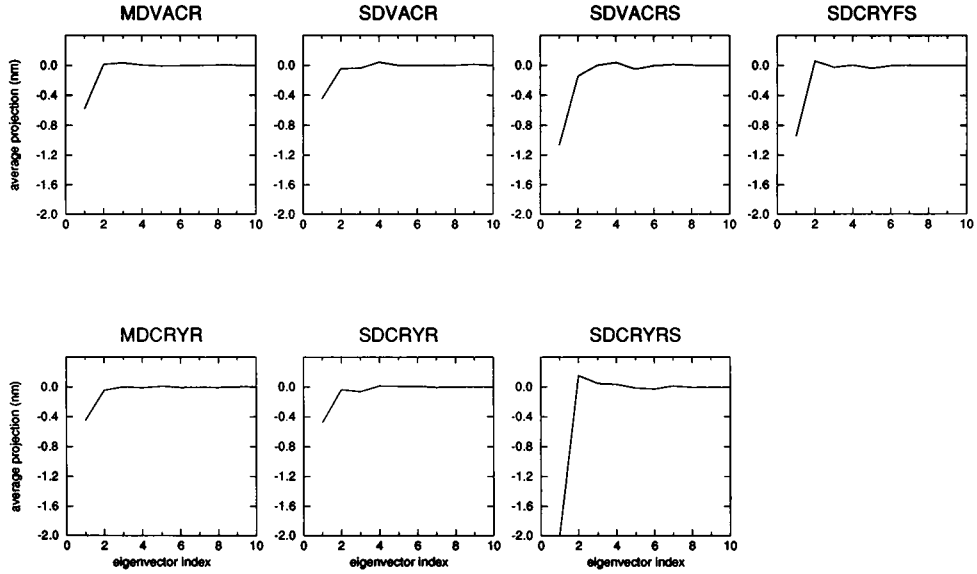

FIGURE 6 Average projection of the single MDVACR through SDCRYFS trajectories onto the first 20 eigenvectors calculated from the MDVACR+MDWATF through SDCRYFS+MDWATF trajectories.

large deviation from zero (van Aalten et al., 1995). SDVACRS, SDCRYRS, and SDCRYFS are examples of this. Fig. 6 shows that the equilibrium structures for these simulations deviate significantly from the MDWATF equilibrium structure along the direction indicated by the first eigenvector. To investigate the meaning of this in 3-dimensional space, the frame from the SDCRYRS trajectory and the frame from the MDWATF that had the largest difference in structure were both projected over the first eigenvector calculated from the SDCRYFS+MDWATF trajectory and superimposed for visual inspection. The result is shown in Fig. 7. Application of stochastic dynamics in combination with a solvation term leads to significant unfolding of various loops, which causes the large increase in total accessible surface (Table 3).

Although there are large deviations for SDVACRS, SDCRYRS, and SDCRYFS, this is not true for all of the simulations. In fact, the deviations for SDVACR are rela-

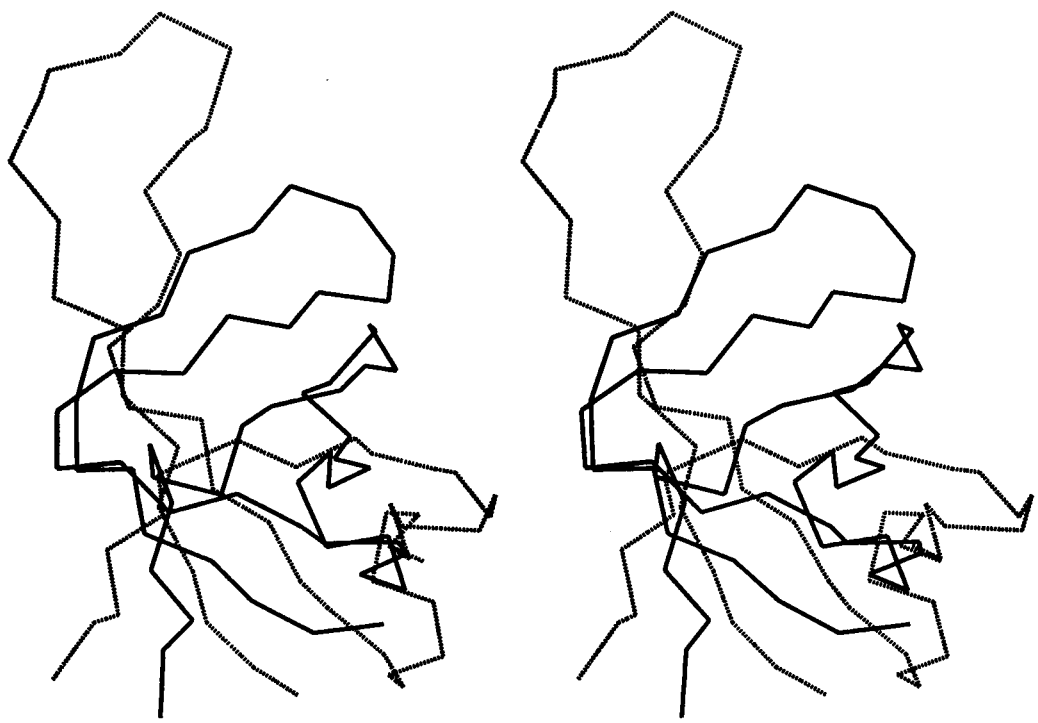

FIGURE 7 Superposition of the last frame of the SDCRYRS (*dotted line*) and the first frame of the MDWATF (*solid line*) trajectories projected onto the first eigenvector of the SDCRYRS+MDWATF eigenvector set.

tively small, indicating that the differences in equilibrium structure will be much smaller.

The root mean squares fluctuations in the projection of the single trajectories onto the eigenvector set calculated from the concatenated trajectories are shown in Fig. 8. These plots indicate differences in mobility of the single trajectories in the combined essential space. Again, there are large differences for the simulations with the solvation term. The fluctuations in the projections of the SDVACR and MDWATF trajectories over the SDVACR+MDWATF eigenvector set are more similar than all the other comparisons in this figure. Taken together with the results from Fig. 6, this indicates that the essential motions in the MDWATF and SDVACR trajectories are the most similar.

CONCLUSIONS

We have investigated different simulation methods and compared the static and dynamic properties. From the conventional analysis (Table 3) it appears that MDWATF and all other simulations not using the solvation term (MDVACR, MDCRYR, SDVACR, SDCRYR) have physical properties that are comparable to those derived from the crystal structure. Including the solvation term is identified as an unrealistic method of simulation. However, the analyses presented in Table 3 mainly concern properties based on structures and database statistics. These methods cannot give insight in the relative dynamical properties of these simulations. In addition, it seems that the various methods listed in Table 3 are not always consistent. In fact, some parameters show a good agreement with the crystal structure properties, whereas for other parameters the opposite is found.

To further investigate the reliability of these simulation methods, we have applied the essential dynamics method (Amadei et al., 1993). We have compared the full solvent simulation (MDWATF), which we consider to be the most realistic simulation method, to all other trajectories. From Figs. 6 and 8, it follows that addition of the solvation term always leads to significant changes of the essential dynamics of the protein, and large static shifts in the structure are produced. In addition, it appears that the use of crystal waters does not produce large effects on the essential dynamics properties. For thermolysin, we observed reasonable reproduction of the essential dynamics properties of a full solvent simulation (equivalent to MDWATF) by a simulation in vacuum with crystal waters (equivalent to MDCRYR) (van Aalten et al., 1995). However, including crystal waters in thermolysin is essential because of their buried nature, that is, they play a role in determining the structure and dynamics of the protein core. In the SH3 crystal structure no completely internal water molecules are present (see Fig. 1). The crystal waters simply form a mono/bimolecular layer on the surface of the protein. So although inclusion of crystallographic water molecules does not produce consistent and/or significant effects on the structure/dynamics of SH3, this does not necessarily mean that crystallographic water molecules are not important.

Before the 500-ps simulations described in detail in this paper, we had carried out shorter (250-ps) runs using a slightly different temperature and a heat-up protocol that did not involve reassigning velocities from a Maxwellian distribution with each increase of temperature (data not shown). Recently it was observed that slightly different initial conditions can lead to markedly different simulation results. However, for the systems we studied, the two different heat-up protocols, final temperatures, and simulation lengths gave essentially the same results.

The observation that SDVACR produces dynamics comparable to the MDWATF simulation is in contradiction to a similar comparison performed on cyclosporin (Yun-yu et al., 1988). In the equivalent of an SDVACR simulation on

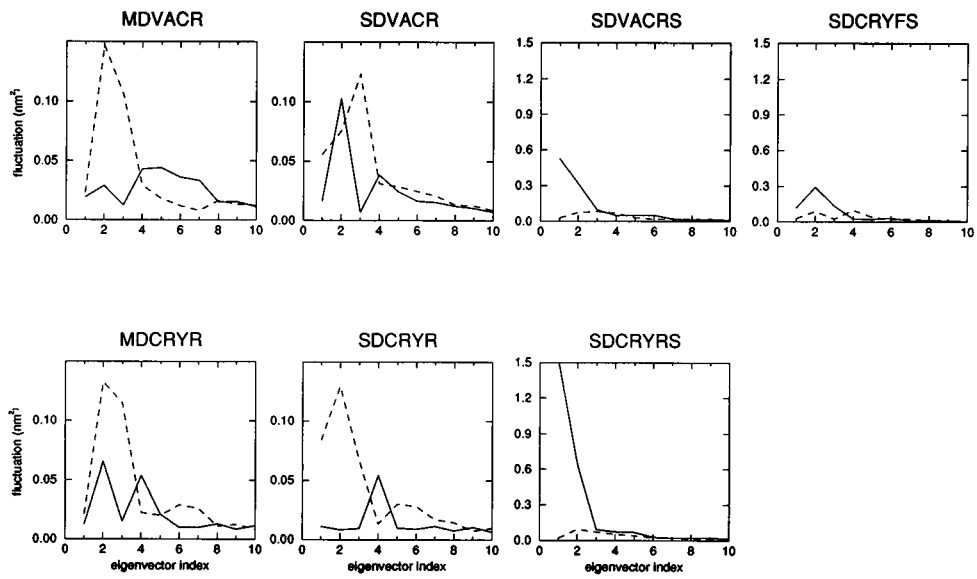

FIGURE 8 Root mean squares fluctuations in the total projection (displacements) of the single trajectories onto the first 10 eigenvectors calculated from the MDVACR+MDWATF through SDCRYFS+MDWATF trajectories. The solid lines indicate the MDVACR through SDCRYFS trajectories; the dotted lines indicate the MDWATF trajectory.

this small peptide, a large deviation in the number of hydrogen bonds with respect to an MDWATF simulation was observed. On the basis of this finding it was concluded that SDVACR did not produce MDWATF-like structures. A similar conclusion was drawn for the BPTI SDVACR simulation (Stouten et al., 1993). In our SH3 SDVACR simulation, we do not find a large deviation of (backbone-backbone) hydrogen bonds as observed for cyclosporin. Moreover, we have compared the dynamics of the two simulations using the essential dynamics method. This comparison shows that the large concerted motions in SDVACR are similar to the motions in MDWATF. On the basis of this observation and the results listed in Table 3, we conclude that SDVACR (and to a lesser extent the other non-solvation-term simulations) is a specially good mimic of MDWATF when simulating SH3. Unfortunately, it is not possible to derive general simulation strategies on the basis of the results presented here. Similar comparisons must be carried out using larger proteins with internal crystallographic water molecules.

In addition to comparing simulation methods, we have used the essential dynamics method to evaluate the motions of possible biological importance from MDWATF. The analysis of the motions along the essential eigenvectors revealed the presence of concerted motions involving two loops that have been shown to be important for substrate binding (Musacchio et al., 1994; Yu et al., 1994a,b).

We thank G. Vriend, B. de Groot, and Z.R. Wasserman for useful suggestions and critically reading the manuscript.

REFERENCES

Amadei, A., A. B. M. Linssen, and H. J. C. Berendsen. 1993. Essential dynamics of proteins. *Proteins Struct. Funct. Genet.* 17:412–425.

Auffinger, P., S. Louise-May, and E. Westhof. 1995. Multiple molecular dynamics simulations of the anticodon loop of tRNA-Asp in aqueous solution with counterions. *J. Am. Chem. Soc.* 117:6720–6726.

Baumann, C., C. Froemmel, and C. Sander. 1989. Polarity as a criterion in protein design. *Protein Eng.* 2:329–334.

Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. 1981. Interaction models for water in relation to protein hydration. In: Intermolecular Forces. B. Pullman, editor. Reidel, Dordrecht. 331.

Garcia, A. E. 1992. Large-amplitude nonlinear motions in proteins. *Phys. Rev. Lett.* 68:2696–2699.

Kabsch, W., and C. Sander. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 22:2577–2637.

Musacchio, A., M. Noble, R. Pauptit, R. Wierenga, and M. Saraste. 1992. Crystal structure of a Src-homology 3 (SH3) domain. *Nature.* 359: 851–855.

Musacchio, A., M. Saraste, and M. Wilmanns. 1994. High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline-rich peptides. *Nature Struct. Biol.* 1:546–551.

Noble, M. E. M., A. Musacchio, M. Saraste, S. Courtneidge, and M. Saraste. 1993. Crystal structure of the SH3 domain in human Fyn. Comparison of the three-dimensional structures of the SH3 domain in tyrosine kinases and spectrin. *EMBO J.* 12:2617–2624.

Ryckaert, J. P., G. Ciccoti, and H. J. C. Berendsen. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23:327.

Stouten, P. F. W., C. Froemmel, H. Nakamura, and C. Sander. 1993. An effective solvation term based on atomic occupancies for use in protein simulations. *Mol. Simulation.* 10:97–120.

Van Aalten, D. M. F., A. Amadei, A. B. M. Linssen, V. G. H. Eijsink, and G. Vriend. 1995. The essential dynamics of thermolysin—confirmation of the hinge-bending motion and comparison of simulations in vacuum and water. *Proteins Struct. Funct. Genet.* 22:45–54.

Van Buuren, A. R., S.-J. Marrink, and H. J. C. Berendsen. 1993. A molecular dynamics study of decane in water. *J. Phys. Chem.* 97: 9206–9212.

Van Gunsteren, W. F., and H. J. C. Berendsen. 1987. GROMOS: Groningen molecular simulation computer program package. University of Groningen, Groningen, the Netherlands.

Van Gunsteren, W. F., and H. J. C. Berendsen. 1988. A leap-frog algorithm for stochastic dynamics. *Mol. Simulation.* 1:173–185.

Vriend, G. 1990. WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* 8:52–56.

Vriend, G., and C. Sander. 1993. Quality control of protein models: directional atomic contact analysis. *J. Appl. Cryst.* 26:47–60.

Yu, H., J. K. Chen, S. Feng, D. C. Dalgarno, A. W. Brauer, and S. L. Schreiber. 1994a. Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell.* 76:933–945.

Yu, H., M. K. Rosen, T. Bum Shin, C. Seidel-Dugan, J. S. Brugge, and S. L. Schreiber. 1994b. Solution structure of the SH3 domain of Src and identification of its ligand binding site. *Science.* 258:1665–1668.

Yun-yu, S., W. Lu, and W. F. Van Gunsteren. 1988. On the approximation of solvent effects on the conformation and dynamics of cyclosporin A by stochastic dynamics simulation techniques. *Mol. Simulation.* 1:369–383.