Electrostatic screening in molecular dynamics simulations

T.Solmajer[1] and E.L.Mehler[2]

Department of Structural Biology, Biocenter, University of Basel, CH-4056 Basel, Switzerland

[1]Permanent address: Boris Kidric Institute of Chemistry, POB 30, 61115 Ljubljana, Slovenia, Yugoslavia

[2]To whom correspondence should be addressed

The screened Coulombic potential has been shown to describe satisfactorily equilibrium properties like pK shifts, the effects of charged groups on redox potentials and binding constants of metal ions. To test how well the screening of the electrostatic potential describes the dynamical trajectory of a macromolecular system, a series of comparative simulations have been carried out on a protein system which explicitly included water molecules and a system *in vacuo*. For the system without solvent the results of using (i) the standard potential form were compared with results of (ii) the potential where the Coulomb term was modified by the inclusion of a distance dependent dielectric, $\epsilon(r)$, to model the screening effect of bulk water, and (iii) standard potential modified by reducing the charge on ionized residue side chains. All molecular dynamics simulations have been carried out on bovine pancreatic trypsin inhibitor. Comparisons between the resulting trajectories, averaged structures, hydrogen bonding patterns and properties such as solvent accessible surface area and radius of gyration are described. The results show that the dynamical behaviour of the protein calculated with a screened electrostatic term compares more favourably with the time-dependent structural changes of the full system with explicitly included water than the standard vacuum simulation.

Key words: bovine pancreatic trypsin inhibitor/dielectric constant/electrostatic interactions/molecular dynamics/electrostatic screening

Introduction

Electrostatic interactions are implicated in a plethora of protein properties (Perutz, 1978; Harvey, 1989). The state of the art molecular dynamics simulations which are performed to describe the behaviour of molecular systems not amenable to more analytical approaches are based on a molecular force field (Karplus and Petsko, 1991; Wendoloski and Matthew, 1989; Kitchen *et al.*, 1990; Struthers *et al.*, 1990; Swaminathan *et al.*, 1990; Dauber-Osguthorpe *et al.*, 1988). Electrostatic interactions play a dominant role in the potential energy function and any attempt to apply this technique is conditioned by a reasonable choice for the parametrization of the electrostatic term.

Although recent advances in computer technology have alleviated somewhat the problem of system size, the explicit inclusion of solvent still presents a formidable computational effort. Therefore, models which can reproduce the effects of bulk solvent with a concomitant reduction in computing requirements continue to be of interest. One promising way to study the solvation problem is the use of integral equation theories for the solvent

structure and its effects on solute behaviour (Pratt and Chandler, 1977; Pettitt and Karplus, 1985). The site–site potential of the mean force for the atoms, of which the solute is composed, is computed from the intramolecular density distribution function. This very elegant approach, however, becomes impractical for larger protein systems. On the other hand, models which explicitly include solvent molecules, also increase the size of the system by roughly a factor of 10, while the number of nonbonded interactions which have to be calculated at each step of the simulation increase the computer time needed between one and two orders of magnitude, depending on the nonbonded cutoff used in the simulation. Thus numerous efforts have been made to reduce the size of the nonbonded interaction calculation by excluding the solvent. Two methods extensively employed to do this are (i) the use of a linear-distance-dependent dielectric function (Weiner and Kollman, 1981; Brooks *et al.*, 1983), $\epsilon(r) = r$, and (ii) use of a constant dielectric, ϵ, with a value of about 4 (Momany *et al.*, 1975). Both were developed on the basis of the assumption that the relative dielectric permittivity in the protein is most appropriately described as being between 2 and 5 (Pethig, 1979). Several lines of evidence (van Duijnen *et al.*, 1979; Warshel, 1979; Rees, 1980; Gilson and Honig, 1987; Sternberg *et al.*, 1987; Pickersgill, 1988) were recently presented which suggest that the value of the dielectric constant in the protein interior could be much higher than 4. Also theoretical estimates of the dielectric permittivity which would best describe independently determined experimental observables such as the pK of acidic and basic amino acids point to a significantly higher value (Mehler and Eichele, 1984; Warshel, 1984; Lavery *et al.*, 1986). Wendoloski and Matthew (1989) have found that the uniform value $\epsilon = 50$ most appropriately describes the static and dynamical behaviour of tuna cytochrome c. They (Northrup *et al.*, 1990) have subsequently used the high uniform dielectric permittivity value of $\epsilon = 80$ to calculate molecular dynamics trajectories and obtained good agreement with experimental values for the energy transfer rate.

The modelling of solvent by a linear, distance-dependent dielectric constant or by a low constant value was shown to introduce artifactual structural distortions (van Gunsteren and Karplus, 1982; Aquist *et al.*, 1985). An alternative possibility is to modify the charge distribution which accurately describes the dynamics of a protein system in solution to the *in vacuo* situation and is based on the experimental observation that counterions neutralize the charged amino acid side chains (Aquist *et al.*, 1985; Aquist and Tapia, 1990), especially those close to the protein surface. However, this approach also reduces or eliminates important ionic interactions in the protein interior or active site regions, where the structural barriers most probably prevent or reduce access of ions. Thus for applications such as simulation of active site processes the full charge force fields appear to be more suitable.

Recently, a nonlinear form of the screened Coulomb potential has been shown to provide a satisfactory description of properties like pK shifts, the effects of charged groups on redox potentials and binding constants of metal ions (Mehler and Eichele, 1984; Mehler and Solmajer, 1991). This nonlinear form is based on

the observed sigmoidal shape of the effective dielectric permittivity obtained from pK shifts in bifunctional organic acids and bases (Conway et al., 1951). Similarly, Hingerty et al. (1985) have developed a sigmoidal-type function which was successfully used in a simulation of a DNA single-base-pair disruption using a molecular mechanics approach (Ramstein and Lavery, 1988) and, most recently, in a molecular dynamics simulation of a small model peptide (Daggett et al., 1991). In view of its effectiveness in reproducing electrostatic interactions, it is of interest to probe its suitability to model solvent effects in dynamic simulations of proteins.

We have chosen the small protein, bovine pancreatic trypsin inhibitor (BPTI) as a test system for several reasons. On the one hand, its three-dimensional structure has been analysed with a multitude of experimental techniques, including X-ray diffraction to 1.5 Å resolution (Huber et al., 1970; Deisenhofer and Steigemann, 1975; Wlodawer et al., 1987), neutron diffraction (Wlodawer et al., 1984) and two-dimensional nuclear magnetic resonance (Wutrich et al., 1982), while on the other hand it has become a benchmark for detailed theoretical investigations. There is a large literature reporting calculations using molecular dynamics methodology in vacuo, in solution with explicit water molecules and in the crystal using periodic boundary conditions (van Gunsteren and Karplus, 1982; van Gunsteren et al., 1983; van Gunsteren and Berendsen, 1984; Levitt and Sharon, 1988; Kitchen et al., 1990; Swaminathan et al., 1990). There is also a wealth of experimental data available for BPTI and the computed properties of this model system could also be compared with results of more time-consuming calculations on the full system including explicit solvent. The primary structure of this 58 residue protein which contains an above-average content of charged amino acid residues at physiological pH, presents a further stringent test for any method aimed at an accurate description of electrostatic effects. Alternatively, another very attractive test system would be a protein with an above-average content of hydrophobic residues, such as crambin.

To test the facility of the screened potential to model the effects of solvent on the dynamical trajectory of a macromolecular system, comparative simulations have been carried out on (i) BPTI with solvent water molecules explicitly incuded (BOX); (ii) BPTI in vacuo using a standard potential with fully charged acidic and basic side chains (fully charge vacuum, FCV); (iii) BPTI in vacuo using a standard potential with neutralized side chains (reduced charge vacuum, RCV); and (iv) BPTI in vacuo using a potential in which the screening effect of the bulk water is modelled by the inclusion of a nonlinear, distance-dependent dielectric, $\epsilon(r)$, but explicit bulk water is excluded (screened Coulombic potential, SCP). The calculated trajectories, averaged structures, hydrogen-bonding patterns and solvent accessibility allow comparison of the dynamical behaviour of the protein with the time-dependent structural changes calculated for the full system with explicitly included water.

Materials and methods

All computations described were performed using the GROMOS87 molecular package (van Gunsteren and Berendsen, 1987). The standard empirical force field in this suite of programs consists of bond stretching, bond angle bending, improper dihedral and dihedral angle bending, and nonbonded interaction energy terms. For the BOX, SCP and FCV simulations, the charge set C with fully charged ionizable side chains at physiological pH (Arg, Lys, Asp and Glu) was used, and for the RCV simulation the charge set D was used, which models

solvent screening by placing zero net charges on all the amino acid side chains (Aquist et al., 1985; Aquist and Tapia, 1990). The electrostatic energy term uses a dielectric permittivity constant $\epsilon = 1$. In order to test the screened form of the potential using a nonlinear dielectric permittivity $\epsilon(r)$, we have modified the electrostatic term in the potential function and its gradient. The Coulomb potential term becomes now

$$E_{ij} = q_i \, q_j/\epsilon_{ij} \, r_{ij} \qquad (1)$$

where the functional form of $\epsilon(r)$ is given by

$$\epsilon(r) = A + B/[1 + k\exp(-\lambda Br)] \qquad (2)$$

and the derivative of the dielectric function is simply

$$d\epsilon/dr = \lambda(\epsilon - A)(\epsilon_0 - \epsilon) \qquad (3)$$

where $B = \epsilon_0 - A$, ϵ_0 is the dielectric constant of water, and A, λ and k are parameters as derived from data of Conway et al. (1951) in the original paper (parameter set ϵ_{wds} in Mehler and Eichele, 1984; see also Mehler, 1990; Mehler and Solmajer, 1991). Thus, such a form of the screening function is conceptually simple and because of equation (3) does not pose any considerable computational demands. On the other hand, its asymptotic behaviour when the solvent continuum limit is reached at $\epsilon = 80$ is essentially correct (Ehrenson, 1989) in contrast to the linear-distance-dependent and uniform dielectric formalism with high values of ϵ (Wendoloski and Matthew, 1989).

Identical simulation conditions were used in all systems considered. The charge distributions on the protein titratable sites were assigned to correspond to BPTI at physiological pH of 7.4. All the simulations were started from energy-minimized coordinates of BPTI, obtained from the crystal structure (Marquart et al., 1983). No switching function was used (Brooks et al., 1983). Such a method could be used in order to allow for a smooth behaviour of the nonbonded terms of the interaction function when applying a finite cutoff radius. The cutoff radius was set to 9.5 Å to compromise between reliability of the model and requirements for computer time.

Starting conformations were assigned initial velocities from a Maxwellian distribution and rotational and translational motions were removed after initialization. The integration time step was 2 fs and the temperature was kept constant at 300 K by coupling to a thermal bath with temperature relaxation time set to 0.025 ps. The extended atom model was used throughout and the SHAKE procedure (Ryckaert et al., 1977; van Gunsteren and Berendsen, 1977) was employed to eliminate the high frequency bond-length vibrations. A 10 ps equilibration period was followed by 100 ps production runs, except for the BOX simulation which included 1748 explicit water molecules and the trajectory was

Table I. R.m.s. deviations between average and crystal structures of BPTI[a]

Structure	No. of atoms	R.m.s. (Å) all atoms	R.m.s. (Å) main chain	Rel CPU time
Expl. H$_2$O (BOX)	5812	1.92	1.19	100
Screened (SCP)	568	2.84	1.64	9
Vacuum (FCV)	568	3.72	2.79	8
Vacuum (RCV)	568	1.99	1.41	8

[a]Average structures obtained from the last 50 ps of each run, residues 1. 57 and 58 were omitted from the r.m.s calculation.

run for only 70 ps. SPC potential parameters were used for the description of water molecules.

The resulting trajectories were analysed for root-mean-square (r.m.s.) deviations from the crystal and averaged structure from the water-box simulation, radii of gyration, R_g, for all atoms, hydrogen-bonding patterns and solvent accessibility. The latter was obtained by using the algorithm to calculate the solvent-accessible surface area (Lee and Richards, 1971) as implemented in the DSSP program (Kabsch and Sander, 1983). Hydrogen-bonding patterns were analysed from the time-averaged structures using a recently proposed graphical method. For a protein consisting of N amino acid residues this representation is constructed by scanning the protein sequence for all potential donor (PD) – proton acceptor (PA) pairs, ij, where i and j refer to residue numbers for the PD and PA respectively. Subsequently, a distance criterion and angle criterion [for main chain – main chain (abbreviated as m-m) H-bonds only] is applied to select the pairs which are actually H-bonded. The selected pairs are then mapped onto an $N \times N$ symmetric matrix, where the ij^{th} element is marked by an appropriate colour or symbol (Factor and Mehler, 1991). The changes in H-bonding resulting from the different simulations were analysed using three different methods for assigning hydrogen bonds: (i) geometry of the hydrogen bond (GROMOS87) (a hydrogen bond was considered as present if the distance between the PA and the proton was < 2.5 Å and hydrogen bond angle A..H-D was > 135°); (ii) energy of the hydrogen bond (the m-m hydrogen bond was considered as present if the electrostatic hydrogen bond energy was < −1.2 kcal/mol) (Kabsch and Sander, 1983); and (iii) the hydrogen bond was considered as present if the distance between PA and PD was < 3.20 Å and for m-m H-bonds the angle was > 110° (Factor and Mehler, 1991). We felt that comparative use

of several criteria which have so far been reported in the literature might give additional insights into this important component of protein structure.

Results

Average structures have been obtained from the molecular dynamics trajectories from the last 50 ps of the simulation for each system, and these are compared with the crystallographically determined structure in Table I. R.m.s. differences which are given for all atoms and main-chain atoms respectively, show that screening the electrostatic contribution yields r.m.s. deviations about half-way between the FCV and BOX results. The FCV force field produces large fluctuations in the electrostatic energy term which correspond to large r.m.s. deviations. The nonlinear distance-dependent screened potential decreases these deviations, thereby yielding smaller r.m.s values.

The average simulated structure which agrees most closely with the atomic positions of the crystal structure is obtained with explicit inclusion of water. However, the relative time needed for this simulation (Table I, last column) is roughly an order of magnitude larger than in the case of the screened or vacuum simulations. Of the various nonsolvent simulations, the r.m.s. differences from the RCV model, correspond most closely to the crystal structure. The scaled-down electrostatic interactions have significantly decreased the average structural deformations observed in the simulation using fully charged side chains. The time dependence of the r.m.s. deviations for the instantaneous structures generated along the trajectories are plotted in Figure 1, reflecting the range of conformations accessible to the protein during the simulation. Placing zero net charges on the side chains yields positional deformations from the crystal structure similar
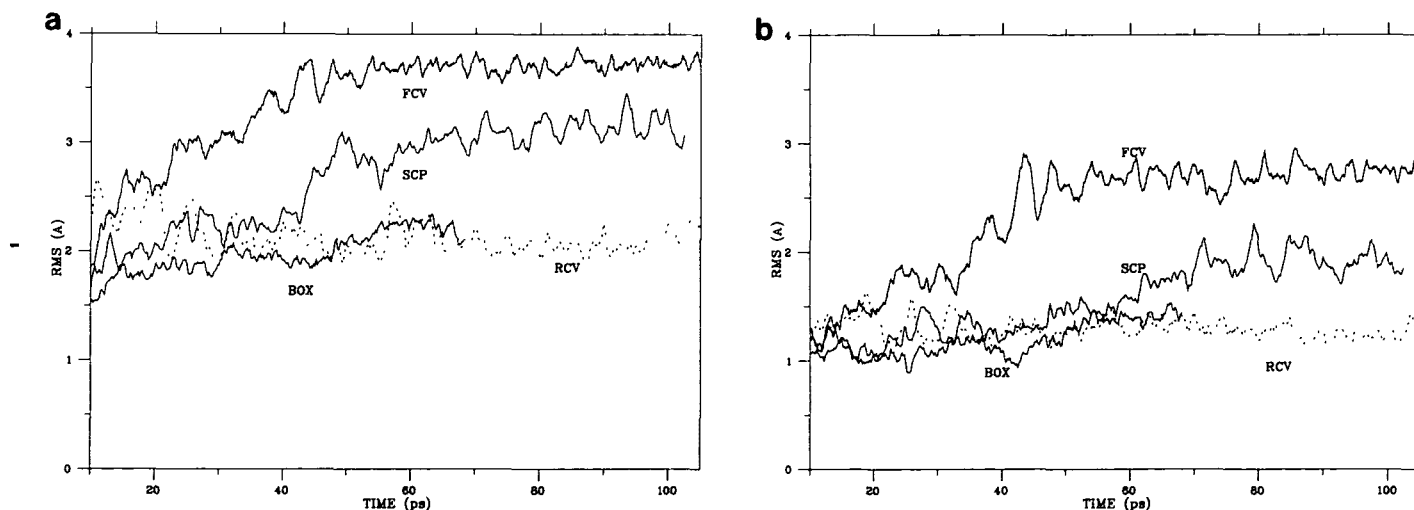

Fig. 1. Comparison of r m s. deviations from crystal structure of BPTI for the simulations in the water box (BOX), in vacuum using screened potential (SCP), in vacuum using standard potential (FCV) and in vacuum using neutral side chain charges (RCV). (a) All atoms, (b) main chain.

Table II. R.m.s deviations of simulated solvent structures from BPTI explicit solvent structure[a]

System	Segment					
	2−56	2−7 3_{10}	15−25 β-strand	25−28 loop	28−36 β-strand	46−56 α-helix
Screened (SCP)	1.97	1 37	1.76	1 37	1.00	1.07
Vacuum (RCV)	1.38	1.30	1.52	1 05	0.81	1.33
Vacuum (FCV)	3 71	1.07	3 43	1.10	2.11	1.61

[a]All atom r.m.s (in Å) differences between explicit solvent model (BOX) and simulated solvents
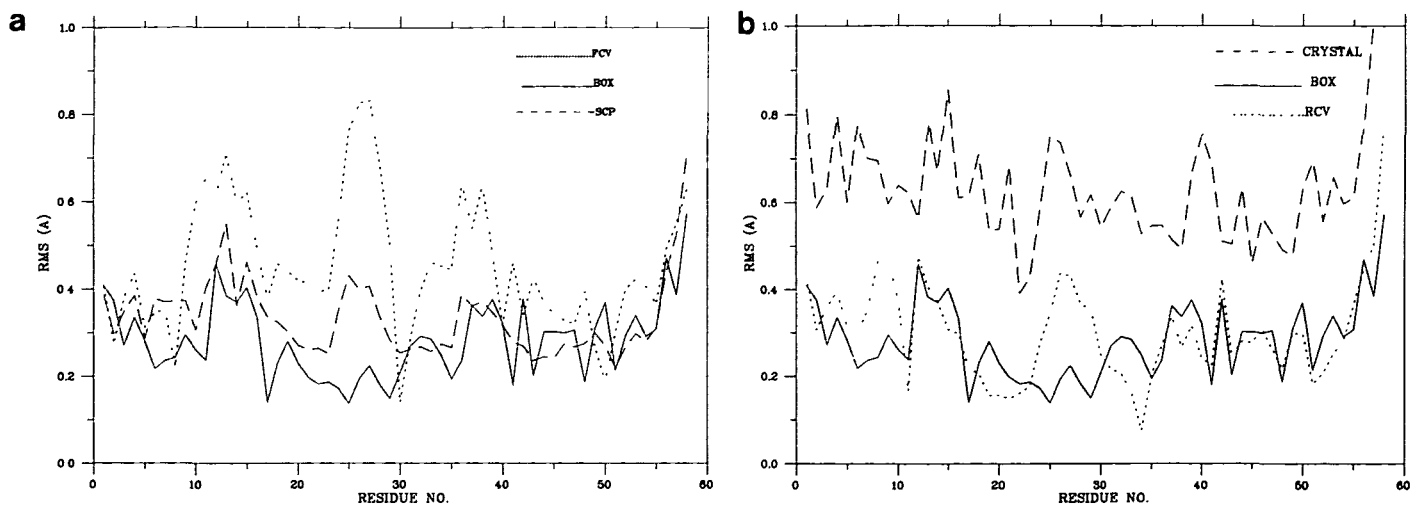
Fig. 2. Root mean square fluctuations of Cα atoms for averaged structures from molecular dynamics runs. (a) Vacuum simulation (FCV), screened potential (SCP) and explicit solvent simulation (BOX), (b) vacuum simulation (RCV), explicit solvent simulation (BOX) and temperature factors from the X-ray experiment (CRYSTAL).

to those obtained in the full solvent simulation, whereas using the SCP reduces these deformations by ~50%. It has to be noted here that the 20 ps simulation of BPTI in the crystal environment using the full unit cell of the PTI crystal including crystal water (van Gunsteren et al., 1983) gave an average structure with r.m.s. deviation of ~1 Å from the X-ray structure. The larger r.m.s. deviations for the isolated solute listed in Table I are partly due to the lack of inclusion of crystal contacts in the simulation.

It is instructive to decompose the overall r.m.s. deviations of the average structures into contributions from different segments of the three-dimensional structure. In Table II the r.m.s. differences between the average structures of the simulations are presented for various elements of secondary structure. The all-atom (excluding H atoms) r.m.s. differences between structures from in vacuo simulations and the structure obtained from simulation with explicit solvent (BOX) are similar to the r.m.s. differences between these simulations and the crystal structure except that the SCP and RCV appear to be closer to the BOX than the crystal. It is also clear that, with a few exceptions, the r.m.s. values of the main secondary structural elements are substantially smaller than for the whole chain. In particular, the values obtained from the SCP simulation are all <2 Å, although the all-atom r.m.s. is 2.0 Å. In contrast, the r.m.s. deviations of segments 2−7, 15−25 and 45−56 obtained from the RCV average structure are about the same or greater than the overall value. Thus the RCV simulation appears to distort some of the secondary structure elements more than in the SCP simulation.

Amplitudes of backbone fluctuations are plotted in Figure 2(a and b) along with temperature factors obtained from the X-ray diffraction data (Marquart et al., 1983). The overall behaviour of the SCP simulation shows similarities to both the observed temperature factors and the BOX fluctuations, although there are quantitative atom-by-atom differences. The secondary structural elements (two β sheets from residue Lys15 to residue Ala25 and Gly28 to Gly36) and helices (a 3−10 helix at position Asp3 to Leu6 and an α-helix from 47 to Gly56) are in the region of low r.m.s. fluctuations and both termini of the backbone are in regions of above-average flexibility. The r.m.s fluctuations of Cα atoms from the BOX simulation are shifted to much lower average values (~0.3 Å) in comparison with average fluctuations obtained from the crystal temperature factors of ~0.6 Å. This may be due to the short duration of the solution simulation (it is estimated

that continuation of the simulation by an order of magnitude more time would be required to explore a sufficiently large portion of conformational space to account for significantly longer relaxation times in the solvent) as well as the influence of the electrostatic interactions on the dynamic structure in the vacuum simulation. The backbone fluctuations in the vacuum simulation with neutral side chains have been significantly lowered by flattening the electrostatic contribution to the force field, although their behaviour still allows for differentiation between structurally more rigid and more flexible parts of the protein; e.g. all three solvent-free simulations show significant flexibility of the extended loop between Ala25 and Gly28. The reduction in the flexibility of the backbone is probably due to the fact that the overall forces have been decreased and, as a consequence, the motions were reduced appropriately.

A detailed analysis of the changes in secondary and tertiary structure observed in the simulations can be obtained from an analysis of the hydrogen-bonding pattern (HBM, Factor and Mehler, 1991) which are given in Figure 3. The HBM of the crystal structure (Figure 3a, lower triangle) clearly exhibits the structural elements described above. In addition there are two m-m H-bonds linking more distant parts of the chain (residues 11,36 and 21,45), and several main chain−side chain (m-s) H-bonds which stabilize the tertiary fold of the protein. The BOX simulation is also given in Figure 3(a) (upper triangle). The most obvious changes are (i) the loss of two H-bonds in the 3₁₀ helix at the N-terminus, (ii) the increased length of the β-sheet, i.e. formation of m-m 16,37, which replaces m-m 16,36 in the crystal structure, and the loss of two i,i + 4 H-bonds in the C-terminus α-helix. Overall, however, it is clear that the crystal structure has been well conserved in the full solvent simulation.

The FCV structure (Figure 3b) exhibits severe distortion with respect to both BOX and crystal structures. The N-terminus 3₁₀ helix is different from either structure and the α-helix has completely vanished. In addition, a large number of m-s and side chain−side chain (s-s) H-bonds have formed between charged and polar groups, which are well-known artifacts of unscreened vacuum calculations (e.g. Levitt and Sharon, 1988).

Figure 3(c and d) presents the HBMs of the SCP and RCV simulations respectively. In general both simulated structures reflect the BOX and crystal structure H-bonding patterns reasonably well. Moreover, the number of additional H-bonds
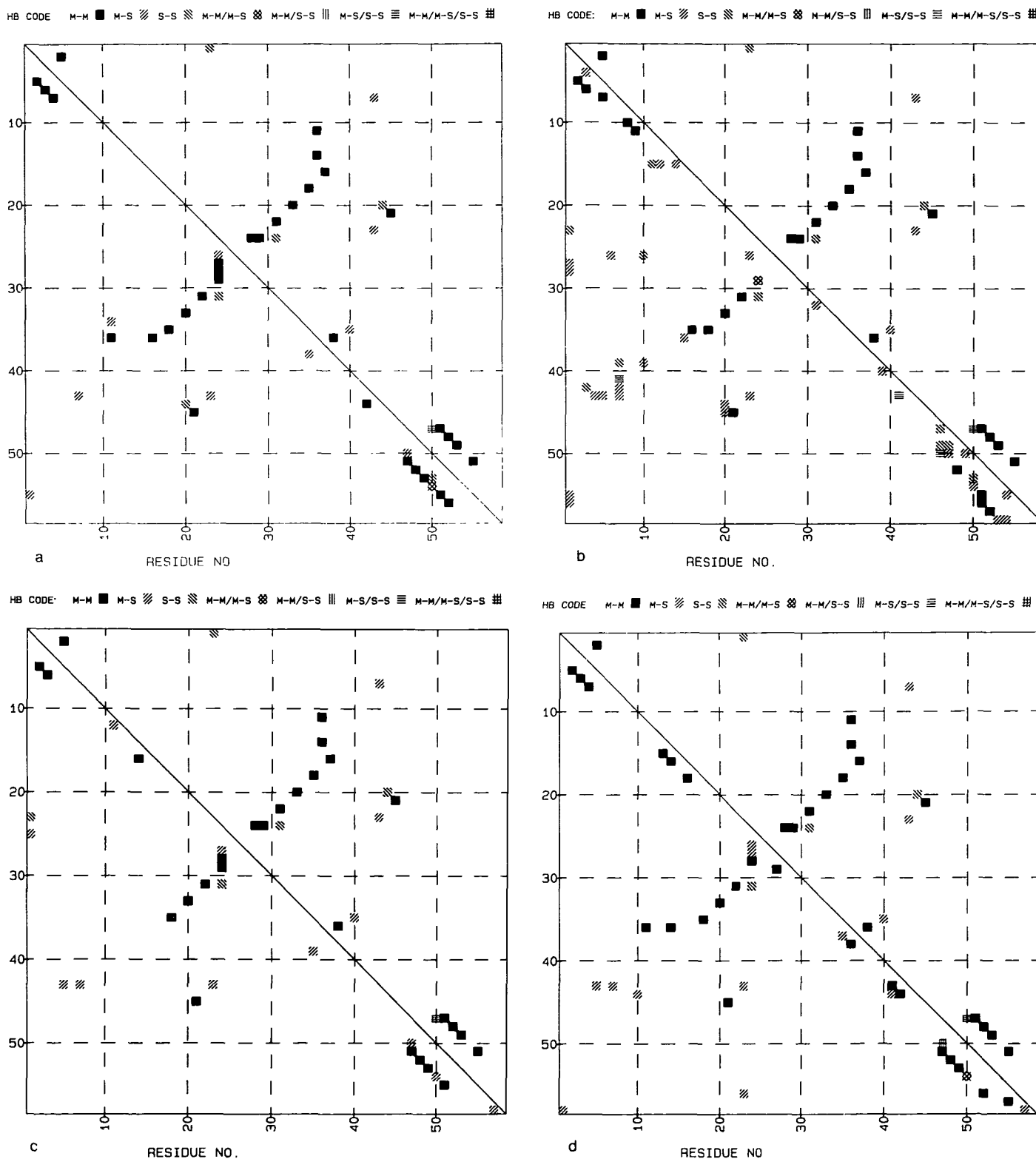
Fig. 3. Hydrogen-bonding matrices from crystal structure and average simulation structures. Upper triangle. BOX average structure in a–d; lower triangle: (a) crystal; (b) FCV, (c) SCP; (d) RCV

which have formed is not unreasonably large and is somewhat less in the SCP than RCV average structure. There are also important differences. The secondary structural elements of the SCP simulation are closer to those of the BOX simulation, whereas the RCV seems to resemble the crystal structure and also has formed several $i,i + 2$ m-m H-bonds which are not

present in either the crystal or BOX average structures. It appears that the SCP is closer to the model solution structure as obtained by the BOX simulation than the RCV.

The H-bonding patterns obtained from the different simulations and crystal structure are summarized in Table III. Results from three algorithms are given and the parameters of the HBM

Table III. Hydrogen-bonding patterns in BPTI

Structure/H-bond[d]	G[a]			K&S[b]	HBM[c]		
	M−M	M−S	S−S	M−M	M−M	M−S	S−S
Crystal	21	8	2	22	19	9	3
Box	19	6	2	19	15	5	4
Screened (SCP)	20	8	3	17	13	13	3
Vacuum (FCV)	16	20	13	16	13	35	17
Vacuum (RCV)	26	12	2	26	24	15	2

[a]Y..H < 2.5 Å, angle 135 < Y..H−X < 180 (Gromos, 1987)
[b]ΔE < −1.2 Kcal/mol (Kabsch and Sander, 1983).
[c]Y...X < 3..20 Å; angle Y..H−X > 110 (Factor and Mehler, 1991)
[d]M−M, main chain − main chain; M−S, main chain − side chain; S−S, side chain − side chain.

(Factor and Mehler, 1990) and Kabsch and Sander (1983) methods have been varied to give as close agreement as possible between the three algorithms for the crystal structure. It is seen that the number of m-m H-bonds calculated with the three methods for the different simulations is in reasonable agreement. For the BOX structure the GROMOS (van Gunsteren and Berendsen, 1987) and HBM algorithms also show good agreement for the number of m-s and s-s H-bonds.

As noted already from Figure 3, in the FCV calculation the increase in m-s and s-s H-bonds is considerable. This increase is much smaller in the SCP and RCV simulations. Also, the number of m-s H-bonds found by the HBM algorithm for the three model solvent simulations is larger than the number calculated by the GROMOS method. The reason for this is that the latter uses the proton to apply an angle criterion in addition to the distance criterion, whereas the HBM program only uses the PA−PD separation for counting the number of m-s H-bonds and therefore tends to overcount. However, in the present case these m-s H-bonds are not seen in either the crystal or BOX structures. The fact that so many PA−PD pairs approach each other to H-bonding distances is therefore an artifact of the simulation and the fact that the GROMOS force field has no terms to account for lone-pair electrons. By applying an additional angle criterion to these interactions the results from the GROMOS analysis tend to be 'over-optimistic' in predicting better results for the SCP and RCV simulations than is actually the case. On the other hand, all three methods point to deviations in H-bonding patterns for the m-m bonds in the RCV structure.

Of the properties presented in Table IV, the solvent-accessible surface area and R_gs provide a qualitative measure of alterations in the overall size and shape of the globular protein. It can be observed that for the FCV simulation the surface area is diminished ~17% from the surface area of the crystal structure and R_g is reduced by ~9%. Both SCP and RCV models fare approximately equally well, and R_g is ~4% smaller, while the solvent accessible surface area is contracted by ~11%. In contrast, the R_g and surface area of the BOX simulation is essentially unchanged from the crystal structure.

Discussion

We have compared three options by which one can hope reasonably well to run molecular dynamics of proteins systems without explicit inclusion of solvent. Two of these, the RCV and FCV have been used previously (van Gunsteren and Berendsen, 1984; Aquist et al., 1985), and our results are in agreement with the earlier calculations, albeit our simulations were continued for a somewhat longer time. The neutralization of the ionized side chains decreases the structural deformations more than the

Table IV. Deviation of overall properties of time-averaged dynamic structures

Structure	SA[a] ($\mathring{A}^2$)	R_g[b] ($\mathring{A}$)
Crystal	3936	11.21
Box	3895	11.26
Screened (SCP)	3519	10.81
Vacuum (FCV)	3248	10.42
Vacuum (RCV)	3484	10.97

[a]Solvent accessibility.
[b]Radius of gyration.

screened potential model and results in smaller r.m.s. differences between the simulated and observed structure. At the same time a more-detailed analysis of r.m.s. differences and changes in H-bonding patterns suggested that both simulations exhibit about the same level of distortion from the BOX simulation and crystal structure, and in fact, the SCP simulation is qualitatively closer to the BOX simulation average structure than the RCV average structure. It is clear, however, that the charge distributions which are in use in solvent simulations have to be scaled down if the structural changes in time are to be studied in vacuo. In this context it is of interest to compare our results with the simulations of Levitt and Sharon (1988) who also obtained a lower r.m.s. deviation (2.5 Å for all atoms) for the simulation in vacuum using a neutral guanidinium group model for the arginine side chains of BPTI. That placing zero net charges on polar side chains reduces r.m.s. deviations is further exemplified from the comparison of our result for the time dependence of r.m.s. differences from the BOX simulation of BPTI with those of Levitt and Sharon (1988). They obtained an all-atom r.m.s. of about 1.4 Å which is considerably smaller than the value of 1.9 Å obtained here, although with the exception of the neutralized Arg side chains, the simulation conditions were similar.

All in vacuo simulations show qualitatively similar trends in overall properties, solvent-accessible surface area and radii of gyration. The surface side chains collapse to the protein surface thus causing the reduction of surface area and r.m.s. distance from the centre of mass of the molecule. The reduction in both surface area and R_g is, however, smaller and much closer to the values from the solution simulation and experiment, for the SCP and RCV models than for the FCV model. Wendoloski and Matthew (1989) observed similarly, that the surface side chains collapsed to the protein surface in their in vacuo simulation of tuna cytochrome c using several low values for the dielectric constant. From the structural point of view the majority of polar and charged side chains are close to the surface of the protein.

It appears that the exponential reduction of the electrostatic interactions (closer to the surface) does have a favourable effect in maintaining protein secondary and tertiary structure as well as hydrogen-bonding patterns. Electrostatic interactions are a driving force for charged substrate−protein association and catalysis (Warshel and Russel, 1984). Such dynamic processes depend crucially on the specific sequence of the protein in question and the actual charge state of ionizable groups, as well as on its three-dimensional structure. Therefore, any model successfully simulating such processes must somehow account for this. The all-neutral side chains approximation seems to be inappropriate for this purpose, as well as for the prediction of pH and ionic strength-dependent properties. We have shown in a previous paper (Mehler and Solmajer, 1991) that the nonlinear screened potential used here satisfactorily reproduces just such properties. Results presented here indicate that by use of such a model for *in vacuo* molecular dynamics simulations one can expect reasonably well-preserved overall structural features, and a partial accounting of the effects of solvent on the protein's structure.

Conclusion

Success in modelling the electrostatically dependant properties of charged groups in proteins by the use of a screened electrostatic potential has prompted us to perform comparative simulation studies, using various approaches for modelling the solvent, which allow the calculations to be carried out without the explicit inclusion of water molecules. While the validity of molecular dynamics simulations has frequently been judged only on the basis of magnitudes of atomic fluctuations from the experimental, usually X-ray, structure the results reported here illustrate that a more extensive comparison is required. If in a protein system the enclosing solvent is removed, but the electrostatic interaction is left unmodified, charges on amino acid side-chain groups lead to unrealistic molecular motions as well as loss of secondary structure. Our results on a protein system simulated by fully charged polar groups corroborate well with previously described simulations in the literature. In contrast, the use of fully neutralized side chains seems to preserve the crystal structure. However, such drastic reduction of the effective electric charges probably introduces a too rigid molecular structure which does not sample conformational space of the molecule in a way appropriate for description of any large-scale conformational changes, such as hinge-bending motion or transformations of secondary structure which frequently are desired to be modelled. Furthermore, electrostatic interactions crucially determine not only the protein's structure but also its function, and these cannot be modelled by neutral amino acid side chains. It appears that the exponential reduction of the electrostatic interactions at larger distances does have a favourable effect on maintaining the protein's secondary and tertiary structure as well as hydrogen-bonding patterns.

Acknowledgements

References

Aquist,J. and Tapia,O. (1990) *Biopolymers*, **30**, 205−209.
Aquist,J., van Gunsteren,W.F., Leijonmarck,M and Tapia,O. (1985) *J. Mol Biol.*, **83**, 461−477.
Brooks,B.R., Bruccoleri,R.E., Olafson,B.D., States,D.J., Swaminathan,S. and Karplus,M. (1983) *J. Comput. Chem.*, **4**, 187−217.

Conway,B.E., Bockris,J.O'M and Ammar,I.A (1951) *Trans. Faraday Soc.* **47**, 756−766
Dauber-Osguthorpe,P , Roberts,V.A., Osguthorpe,D.J., Wolff,J., Genest,M and Hagler,A.T. (1988) *Proteins*, **4**, 31−47.
Dagett,V , Kollman,P.A. and Kuntz,I.D. (1991) *Biopolymers*, **31**, 285−304.
Deisenhofer,J. and Steigemann,W. (1975) *Acta Crystallogr.*, *B*, **31**, 238−250.
Ehrenson,S. (1989) *J. Comp. Chem.*, **10**, 77−93.
Factor,A. and Mehler,E.L. (1991) *Protein Engng*, **4**, 421−425.
Gilson,M.K and Honig,B. (1987) *Nature*, **330**, 84−86.
Harvey,S. (1989) *Proteins*, **5**, 78−92
Hingerty,B.E., Ritchie,R.H., Ferrel,T.L. and Turner,J.E. (1985) *Biopolymers*, **24**, 427−439.
Huber,R., Kukla,D., Ruhlmann,A., Epp,O. and Formanek,H. (1970) *Naturwissenschaften*, **57**, 389−392.
Kabsch,W. and Sander,C. (1983) *Biopolymers*, **22**, 2557−2637.
Karplus,M. and Petsko,G.A. (1991) *Nature*, **347**, 631−639.
Kitchen,D.B., Hirata,F., Westbrook,J.D., Levy,R., Kofke,D. and Yarmush,M. (1990) *J. Comp. Chem.*, **11**, 1169−1180.
Lavery,R , Sklenar.H., Zakrzewska,K. and Pullman,B (1986) *J. Biomol. Struct. Dyn.*, **3**, 989−1014.
Lee,B K and Richards,F.M. (1971) *J Mol Biol.*, **55**, 379−400
Levitt,M (1981) *Nature*, **253**, 694−698
Levitt,M. and Sharon,R (1988) *Proc. Natl Acad. Sci. USA*, **85**, 7557−7561.
Marquart,M., Walter,J , Deisenhofer,J., Bode,W. and Huber,R. (1983) *Acta Crystallogr.*, *B*, **39**, 480−490.
Mehler,E L. (1990) *Protein Engng*, **3**, 415−417
Mehler,E.L. and Eichele,G (1984) *Biochemistry*, **23**, 3887−3892
Mehler,E.L. and Solmajer,T (1991) *Protein Engng*, **4**, 903−910
Momany,F.A., McGuire,R.F., Burgess,A.W. and Scheraga,H.A (1975) *J. Phys. Chem.*, **79**, 2361−2381
Northrup,S.H., Wensel,T.G , Meares,C.F., Wendoloski,J J. and Matthew,J B. (1990) *Proc Natl Acad Sci. USA*, **87**, 9503−9507
Perutz,M. (1979) *Science*, **206**, 1187−1191.
Pethig,R. (1979) *Dielectric and Electronic Properties of Biological Materials*. Wiley, Chichester.
Pettitt,B.M. and Karplus,M. (1985) *Chem. Phys. Lett*, **121**, 194−201.
Pickersgill,R.W. (1988) *Protein Engng*, **2**, 527−530
Pratt,L. and Chandler,D. (1977) *J. Chem. Phys.*, **67**, 3683−3709.
Ramstein,J. and Lavery,R (1988) *Proc. Natl Acad. Sci. USA*, **85**, 7231−7235.
Rees,D.C (1980) *J. Mol Biol.*, **141**, 323−326.
Ryckaert,J -P., Ciccotti,G and Berendsen,H J.C. (1977) *J Comput Phys*, **23**, 327−341.
Sharp,K. and Honig,B. (1990) *Annu. Rev. Biophys Biophys. Chem*, **19**, 301−332.
Sternberg,M J.E., Hayes,R.F F., Rusell,A J., Thomas,P.G. and Fersht,A R. (1987) *Nature*, **330**, 86−88
Struthers,S , Tanaka,G., Koerber,S.C., Solmajer,T., Baniak,E.L., Gierasch, L M., Vale,W., Rivier,J. and Hagler,A.T. (1990) *Proteins*, **8**, 295−304.
Swaminathan,S., Ravishankar,G , Beveridge,D.L., Lavery,R , Etchebest,C. and Sklener,H. (1990) *Proteins*, **8**, 179−193.
van Duijnen,P. Th., Thole,B Th. and Hol,W.G.J. (1979) *Biophys. Chem.*, **9**, 273−280.
van Gunsteren,W.F. and Berendsen,H.J.C (1977) *Mol. Phys.*, **34**, 1311−1327.
van Gunsteren,W.F and Berendsen,H.J.C. (1984) *J. Mol. Biol.*, **176**, 559−564.
van Gunsteren,W.F and Berendsen,H.J.C. (1987) *Groningen Molecular Simulation (GROMOS) Library Manual*. Nijenborgh 16, Groningen, The Netherlands
van Gunsteren,W.F and Karplus,M. (1982) *Biochemistry*, **21**, 2259−2274.
van Gunsteren,W.F., Berendsen,H.J.C., Hermans,J., Hol,W G.J. and Postma,J.P.M. (1983) *Proc. Natl Acad. Sci. USA*, **80**, 4315−4319.
Warshel,A (1979) *Photochem. Photobiol.*, **30**, 285−290.
Warshel,A. (1984) *Proc. Natl Acad. Sci. USA*, **81**, 4344−4348.
Warshel,A. and Rusell,S.T (1984) *Qt. Rev. Biophys.*, **17**, 283−697
Weiner,P.K. and Kollman,P.A. (1981) *J. Comput. Chem.*, **2**, 287−303
Wendoloski,J.J. and Matthew,J.B. (1989) *Proteins*, **5**, 313−321.
Wlodawer,A., Deisenhofer,J. and Huber,R. (1987) *J. Mol. Biol.*, **193**, 145−156.
Wlodawer,A., Walter,J. Huber,R. and Sjolin,L. (1984) *J Mol. Biol*, **180**, 301−331.
Wutrich,K , Wider,G., Wagner,G and Braun,W. (1982) *J Mol. Biol.*, **155**, 311−319.