

A study of the electrostatics treatment in molecular dynamics simulations.

Robert Garemyr and Arne Elofsson*

Department of Biochemistry, Stockholm University, 106 91 Stockholm, Sweden

February 4, 2003

*To whom correspondence should be addressed

Fax: +46-8-153679

Tel: +46-8-161553

Email: arne@biokemi.su.se

Running Title: Cutoff in MD-simulations.

Abbreviations: MD, molecular dynamics simulations. th1 first molecule
from 2trx protein set; th2 second molecule from pdb set

Keywords: Molecular dynamics simulations, cutoff, dielectric, electro-
static, thioredoxin.

Abstract

This paper considers the treatment of long range interactions in molecular dynamics simulations. We investigate the effects of using different cutoff distances, constant vs. distance dependent dielectric and different smoothing methods. In contrast to earlier studies we find that increasing the cutoff over 8 Å does not significantly improve the accuracy, [Arnold, G.E., Ornstein, R.L. Proteins 18:19-33,1994], nor does using a distance dependent dielectric instead of a constant dielectric, [Guenot, J., Kollman, P.A. Protein Sci. 1:1185-1205, 1992]. This might depend on differences in simulation protocols and/or force field, as we use the CHARMM22 force field with stochastic boundary conditions, whereas earlier studies have used other protocols and energy functions.

We also note that the stability of the simulations is highly dependent upon the starting structure, showing that accurate molecular simulations not only depends on a realistic simulation protocol, but also on correct initial conditions.

INTRODUCTION

Computer simulation techniques, such as molecular dynamics (MD), are nowadays used as a standard tool for studying the structure and dynamics of biological systems. It has been shown in earlier studies,^{1, 2} that MD simulations using explicit water molecules describes protein dynamics more accurately than in vacuo simulations. One reason for this is that without explicit water the hydrophilic sidechains have no solvent molecules to interact with, and are therefore often found compacted onto the protein surface rather than being extended out into the solvent. Although computer power is rapidly increasing, inclusion of explicit solvent molecules put severe restrictions on the length of feasible sampling periods. Aqueous simulations are still limited to sampling periods of nanoseconds or below for most systems of biological interest. This stems from the fact that most of the computational burden is due to the evaluation of nonbonded interactions, which scales as N^2 . A number of different methods have been developed aiming to reduce the computational cost for evaluation of long range interactions. Examples are truncation of the nonbonded interactions beyond a certain cutoff radius; reaction field methods, which replaces long range electrostatic interactions by those with a dielectric continuum,³ as well as Ewald summation techniques,⁴ which assumes periodicity and long range order. The most common method is the truncation method, however, there are several problems with this method, including the introduction of non physical forces, due to the discontinuity in the gradient of the energy function, as well as the problem of switching of one atom-atom interaction before the other in dipolar systems such as water. To overcome the first problem one often use a smoothening function in order to get a smooth transi-

tion of the energy function to zero. The second problem can be handled by not allowing dipolar groups to be split by the cutoff criterion, so called group based truncation.[?]

Several studies have aimed to find simulation methods that strikes a good balance between computational efficiency and high accuracy. These studies include different methods to incorporate water as well as treating the long range interactions. Examples are studies on pure water systems,[?] relatively small polypeptide-water systems,^{?, ?} larger systems using either a water shell around the protein,^{?, ?, ?, ?} or periodic boundary conditions,^{?, ?}. This study is similar in nature to the later of these studies, but also extended to include the dependency of the starting structure. We have used two different starting structures from the two different molecules in the asymmetric unit of 2trx, and in oxidized/reduced state of the the active disulphide bridge, to study the accuracy of different simulation protocols. We have chosen to perform our simulations with the stochastic boundary method,[?] and using the CHARMM22 parameters. This method places a water sphere around the protein and treats the outer layer of the water with Langevin dynamics, i.e trying to reproduce interactions with water outside the sphere. To the best of our knowledge this is the first extensive study of cutoff interaction using the CHARMM param22 force field in conjunction with the stochastic boundary method.

Two earlier studies, Arnold&Ornstein,[?] and Guenot and Kollman,[?] render special interest for comparison of results. Arnold&Ornstein studied several solvent models including a series of nonbonded cutoff distances ranging from 9 Å to 20 Å, using constant or distance dependent dielectric and with or without a switching function. Solvent was incorporated into the models either by X-ray crystal waters alone or in combination with

different water shells with thickness in the range of 3 to 10 Å. The protein used in this study was phage T4 lysozyme and 100 ps molecular dynamics trajectories were calculated using the the Discover program and the consistent valence force field⁷. Amongst the results of this study was that (1) for models using a constant dielectric, a minimum cutoff radius of 15 Å and no switching function was required in order to get reasonable agreement with the crystal structure; (2) models using a constant dielectric and a cutoff radius of less than 15 Å gave large temperature differences between protein and water; and (3) using a distance dependent dielectric resulted in significant reduction of atomic fluctuations compared to results obtained using a constant dielectric. Guenot and Kollman used two different proteins, the trp repressor and BPTI, in order to confirm that the results were not protein dependent. Twelve different models were used for the trp repressor. Two models involved explicit water, modeled by a 4 Å thick water shell, using constant and distance dependent dielectric respectively. Forty ps trajectories were calculated using the AMBER program,⁸ and the all-atom force field by Weiner et al..⁹ Among the main results from the trp simulations were that using a distance dependent dielectric resulted in smaller rms deviations, smaller rms fluctuations and significantly smaller temperature difference between protein and water, than when using a constant dielectric. In the BPTI simulations four different models were used, all using a 4 Å water shell, but differing in that they used constant or distance dependent dielectric, and either an 8 Å cutoff or infinite cutoff. Fifty ps trajectories were calculated for these models, and for the distance dependent dielectric model with an 8 Å cutoff the simulation was continued to 121.5 ps. The BPTI results confirmed the results from the trp study. Further it was

found that, in contrast to the constant dielectric model, the distance dependent dielectric model showed essentially no conformational dependence on the nonbonded cutoff.

Evaluation of the ability to reproduce experimentally obtained structural and dynamical properties is complicated by several issues, e.g (1) the significance of using crystal structures for evaluation of simulations performed in water and (2) the lack of experimental data of structural or dynamical parameters with atomic resolution on MD time scales. Concerning (1), the following motivates our choice of simulating in solution: Firstly, we wanted to examine conditions that are commonly used in molecular dynamics studies. Secondly, it is not obvious that the same set of parameters are the best for a simulation in a crystal environment as in an aqueous environment and, lastly, the fact that high coincidence has been shown between structural data for globular proteins obtained by X-ray crystallography and solution NMR, despite the different environments.⁷ Concerning (2), one can argue that the production of stable trajectories, not deviating far from the crystal structure, is certainly a necessary condition for a good simulation. This motivates for instance the use of the rms deviation, between generated structures and the crystal structure, as one evaluation parameter. Another frequently used measure is the change in accessible surface area. Further, it seems plausible to demand agreement between a simulation's rms fluctuations about the average structure and experimental temperature factors, at least qualitatively, even though the former lack contributions from external motions. Most other experimental data lack the atomic resolution to be useful to distinguish the small differences in accuracy between different simulation models.

In this article we address the behavior of MD simulations using a series of differ-

ent cutoff distances, constant vs. distance dependent dielectric and different methods for smoothening of the long range potentials at the cutoff distance. We have used the following measures for evaluation of different methods: (1) the rms deviation of the calculated structures from the experimental X-ray structure, (2) correlations between calculated rms fluctuations about the average coordinates and the experimentally obtained temperature factors, (3) accessible surface area, (4) temperature gradient between protein and water, (5) 'rms-drift' and (6) CPU-time.

MATERIALS AND METHODS

We chose to perform all our simulations using *Escherichia coli* thioredoxin (Trx) since it is a fairly small and well characterized redox protein. Trx contains 108 amino acid residues, with a single disulphide bridge (C32-C35). The molecule is globular and has five β -pleated sheet strands and four α -helices. It has been shown that the structural changes occurring with the reduction of the disulphide bond do not affect the general structure of Trx, but that the structure around the active site is slightly changed.[?]

The simulation method has been thoroughly described in previous studies,^{?,?} and is therefore only described briefly in the present report. Starting from the crystal coordinates of oxidized Trx (Trx-S₂, pdb code 2trx), refined at 1.7 Å,[?] the following process was applied to each of the two asymmetric crystal units, denoted by th1 and th2. A 25 Å sphere centered at a point between the aromats (Y49, Y70, W28, W31) and the active site was filled by overlaying a 25 Å radius sphere of TIP3 waters. The system was then partitioned into two parts, a full MD region of both protein and water atoms included

in a 23 Å radius sphere and a Langevin region between 23 Å and 25 Å. Following the overlaying and the partitioning of the system, the positions of the water molecules were energy minimized for 50 steps of steepest descent minimization while the protein conformation was constrained. Reorientation of the water ball was then performed followed by repartitioning of the system[?]. The entire system was then energy minimized for 200 steepest descent steps with harmonic constraints on the protein atoms which were gradually decreased every 40 steps. The whole process was repeated for two cases where the disulphide bridge had been reduced. The resulting four systems, which we denote by th1ox, th2ox, th1red and th2red, comprised the starting points for the MD runs.

All simulations were performed using CHARMM[?], parameters (PARAM22), including all hydrogens. Initiation of the MD simulations was performed by instantaneously assigning all atoms a random velocity yielding an overall kinetic energy corresponding to 300 K, (Maxwell Boltzmann distribution).

Stochastic boundary conditions were used by propagating all water atoms in the Langevin region with a friction constant of 50 ps⁻¹,[?] while atoms within the reaction zone were propagated by the Verlet algorithm[?]. The frequency for checking whether an atom is in the Langevin region was set to 5 time steps, i.e every 10 fs. For all atoms in the Langevin region the temperature was checked every 20 fs and allowed to vary ±10 K. All covalent bonds between hydrogens and heavy atoms were constrained using the SHAKE method.[?]

Average coordinates were calculated for each trajectory by sampling coordinates every ps, rotating each frame for best fit relative to the starting structure, followed by 200 steps

steepest descent minimization for the resulting structures.

One of the most frequently used measures in order to assess the stability of a MD simulation over the course of time is the rms deviation between experimental coordinates and the generated structures:

$$rmsd = \sqrt{\frac{1}{N} \sum (r_i^{exp.} - r_i^{gen.})^2}$$

where $r_i^{exp.}$ and $r_i^{gen.}$ denotes the Cartesian coordinates for atom i in the experimental and generated structures respectively. After rotation for best fit, averages were calculated over all non hydrogen atoms and all backbone atoms respectively.

In order to assess the agreement between atomic fluctuations and experimental B-factors we calculated the root mean square (rms) fluctuations about the average structure:

$$rmsf = \sqrt{\frac{1}{N} \sum \langle (r_i^t - r_i^{aver})^2 \rangle_t}$$

where $\langle \dots \rangle_t$ denotes time average. The fluctuations was averaged over non hydrogen atoms for each residue. If the fluctuations are isotropic there is an exact connection between the rms values for an individual atom i and the corresponding experimental temperature factor B_i : $rmsf_i = \sqrt{3B_i/8\pi^2}$. The B-factors for all non hydrogen atoms, all backbone atoms and all sidechain atoms respectively was averaged per residue and converted to rms fluctuations using the formula above.

As a measure of how fast each model spans conformational space we have chosen the following measure: for each trajectory the rms deviation between the structure at time t and the structure at time $t+\Delta t$ is calculated.[?] Averages were calculated for different values of Δt . Solvent accessible surface areas were calculated using a 1.4 Å probe.[?]

In the initial phase of this work we tried 23 different models simulating for 50 ps and collecting data from the last 25 ps, see Table I. The simulation models tested the following parameters: (1) constant versus linear distance dependent dielectricity, (2) a series of nonbonded cutoff distances ranging from 6.0 Å to infinity, (3) smoothening the nonbonded truncation by using a shifting or a switching function as well as doing this on a group by group basis or on an atom by atom basis, (4) using a shorter cutoff distance for water-water interactions than for protein-water interactions and protein-protein interactions, (5) different force fields. Table I summarizes the different parameters used. For models using a switching function the default cuton-cutoff distance is 2.0 Å. However, the last five models uses a larger switching region, e.g the 8-14rswg-model with cuton = 8 Å and cutoff = 14 Å. The 8ge-model invokes the extended electrostatics method built into CHARMM. This method uses a time saving multipole approximation for electrostatic interactions between pairs of particles separated by distances larger than the cutoff radius.

We decided to proceed further with seven of the models from trial 1 plus three new models, this time simulating for 250 ps collecting data from the last 150 ps. Table II shows the parameters for the models used in trial 2. MD-trajectories were calculated for each model using each of the four starting systems, saving coordinates at 1 ps intervals. All average values were taken over the last 150 ps. In the last stage of this study we performed four 1ns simulations. This was only applied to the 8c and 14c models, using th1ox and th1red.

RESULTS

Simulating for 50 ps is clearly too short in order to obtain sufficient sampling of conformational space. Further, the results from trial 1 showed that for most models the rms deviation was still increasing at the end of the trajectory, data not shown. However, based on average rms deviation, illustrated by th1ox (heavy atoms) in Table I, and rms fluctuation, we used the results from trial 1 as a discriminator, sorting out clearly unstable models as well as models with questionable gain in performance taking computational cost into account.

Table I. The simulation models tried in the initial phase of this study.

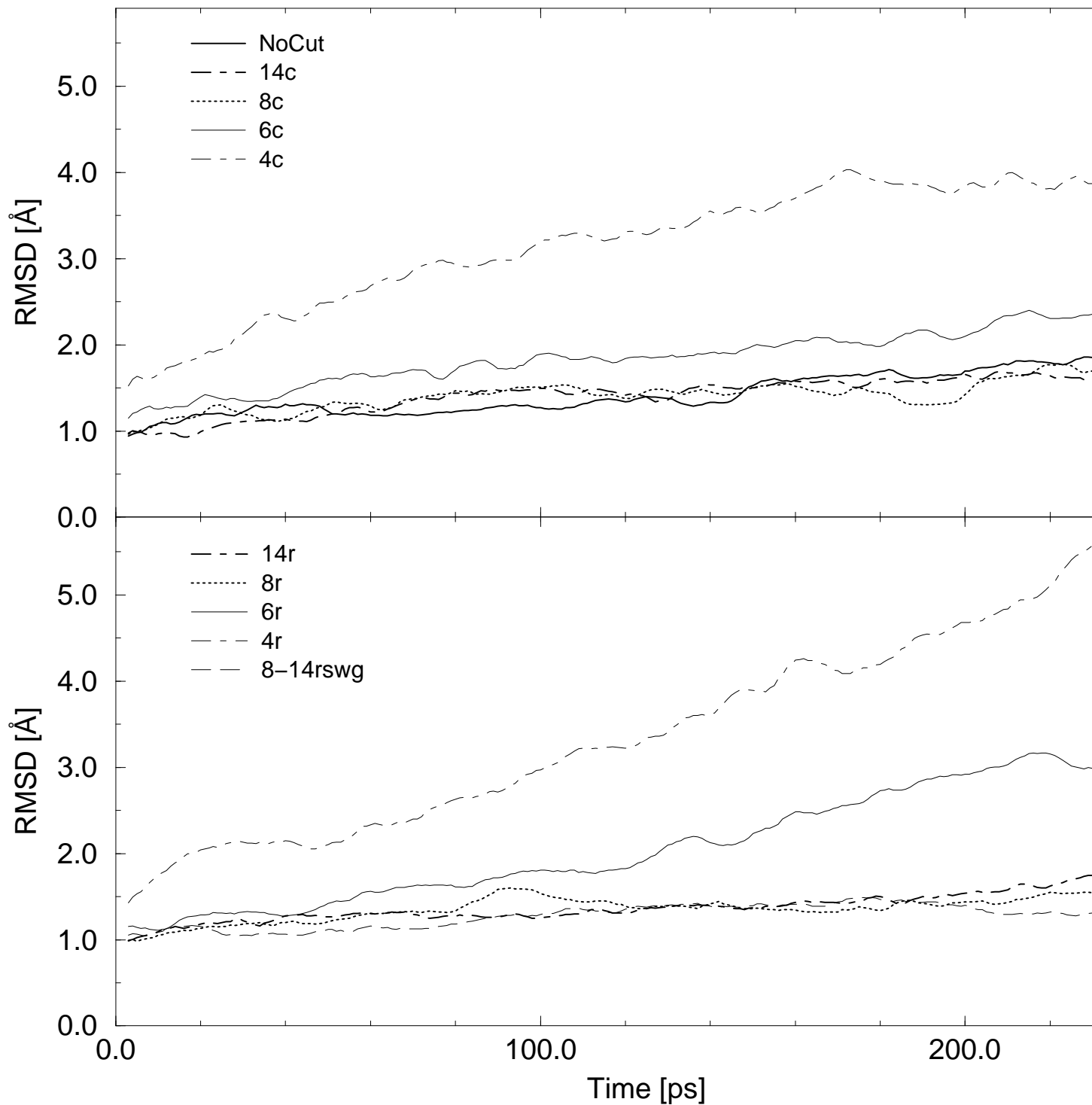
Model Name
\textbf{NoCut}& \textbf{1.0}& \textbf{100.0}& \textbf{fshift/atom}& \textbf{3.4}& \textbf{1.16}

The following models were disregarded from further studies since they were less stable than other models: 14p19, 14p20, 8ge, and 14w8. The 14c1fs and 8r1fs models were abandoned because of large computational cost and limited gain in accuracy. The 14g and 8g models were abandoned since they did not improve the results compared to the 14c and 8c models. Among the remaining models using a switching function, we decided to proceed further with the 8-14rswg model since the results indicated both computational efficiency and good stability, (lowest rmsd of all models considered). Models 8r and 6c were kept for further studies solely because they are interesting as reference models. In order to estimate the error in our data we performed ten simulations of the 8c-model (th2red) using different seed numbers for the assignment of initial velocities.

Rms deviations.

Time series for the rms deviation, averaged either over all heavy atoms or backbone atoms, were calculated relative to the starting structure. Figure 1 shows the rms deviation for all heavy atoms, plotted as a function of time and smoothed by averaging over 5 ps windows, for all simulations starting from th1red.

Figure 1. Rmsd development for all heavy atoms of different simulation models for th1red.



The 4c-model exhibit a steady increase for the first 170 ps and then reaches a plateau of about 3.9 Å. The 4r-model never reaches such a plateau, the rms deviation increases over the whole 250 ps interval of the simulation, ending at 5.9 Å. With the 6c and 6r

models the rms deviations gets considerably lower, although the 6r model exhibits the same behavior as the 4r model, i.e a steady increase over the 250 ps of the simulation. The NoCut, 14c, 14r, 8c, 8r and 8-14rswg models all show lower rms deviations than the 6c model over the last 150 ps, even though the stability of these simulations is still questionable, with a possible exception for the 8-14rswg simulation. We observed the same trends for all simulations starting from the th1ox structure. For simulations using th2ox as starting structure, the main difference is that the 4c-model exhibits a steady increase and ends up with a higher rms value than the 4r-model (data not shown). For th2red we observed very high rms deviations for the 6c-model over the last 150 ps, reaching a plateau of about 4.5 Å (data not shown).

Figure 2 summarize the average rms deviation for the last 150 ps of each trajectory relative to the crystal structure, for backbone atoms (bottom panels) and heavy atoms (top panels). It can be seen that the rms deviations are larger for th2 than for th1, for all simulations using a cutoff of at least 8 Å. For the models using very short cutoff radiuses, i.e 6c, 4c, 6r and 4r, we can not see this pattern.

In order to see if it was possible to discriminate the behavior of models using a cutoff of 8 Å and 14 Å we decided to use the 8c and 14c-models and simulate for 1ns. Figure 3 shows the backbone rms deviations as a function of time for th1ox and th1red, smoothed by averaging over windows of 5 ps. For th1ox it is hard to see any significant difference in behavior due to different cutoff radiuses, even though there is a weak indication of larger instability for the 8c simulation over the last 200 ps. For th1red we observe a large bulge starting at ~ 840 ps, reaching a first peak at ~ 900 ps and its

maximum at ~ 940 ps for the 8c simulation, whereas the 14c model remains quite stable over the last 250 ps. In order to investigate the reasons for the rmsd-bulge in the th1red 8c simulation we looked at the backbone rms deviations averaged by residue, over the 841-901 ps period. These results showed that the largest contributions to the rms deviation stems from one region, residue 30-41, with the largest values for residues 37-38. Further, concentrating on this region and the 841 ps and 901 ps structure, a dihedral angle analysis revealed large changes for Gly33 and Met37, $\Delta\phi \geq 70^\circ$, and for Cys32, Gly33, Cys35 and Pro40, $\Delta\psi \geq 70^\circ$. Figure 4 shows the differences between the 841 ps and 901 ps structures for residue 30-42.

Rms fluctuations. Reproduction of the experimental structure alone is not enough in order to discriminate a good MD-model from a bad model. Looking only at rms deviations from the crystal structure would judge a model that stays close to the crystal structure as good, even if this positive result is due to damping of important intramolecular motions or freezing out molecular degrees of freedom. For this reason an important measure of the correct dynamic behavior is the correlation between the rms fluctuations about the average coordinates and experimental B-factors. Figure 5 shows the correlation coefficients between converted B-factors and the rms fluctuations around the average structure for the last 150 ps of each simulation.

In Table III we report the average rms fluctuations over the last 150 ps for each model, as well as the average B-factors converted to rms fluctuations from the crystal structures. The estimated uncertainty for the average rms fluctuations is ± 0.03 . All models using an inner cutoff radius of at least 8 \AA show equal or lower rms fluctuations

than the experimental data, whereas the opposite is true for the 6c, 6r, 4c, and 4r models. The 4c and 4r models shows very large average rms fluctuations indicating substantial loss of stability due to the truncation of long range electrostatic interactions.

Accessible Surface Areas. Figure 6 summarizes the average accessible surface area over the last 150 ps for each model. The estimated error is 84.3 \AA^2 and the solvent accessible surface areas for the starting structures are: 5687 \AA^2 (th1ox), 5685 \AA^2 (th1red), 5762 \AA^2 (th2ox) and 5760 \AA^2 (th2red). Averages for the th1ox/th1red and th2ox/th2red starting structures are shown as horizontal lines. It can be seen that the accessible surface area increases for all models and, with an exception for the models using a 4 \AA cutoff, that the expansion is larger with a constant dielectric than with a distance dependent dielectric. Using a cutoff radius of 4 \AA leads to very large expansions even using a distance dependent dielectric.

The behavior as a function of time, data not shown, was very similar for all models using a cutoff radius of at least 8 \AA . For these models the accessible surface area was very stable during the whole simulation time. The 6c and 6r models were quite stable whereas the 4r and especially the 4c model showed dramatic expansion of the accessible surface area.

Temperature difference protein-water. Earlier studies,^{?, ?} have shown that, especially for constant dielectric models, the truncation of the nonbonded interactions may lead to severe problems in partitioning the kinetic energy between water and protein. In both these studies the whole system was coupled to a single temperature bath. This 'hot solvent-cold solute problem' can be avoided by separately coupling the solvent and solute

to two temperature baths, as in later studies by Kollman et al.[?] Since we do not use separate coupling for protein and water in this study we were worried that maybe our models exhibits large solvent-protein temperature differences. The result showed that the temperature difference was about the same, ~ 10 K, for the 14c, 6c and 14r models. We were a bit surprised that the 6c model showed this small temperature difference, as we intuitively expected the short cutoff radius to cause bad coupling between water and protein.

'Rms-drift'. Time averages over the last 150 ps were calculated for six different values of Δt . The results are summarized in Figure 7 for the th1ox-simulations. The qualitative behavior for the other three cases was very similar (data not shown). The behavior, both qualitatively and quantitatively, are very much the same for all models using a cutoff radius of at least 8 Å. The 6c, 6r, 4c and 4r models shows very large rms deviations for all six values of Δt relative to the other models, with a possible exception for the 6c model $\Delta t = 100$ ps.

Structural differences between the oxidized and reduced form. The solution structures of oxidized and reduced trx have been shown to be very similar.[?] Slight differences were found to be located in the region of the active site, which include Cys32 and Cys35, with an increase of the S-S distance upon reduction of the disulphide. The χ^1 angles of the two cysteines were found to remain the same in the two forms. Further, it was found that the overall backbone structure of oxidized *E. coli* thioredoxin in solution is very similar to molecule A in the asymmetric crystal unit, i.e th1, whereas the difference to molecule B, i.e th2, is considerably larger. This motivates a comparison between the

solution structures and the structures of the present study, considering only th1ox and th1red.

We have calculated the S-S distance and χ^1 angles for Cys32 and Cys35 for (1) the average solution structures of Jeng et al., pdb code 1xoa and 1xob; (2) the th1ox and th1red starting structures and (3) the average structures from the simulations using a cutoff radius of at least 8 Å. The results are summarized in Table IV.

DISCUSSION

Considering only the 250 ps simulations, almost all of the measures used for evaluation in this study have showed the same picture: the results for th2 are not as good as those for th1. To investigate the origin of this difference we calculated the rms deviations between the th1 and th2 crystal structures and found the largest values for residues 6, 10-13, 18-20, 52, 73, 82-85, 101, 104 and 107-108, see Figure 8 top panel. We also note that, for all models using a cutoff radius of at least 8 Å, there is a clear tendency for higher correlations between generated rms fluctuations and experimental B-factors for th1ox and th1red than for th2ox and th2red, see Figure 5. A closer look on a residue basis reveals a large bulge in the experimental B-factors for the thio2 crystal structure, located between res 10 to res 20, see Figure 8 bottom panel. This bulge is not reproduced by our models. Further, this is the same region Katti et al.⁹ refers to as the “disordered region” of th2. It is possible that there are strain inherent to the th2 structure, possibly due to crystal contacts, especially in the region between residues 10-20. Therefore, simulating in a different media than the actual crystal environment gives results deviating significantly

from experimental values, e.g large rms deviations between the crystal structure and generated structures. Further, as mentioned in the results section, experimental data has shown that the overall backbone structure of oxidized *E. coli* thioredoxin in solution is very similar to the th1 crystal structure whereas the difference to th2 is considerably larger.[?] This motivates using only results for th1 in the evaluation of the different models.

Considering only th1, it seems as if an 8 Å cutoff is sufficient in order to obtain rms deviations comparable to results achieved with larger cutoffs in earlier studies. The averages for th1ox/th1red are 1.28 Å and 1.56 Å for backbone and heavy atoms respectively. As a comparison Arnold&Ornstein reports backbone rms deviations of ~ 1.5 Å from simulations performed using cutoff radiuses in the range 11.5-20 Å.[?] Still, one can argue that the rms deviations in the present study are significantly larger than the ones reported by e.g Kitson et al.,[?] or by Guenot and Kollman[?]. However, the results by Kitson et al. were obtained simulating in the actual crystal environment and without any smoothening at the cutoff. In fact, on comparison of two simulations, both using a cutoff radius of 15 Å, but one with and one without a switching function, Kitson et al. found significantly lower rms deviations for the model using no switching function. We also tried to use abrupt truncation, with cutoff radiuses set to 8 Å and 14 Å, but the simulations turned out to be too unstable, forcing us to use smoothening functions. Further, the protein used in the second part of the study by Guenot and Kollman, (BPTI, pdb code 5pti), is smaller than trx and contains three cysteine bridges which may explain the smaller rms deviations compared to the ones in the present study. For the trp repressor, using distance dependent dielectric, Guenot&Kollman reports rms deviations to the crystal

structure that are similar to our results. However, the corresponding results using a constant dielectric are considerably larger, (2.23 Å).

In contrast to the results from earlier studies,^{?,?} we did not observe smaller rms fluctuations when using a distance dependent dielectric, compared to using a constant dielectric, although the model that exhibits the lowest fluctuations is the 8-14rswg model. This particular model seems to be quite stiff, with average rms fluctuations of 67-72 % of the experimental data for th1, indicating that there may be inappropriate damping of intra molecular motions. However, even though there are some damping, Figure 5 shows that the 8-14rswg model show good correlations with the experimental B-factors for th1ox and th1red. Further, it is interesting that models using a distance dependent dielectric exhibits smaller rms deviations for the rms-drift, for $\Delta t \leq 10$ ps and for each value of the cutoff radius, than the corresponding model using a constant dielectric.

Earlier results,[?] showed that rms deviations to the crystal structure, fluctuations about the mean coordinates and fluctuations in solvent accessible surface area were highly correlated. The results in the present study seems to support this, as Figure 2, Table III and the error bars in Figure 6 indicates. A thorough calculation of the correlation coefficients between the three measures gave results in the range 0.80-0.99. The lowest correlations is found for th1red and has its origin in the surprisingly low rms fluctuations in the accessible surface area for the 4c-th1red simulation.

The structural comparisons of the two cysteines in oxidized and reduced form, see Table IV, shows that data for the models of the present study, using a cutoff of at least 8 Å, are in good agreement with the experimental solution data. Considering

the th1red simulations, it is noteworthy that the S-S bond distance for the average structures is longer but quite close to the experimental value, despite the considerably shorter S-S bond distance for the starting structure. Further, for the th1ox simulations, we note a tendency where the C32 χ^1 angles approaches the experimental value, with an exception for the 14r simulation. It has to be stressed that these results are to be taken as indications of tendencies, and not in too much detail, as the sampling is limited.

Taking computational efficiency into account, the above results seems to suggest that the optimal cutoff radius is 8 Å. It is hard to see any improvement of the results if the cutoff is extended to 14 Å, or even further as in the nocut simulation. This is a bit surprising taking the results of Arnold,⁷ into account. Using a constant dielectric they found reasonably good results only for models using a cutoff radius of 15 Å or more. This difference might depend on the difference in simulation protocol and/or potential function.

Concerning the 1ns simulations, the aim was to further support the conclusion that no significant improvement of the results is obtained by increasing the cutoff radius from 8 Å to 14 Å. However, the results of these studies were not conclusive. We observe a tendency for less instability for simulations using a 14 Å cutoff than for the ones using an 8 Å cutoff during the last 200 ps (Figure 3). Further studies are required in order to determine whether this difference is significant or not, e.g running a series of simulations with different assignments of the initial velocities. The bulge in the rmsd for the th1red-8c simulation may or may not be a result of either the initial assignment of velocities or it may in fact be a direct consequence of neglecting important stabilizing long range

interactions. Further, one might ask if the increased stability for the th1ox-8c simulation is linked to the stabilizing effect of the disulphide bond.

CONCLUSIONS

The primary objective of this study was to find a model that strike a good balance between computational efficiency and good agreement with experimental data. The overall picture from the measures used in this study is that the models using a cutoff distance of at least 8 Å falls into a different class than the rest of the models. Further, using a cutoff of 14 Å does not seem to improve the results significantly over those obtained using a 8 Å cutoff. The 8 Å cutoff models are three times faster than the corresponding 14 Å cutoff model (Table II). The models using a 6 Å cutoff radius shows reasonable good results for average rms deviations to the crystal structure, correlations with experimental B-factors and protein-water temperature differences. However, the time development of the trajectories from these simulations indicates great instability, which is also revealed looking at the accessible surface areas. Models using a cutoff radius of 4 Å exhibits really poor performances for all measures used in this study, with a possible exception for the protein-water temperature difference which was not calculated for these models. At least within the framework of stochastic boundary simulations we do not recommend using a cutoff radius of 6 Å or shorter. Finally, almost all measures used for evaluation in this study showed that the stability of the simulations were highly dependent upon the starting structure. The results for th2 were not as good as those for th1.

Figure Legends

1. Rmsd development for all heavy atoms of different simulation models for th1red.
2. Rmsd between the average structure and the starting structure for all models. All averages were taken over the last 150 ps of each trajectory. Top panels, all heavy atoms. Bottom panels, only backbone atoms.
3. Rmsd from the starting structure for the 1ns simulations of th1ox and th1red.
4. The 841 ps and 901 ps structures from the th1red-8c simulation, residues 30-42.
5. Correlation between rmsf fluctuations and temperature factors. The estimated uncertainty for these correlations are ± 0.13 .
6. Solvent accessible surface area of the average structures.
7. Rmsd-drift for the different simulation models.
8. Top panel, rmsd between the th1 and th2 crystal structures versus residue number. Bottom panel, converted B-factors for th2 and rmsf for the corresponding NoCut simulation versus residue number.

Table II. The simulation models tried in the second phase of this study.

Model Name	Dielectric	Electrostatic cutoff (\AA)	Smoothing	CPU time vs. 8r
NoCut	1.0	100.0	fshift/atom	23.0
14c	1.0	14.0	fshift/atom	3.4
8c	1.0	8.0	fshift/atom	1.1
6c	1.0	6.0	fshift/atom	0.8
4c	1.0	4.0	fshift/atom	0.5
14r	$1.0r_{ij}$	14.0	shift/atom	2.9
8r	$1.0r_{ij}$	8.0	shift/atom	1.0
6r	$1.0r_{ij}$	6.0	shift/atom	0.7
4r	$1.0r_{ij}$	4.0	shift/atom	0.5
8-14rswg	$1.0r_{ij}$	14.0	switch/group	3.0

Table III. Average rms fluctuations for all heavy atoms.

Model	th1ox	th1red	th2ox	th2red
NoCut	0.74	0.75	0.76	0.75
14c	0.68	0.64	0.72	0.88
8c	0.70	0.79	0.74	0.81
6c	0.97	0.99	0.96	1.29
4c	1.45	1.39	1.61	1.42
14r	0.67	0.67	0.69	0.70
8r	0.67	0.66	0.76	0.86
6r	1.04	1.11	1.00	1.06
4r	1.61	1.77	1.65	1.54
8-14rswg	0.60	0.56	0.62	0.60
B-fact.→rmsf	-	0.83	-	0.88

Table IV. Structure data for the two cysteines C32 and C35. The experimental structures are average structures of 20 experimental solution structures in oxidized and reduced form respectively.

	th1ox			th1red		
	dist. S-S	C32 χ^1	C35 χ^1	dist. S-S	C32 χ^1	C35 χ^1
Experimental	2.07 Å	-175.5°	-59.7°	3.67 Å	-171.1°	-62.5°
Start	2.03 Å	-189.1°	-60.2°	3.19 Å	-171.6°	-70.1°
NoCut	2.03 Å	-171.1°	-54.3°	4.07 Å	-169.2°	-63.5°
14c	2.03 Å	-183.1°	-56.6°	3.97 Å	-165.4°	-60.6°
14r	2.03 Å	-189.9°	-54.5°	3.82 Å	-173.3°	-61.8°
8c	2.02 Å	-177.1°	-53.7°	3.74 Å	-168.1°	-60.0°
8r	2.03 Å	-178.8°	-48.6°	3.92 Å	-168.4°	-57.0°
8-14rswg	2.03 Å	-185.5°	-56.1°	3.85 Å	-169.4°	-64.3°