PROTEINS: Structure, Function, and Genetics 26:314–322 (1996)

An Extended Sampling of the Configurational Space of HPr From *E. coli*

B. L. de Groot,[1] A. Amadei,[1] R. M. Scheek,[1] N. A. J. van Nuland,[2] and H. J. C. Berendsen[1]

[1]Groningen Biomolecular Sciences and Biotechnology Institute (GBB), Department of Biophysical Chemistry, the University of Groningen, AG Groningen, The Netherlands; [2]Oxford Centre for Molecular Sciences, New Chemistry Laboratory, University of Oxford, Oxford, United Kingdom

ABSTRACT **Recently, we developed a method (Amadei et al., J. Biomol. Str. Dyn. 13: 615–626; de Groot et al., J. Biomol. Str. Dyn. 13: 741–751, 1996) to obtain an extended sampling of the configurational space of proteins, using an adapted form of molecular dynamics (MD) simulations, based on the essential dynamics (ED) (Amadei et al., Proteins 17:412–425, 1993) method. In the present study, this ED sampling technique is applied to the histidine-containing phosphocarrier protein HPr from *Escherichia coli*. We find a cluster of conformations that is an order of magnitude larger than that found for a usual MD simulation of comparable length. The structures in this cluster are geometrically and energetically comparable to NMR structures. Moreover, on average, this large cluster satisfies nearly all NMR-derived distance restraints.** © 1996 Wiley-Liss, Inc.

Key words: protein structure; molecular dynamics; essential dynamics; nuclear magnetic resonance; nuclear Overhauser effect; distance restraints

INTRODUCTION

HPr is a component of the phosphoenolpyruvate (PEP)-dependent phosphotransferase system (PTS). The PTS is responsible for the phosporylation and translocation of sugars from outside the cell to the inside, and the subsequent phosphorylation of these molecules (reviewed in ref. 1). The role of HPr in the PTS is the transportation of a phosphoryl group from enzyme I to enzyme II, the enzyme that performs the actual transport and phosphorylation of the carbohydrates.

The high resolution structure of HPr from *Escherichia coli* was elucidated both from NMR data[2] and x-ray data[3]. *E. coli* HPr consists of 85 residues and shows an open-faced β sandwich fold, with three α helices on top of a four-stranded β sheet.

Protein dynamics as derived from MD simulations can be split into two classes of motion.[4,5] A low-dimensional "essential" subspace, in which most of the fluctuations are concentrated, is distinguished from a high-dimensional "near-constraints" subspace in which merely small amplitude, fast equilibrating motions occur. Essential subspaces have been shown[4,6,7] to contain functionally relevant information for the simulated proteins. Discussion has been going on about how robust the definition of the essential subspace is[8–10] in view of limited sampling time currently available in MD. We have shown[6,11,12] that the description of the two subspaces approximately converges after only a few hundred picoseconds of simulation (although motions within the essential subspace are not adequately equilibrated in such short time, and therefore the definition of the single essential eigenvectors is far from being converged). We have developed a technique[11,12] that performs an adapted form of MD with constraint forces in the approximated essential subspace. This method yields an enhanced sampling of the configurational space as compared to usual MD, and the extrapolation of the essential subspace, as approximated from a relatively short MD simulation, yields a refined picture of this subspace. When the method was introduced, we presented an application to HPr, where we showed that a cluster of structures generated in this way samples a significantly larger part of the essential subspace, while all structural properties remain intact.[11] Subsequently, we applied a modification of the method to a peptide hormone, for which we obtained an almost complete sampling of the available space.[12]

In this article, we present an extended application to HPr where we study the quality of the obtained structures, the efficiency of the used protocol compared to usual MD, and the reproducibility of the results. Physical correctness of the obtained structures is not only measured in terms of geometrical checks, but also in terms of ensemble-averaged atomic distances, which are compared to experimental NMR data. The ED sampling technique was started from two different starting positions, using different initial approximations of the essential sub-

Received May 30, 1996; accepted May 31, 1996.

Address reprint requests to Dr. H. J. C. Berendsen, GBB, Department of Biophysical Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands.

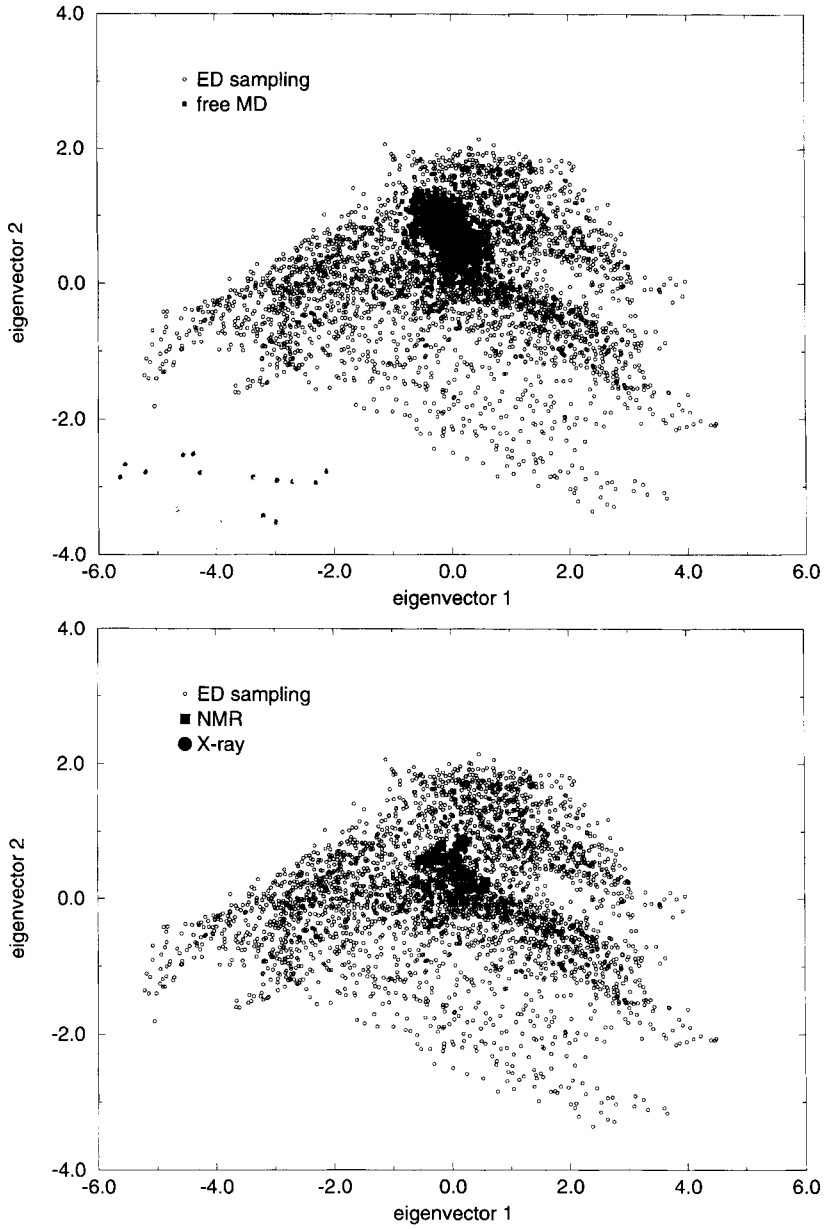

Fig. 1. **a:** Projection *(in nm) of the* collection of structures produced both by free MD (■, simulation time of 1.0 ns) and ED sampling (○, sampling "time" corresponding to 3.1 ns) onto the plane spanned by the two eigenvectors with largest eigenvalues from ED sampling. **b:** Projection (in nm) of NMR (■, pdb entry 1hdn) and x-ray (●, 1poh) structures compared to structures obtained from ED sampling (○) onto the plane spanned by the two eigenvectors with largest eigenvalues from ED sampling.

space, obtained from two MD simulations. A comparison is presented between the two clusters thus obtained, providing insight into the convergence of the essential subspaces.

METHODS

An MD simulation in solvent was started from one of the structures in the deposited NMR cluster (pdb entry 1hdn). This simulation of 350 ps is described elsewhere.[13] Over the last 300 ps of this simulation, a covariance matrix was built and diagonalized. Obtained eigenvectors are directions in configurational space, and corresponding eigenvalues give the mean square positional fluctuation for each direction.[4] The three eigenvectors with highest eigenvalues (the first three; eigenvectors are ordered to decreasing eigenvalue) were used in a constraint dynamics procedure, where constraint forces are only applied in this essential subspace (as described in ref. 11). The practical implementation is identical to that described in detail in ref. 12, except for two modifications. Briefly, the algorithm consists of the following steps: a starting position is defined as the set of essential coordinates of the starting conformation; a number of regular MD steps is performed; for each step, a new position is accepted only if it is not closer to the starting position than the previous position, in the subspace defined by the first three eigenvectors (i.e., if the distance from the starting position in this subspace does not decrease). If the new position is closer to the starting position, a correction is applied only in the subspace defined by the first three

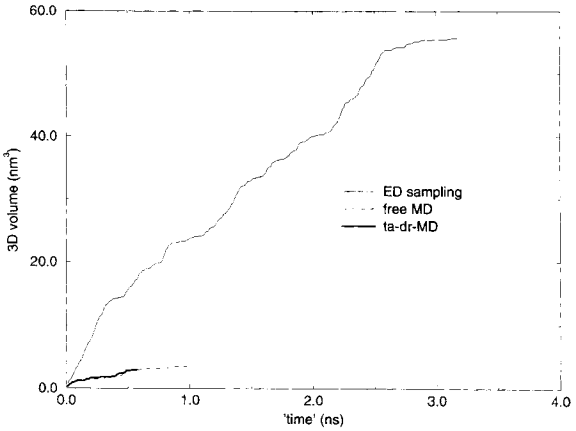

Fig. 2. Sampled configurational volume (in nm^3) in the space spanned by the first three eigenvectors from ED sampling.

eigenvectors with least perturbation.[11] The correction is applied along the radius direction such that the position after correction is at the same distance (in the essential subspace) from the starting position as the previous position. In this manner, the system is encouraged to sample new regions in the essential subspace, and prevented from going back to places that have been visited before. When the distance from the initial position does not increase spontaneously anymore, the cycle is finished, and a new cycle is started with the current position being the new starting position. Because of the applied correction the dynamics of the system is altered and therefore time information is lost, except for local, fast equilibrating motions. Integration steps are therefore merely dynamic sampling steps that cannot directly be interpreted as time steps.

Here we use the same algorithm, with two modifications. First, backbone N, Cα and carbonyl C atoms were used in the analysis instead of Cα atoms only, to make sure that all backbone motions are included in the analysis. We have shown previously[4] that Cα-only analyses usually yield all necessary information, but in the case of HPr, unusual ϕ/ψ combinations have been observed,[2,13–15] and inclusion of the other backbone atoms in the ED analyses is necessary for description of these angles. Second, the criterion for finishing a cycle to start another one was altered. The rate at which the distance from the starting position increases in time is taken as a criterion. This rate was allowed to decrease to zero in the case of the peptide hormone we studied previously,[12] but since such a low value resulted in denaturation of HPr, we chose for an average value of $2.5 \cdot 10^{-4}$ nm per step (for comparison, this value is typically $7.0 \cdot 10^{-4}$ nm per step in the initial phase of each cycle). We produced 125 of such cycles, in which an equivalent of approximately 1.8 ns of simulation time was reached. At a position distinct from the starting position of the first MD simulation, a second MD simulation of 350 ps was started. The eigenvectors derived from the covariance matrix built over the last 300 ps of this trajectory were used to direct a second sampling of 75 cycles (consisting of 650,000 steps, equivalent to 1.3 ns). A free MD simulation of 1.0 ns was performed for comparison.

All generated structures were subjected to energy minimisation with extra restraining forces (force constants of 4000 $kJ \cdot mole^{-1} \cdot nm^{-2}$) on distances for which experimental NOE data were available,[2] until no significant energy change was obtained.

Comparison with time-averaged distance restrained MD is presented. In total, three pieces of 200 ps (which originated from three distinct distance geometry structures) were used in these analyses. The distance restrained simulations are described elsewhere.[2]

The software used is an adaptation of the simulation package Gromos.[16] Essential dynamics analyses, structural checks and visualizations were performed with the molecular modeling program WHAT IF.[17] Secondary structure assignment and solvent accessible surface calculations were calculated by DSSP[18] and dihedral angle evaluations were carried out by PROCHECK.[19] Visualization of secondary structure maps was performed with a gromacs[20] analysis tool.

RESULTS

From the set of structures produced by ED sampling, a covariance matrix was built and diagonalized. The two independent pieces of sampling were combined in this analysis to yield a good definition of both the essential and near-constraints subspaces. Figure 1a shows the structures produced by ED sampling and free simulation, projected onto the plane defined by the first two eigenvectors. The region sampled during free MD is more compact than that obtained from ED sampling. The region obtained from ED sampling includes the area sampled by free MD, and extends it in every direction. The set of NMR structures[2] and the x-ray structure[3] lie close to the center of the region spanned by ED sampling (Fig. 1b). The sampled configurational volume,[12] expressed in the space spanned by the three most prominent directions in the cloud of structures, is depicted in Figure 2. The slope of the curve obtained from ED sampling is approximately seven times larger than that from free MD and time-averaged distance restrained MD, in the first nanosecond, demonstrating the enhanced efficiency of sampling using the ED approach.

The structures produced by both ED sampling and free MD were validated in several ways. Figure 3 shows the secondary structure as a function of time. Both during free MD (Fig. 3a) and ED sampling (Fig. 3b), the fold of HPr remains essentially stable. The amount of fluctuation around the mean structure is larger for the ED sampling than for the free

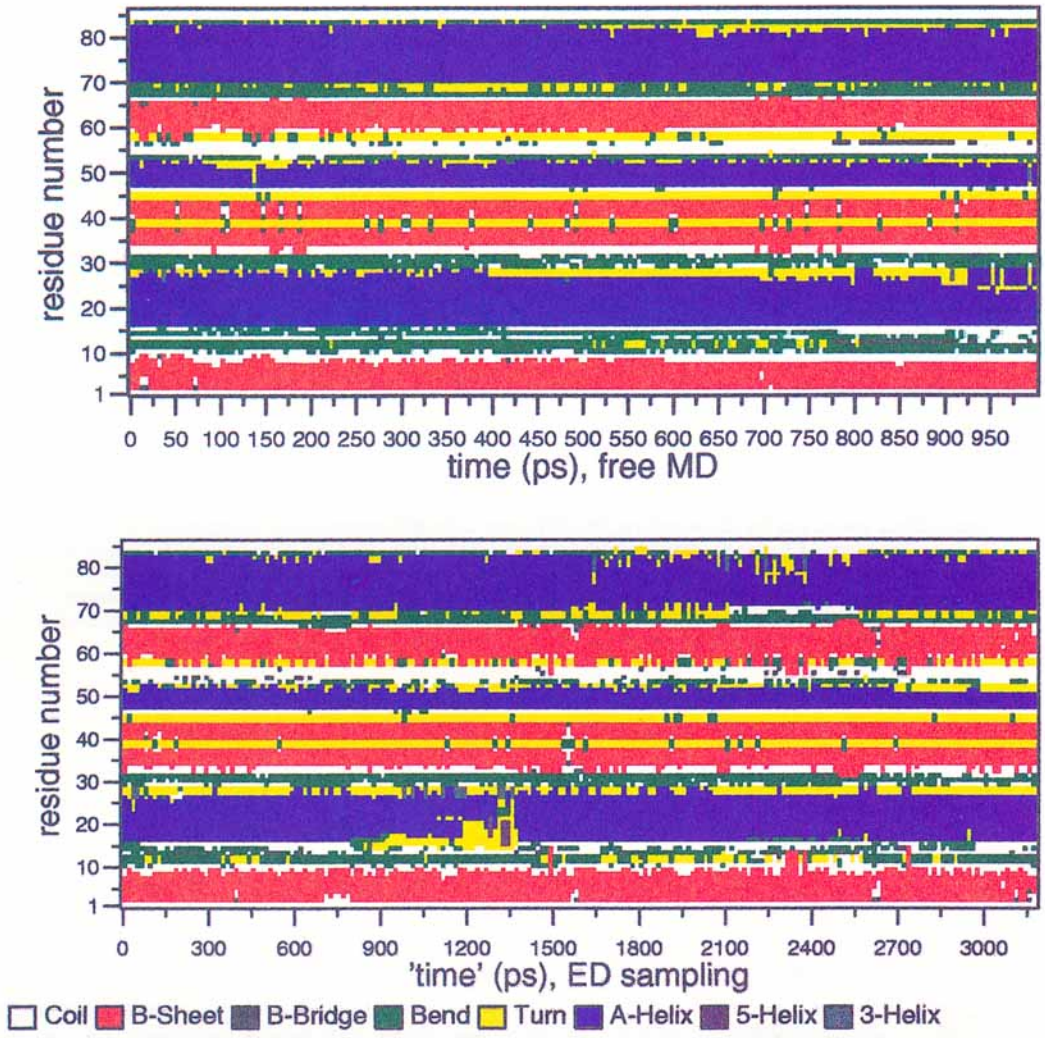

Fig. 3. Secondary structure as a function of time. **Top:** Free MD (1.0 ns). **Bottom:** ED sampling (3.1 "ns"). Data shown are for the reference simulation of 1 ns, and for the ED sampling, after combination of the two independent pieces.

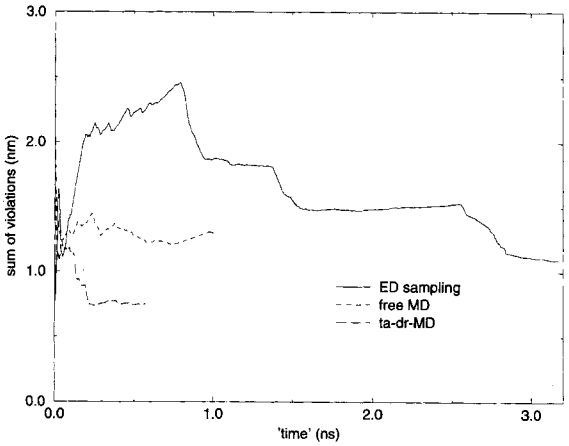

Fig. 4. Sum of violations of $<r^{-6}>^{-1/6}$ averaged distances with respect to upper limits derived from NOE data as a function of the number of integration steps in the calculation. Data shown are for ED sampling, free MD, and time-averaged distance restrained MD. For free MD and time-averaged distance restrained MD, the number of integration steps is proportional to time.

MD simulation. All other geometrical properties are summarized in Table I. For all the checked properties, averages for the free MD simulation, time-averaged distance restrained MD, and the ED sampling procedure are very similar. Only the RMSD from the mean structure is on average significantly larger in the ED sampling procedure. As observed for the secondary structure (Fig. 3), fluctuations of all geometrical properties in Table I are larger during the ED sampling than during free MD. Together with the higher average RMSD, this again demonstrates larger conformational freedom during ED sampling. Energies (Table I) are indistinguishable for the two forms of simulation. Time-averaged distance restrained MD shows even smaller fluctuations, in agreement with previous findings.[21]

As a further validation of the produced structures, we monitored correspondence to distances derived from experimentally observed NOE data. Results

TABLE I. Comparison of Geometrical and Energetical Properties Between ED Sampling, Free MD, and Time-Averaged Distance Restrained MD (ta-dr-MD)*

	Mean ED	σ ED	Mean MD	σ MD	Mean ta-dr-MD	σ ta-dr-MD
NRC	14.36	2.44	14.88	1.88	12.14	1.93
ACC	5551	184	5490	134	5635	134
DIH	5.08	1.21	4.70	1.11	4.60	0.92
HBO	69.4	4.2	67.6	3.7	71.0	4.0
GYR	1.143	0.011	1.140	0.008	1.165	0.007
RMS	1.715	0.469	1.053	0.150	0.762	0.124
EPOT	−130.48	0.42	−130.56	0.43	NA	NA
EPW	−17404	370	−17391	375	NA	NA

*The first column contains abbreviations of the studied properties: NRC, number of residues (out of 85) adopting random coil conformation; ACC, solvent accessible surface (Å^2); DIH, number of residues in unfavorable regions of a Ramachandran plot [22]; HBO, number of backbone-backbone hydrogen bonds; GYR, radius of gyration (nm); RMS, root mean square deviation (Å) with respect to mean structure; EPOT, total potential energy (all interactions, including solvent) in GJ/mole; EPW, All energy terms involving the protein in kJ/mole. The other columns show the averages and root mean square fluctuations of these properties for ED sampling, free MD, and ta-dr MD respectively. Energies are not shown for the time-averaged distance restrained MD simulations because in these calculations, an extra term is present in the potential energy function, making direct comparison impossible. It should be noted that the values for the time-averaged distance restrained MD simulations were averaged over three simulations that started from distinct initial NMR structures, making fluctuations larger than what would have been found for a single simulation.

are summarized in Table II. For both the ED sampling and the free MD simulations, violations with respect to the experimentally derived distances were evaluated after averaging over the whole set of structures, that were energy minimized with extra potentials on NOE derived distances. Only a very small fraction of distances (out of 1108 experimentally observed NOEs) is on average larger than the experimental upper limit, both during free MD and ED sampling. The sum of violations is in both cases of the same order of magnitude as calculated from the cluster of structures produced by time-averaged distance restrained MD simulations,[2] the last step in the refinement of the NMR structures. The running average of the total sum of violations decreases in time (Fig. 4), which suggests that a more complete sampling will yield even fewer and smaller violations.

Eigenvectors obtained from two independent ED sampling runs and two independent MD simulations (from which the eigenvectors used to direct the two sampling runs were extracted) were compared to assess whether the definition of the refined essential subspace depends on their initial approximation. Figure 5 shows a comparison between the essential subspaces from MD and ED sampling. For the first MD simulation (Fig. 5a), the overlap between the essential subspaces from the ED sampling runs and from MD is approximately equal. This indicates that during sampling, the eigenvectors which are used to direct the sampling do not dominate the produced cloud of structures. For the second MD simulation (Fig. 5b), the overlap between the essential subspace of the second sampling is somewhat larger than that for the first sampling. Figure 5c shows that the eigenvectors obtained from the two sampling runs are as similar to one another as are the eigenvectors from two free MD simulations. It should be noted that in all graphs of Figure 5, the overlap between the first 10 eigenvectors of each set is always higher than 40% and that within the first 50 eigenvectors (out of 765), more than 80% of the essential subspace of the other set can be rebuilt.

All trajectories were projected onto the eigenvector with highest eigenvalue from the whole cluster of ED sampling structures. Figure 6 shows the structures corresponding to the minimum and maximum position in this direction that were visited both during free MD and ED sampling.

As observed earlier[11] motion inside the essential subspace is of diffusive nature. Figure 7a shows that, for eigenvectors, 1, 2, and 3, during pieces of free simulation, the average square displacement increases linearly with time, indicative of diffusive behaviour. For comparison, the same evaluation was done for a combination of near-constraint eigenvectors (Fig. 7b). In this case, a linear dependence was observed for a few picoseconds, after which the curve levels off, indicating the presence of a significant free energy gradient for these coordinates, as has been postulated.[4]

CONCLUSIONS AND DISCUSSION

As we found in an initial study of HPr[11] and for a small peptide (13 residues),[12] with the ED sampling technique it is possible to extend the amount of sampling in the essential subspace of proteins. For the peptide it was even possible to approach a complete sampling within a simulation time corresponding to a few nanoseconds. We showed in the present study

TABLE II. Violations With Respect to NOE-Derived Distance Restraints*

	ED	MD	ta-dr-MD		
Sum of violations (nm)	1.093	1.300	0.739		
Largest violation (nm)	0.161	0.165	0.140		
	Residue 1	Residue 2	Upper limit	Average	Violation
	Violations larger than 0.1 nm (ED)				
1	10	85	0.580	0.705	0.125
2	14	57	0.650	0.764	0.114
3	14	80	0.770	0.918	0.148
4	22	55	0.670	0.819	0.149
5	23	55	0.870	1.031	0.161
	Violations larger than 0.1 nm (MD)				
	14	80	0.770	0.929	0.159
2	22	55	0.670	0.809	0.139
3	23	55	0.870	1.035	0.165
4	29	33	0.540	0.643	0.103
5	29	77	0.450	0.592	0.142
6	76	77	0.450	0.556	0.105
7	76	80	0.450	0.592	0.142
	Violations larger than 0.1 nm (ta-dr-MD)				
1	26	76	0.650	0.756	0.106
2	73	76	0.450	0.590	0.140

*For comparison, averaged distances were calculated as $<r^{-6}>^{-1/6}$ over clusters generated by ED sampling, free MD, and time-averaged distance-restrained MD.

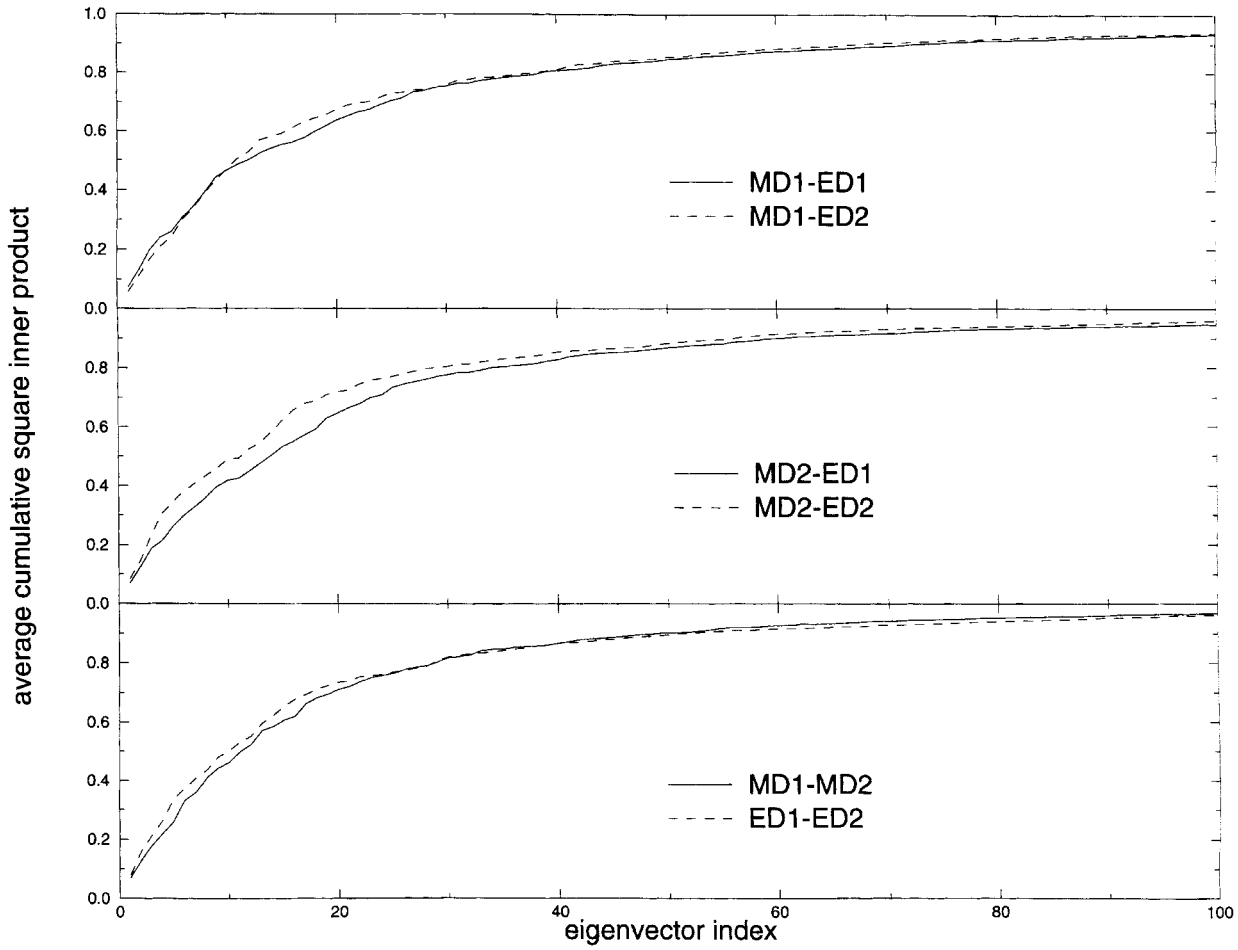

Fig. 5. Cumulative average square inner product between different sets of eigenvectors. The first 100 eigenvectors (of 765) of one set (the set named first in the legend) are compared to the first 10 eigenvectors of the other set. **Top:** Compares one MD simulation of 300 ps to two ED sampling runs. **Middle:** Shows the same, now for another simulation of 300 ps. **Bottom:** Compares the two MD simulations and the two ED sampling runs.

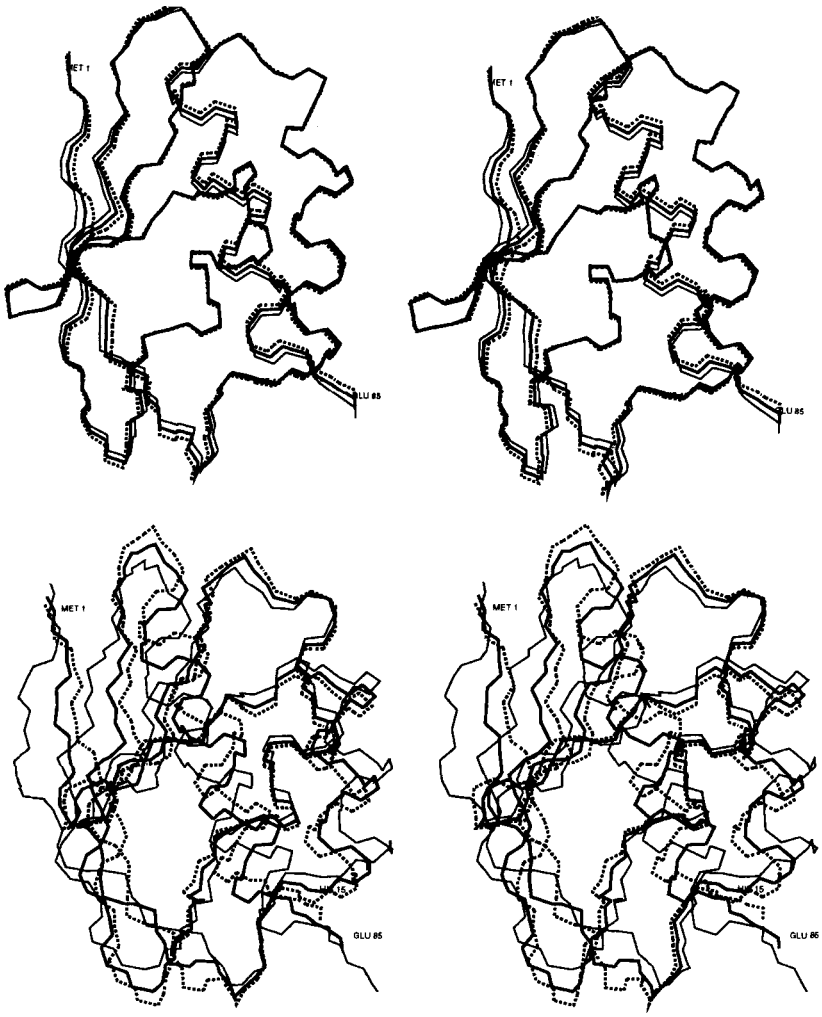

Fig. 6. Stereo representation of the backbone structures corresponding to minimum (*thin line*), maximum (*dashed line*), and average (*wide line*) position along the first eigenvector from ED sampling. **Top:** Corresponds to free MD. **Bottom:** Corresponds to ED sampling.

that also for HPr, a significant gain in the rate in which the essential subspace is filled during simulation is attained with respect to usual MD.

From a structural and energetical point of view, the clusters produced by ED sampling show similar properties as those from MD and time-averaged distance restrained MD. This makes the set of structures obtained by ED sampling equally acceptable as the smaller cluster from free MD or an even smaller cluster (e.g., from distance restrained simulations, for which we showed that motions inside the essential subspace may be damped[21]). So far, HPr is the only protein on which the ED sampling technique has been applied, but there is no reason to assume that other proteins will show significantly different dynamic behaviour.

During distance restrained energy minimisation, positions in the (backbone) essential subspace change only marginally (data not shown). This, together with the observation that violations of the experimental data are small, means that the averages of most of the experimentally derived distances are hardly affected by motions in the essential subspace. In a previous study,[21] we showed that distance restraining during MD reduces fluctuations in the essential subspace, due to a few restrained proton pairs, whose distances were affected by fluctuations in this subspace. Here, we demonstrate that fluctuations within the essential subspace are possible without violating the large majority of experimental restraints. In the graph of the configurational volume (Fig. 2), the curve corresponding to free MD is hardly any steeper than that obtained for time-averaged distance restrained MD. This is mainly because the graph consists of a combination of three pieces of simulation, started from different distance geometry structures. In general, a time-averaged distance restrained MD simulation can be expected to sample a smaller fraction of the configurational volume than a free MD simulation of comparable length would. This suggests that the reduced fluctuation in the essential subspace during distance restrained MD is mainly due to the attempt to fulfil all distances simultaneously (or averaged over too short time scales in time-averaged distance-restrained MD). The small number of distances which on average exceed their upper limits by more than 1.0 Å, are among the largest distances derived

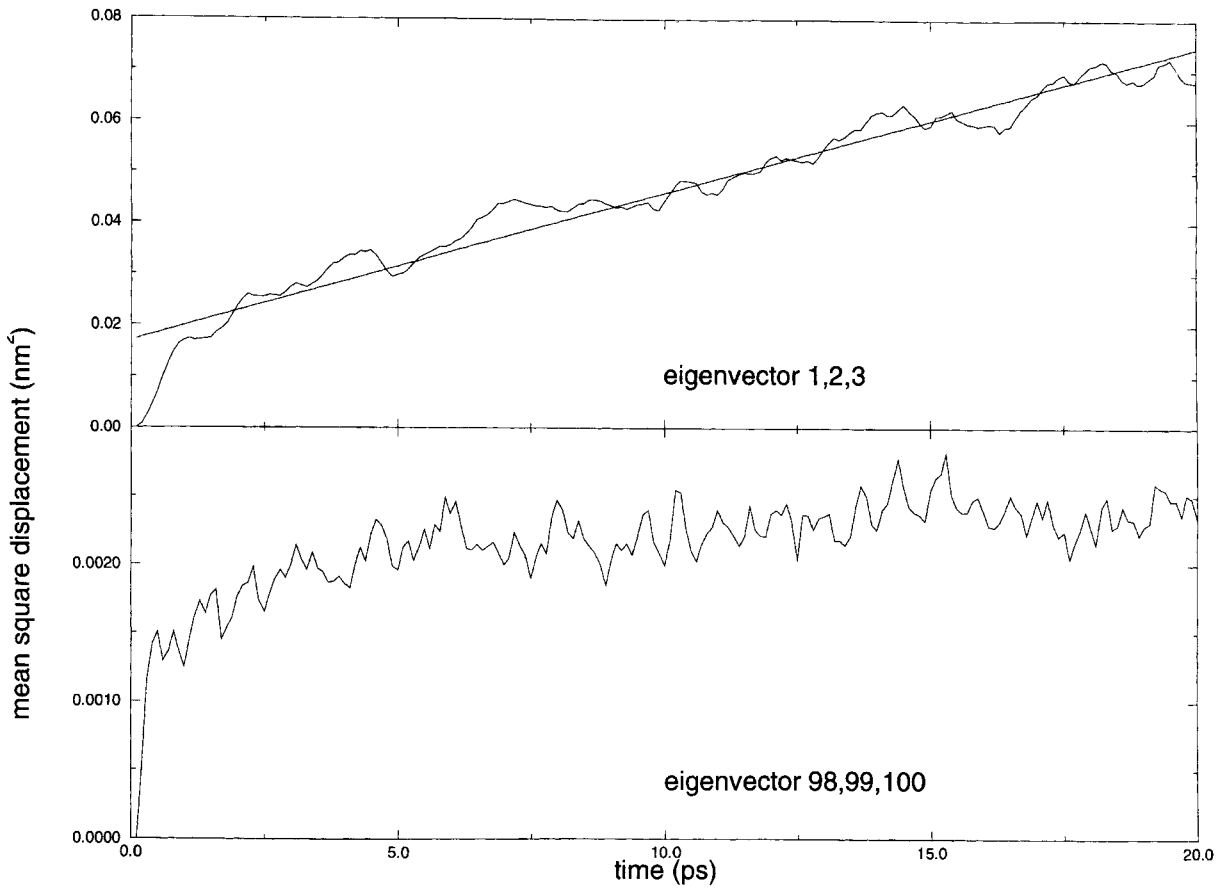

Fig. 7. Mean square displacement along eigenvectors (in nm^2) averaged over 80 pieces of 20 ps of free simulation. **Top:** Corresponds to the displacements along the first three eigenvectors. **Bottom:** Corresponds to the displacements along three near-constraints.

from the set of observed NOEs. These NOE's are the weakest observable peaks in the NOE spectra and are therefore most susceptible to misinterpretation. Therefore we conclude that there is no good reason, based on NMR data, to favor the set of structures obtained from time-averaged restrained MD refinement over the present set of structures, as obtained from our ED sampling algorithm. Moreover, the few remaining violations decrease as the ED sampling proceeds (Fig. 4), suggesting that averaging over an even larger set of structures could enhance correspondence with the experimental data further.

Other analyses, like the calculation of the configurational volume sampled (Fig. 2) and comparison of different sets of eigenvectors (Fig. 5) suggest that the sampling is not complete yet. In a peptide of 13 residues, we observed that the curves in the plot of the configurational volume leveled off after a sampling time corresponding to 2 ns of free MD,[12] indicating an almost exhaustive sampling. Although the slope of the curve of the volume for ED sampling does decrease with time for HPr, it is not approaching an asymptotic value yet. This suggests that the sampling is still not complete in this subspace and that only a small fraction of all possible configurations has been visited.

The fact that eigenvectors extracted from two independent MD simulations are as similar to each other as are eigenvectors from two independent ED sampling runs (Fig. 5), is another indication for this conclusion. The description of the essential subspace has approximately converged, even in a simulation time of a few hundred ps, although individual directions (eigenvectors) spanning this subspace are not fully converged yet. Even in a relatively extensive sampling, complete convergence is not yet reached. The overlap between two sets of eigenvectors, measured as the cumulative mean square inner product between a subset of (10) eigenvectors of one set with all eigenvectors out of the other set (Fig. 5), shows that the largest contribution is always obtained from the first part of the second set. In all graphs, more than 90% of the overlap is found between the subspace spanned by the first ten eigenvectors of one set and less than 10% of the (largest eigenvalue) eigenvectors of the other set. It is important to note that in all compared sets, an amount of noise is present. In a pair comparison as presented in Fig. 5,

this will have a more serious effect than in other forms of comparison.

The dynamic behaviour (in the essential subspace) during the ED sampling procedure is affected by the presence of constraint forces. Therefore, the density of sampled configurations cannot be used directly for exact thermodynamic evaluations. However, the dynamic behaviour of essential coordinates (Fig. 7) appears to be a form of diffusion, indicating that no significant free energy gradients exist in the essential subspace. Hence, a homogeneous equilibrium density distribution in the available essential subspace (as obtained by ED sampling) is to be expected. The sampling of the essential subspace of HPr as presented in this paper is not complete enough to provide insight into the details of the free energy surface in the complete essential subspace. The results suggest however, that the sampled fraction is a rather flat surface with at most many small ($< kT$) local minima. In a previous paper we showed[12] that for a small peptide the borders of this space are well defined and steep. For HPr, the borders seem less well localized and rather soft in some directions, leaving routes for denaturation.

ACKNOWLEDGMENTS

We thank Dr. G. T. Robillard for many stimulating discussions, and Daan van Aalten for support with a number of analysis programs.

REFERENCES

1. Postma, P.W., Lengeler, J.W., Jacobson, G.R. Phosphoenolpyruvate: Carbohydrate phosphotransferase systems of bacteria. Microbiol. Rev. 57:543–594, 1993.
2. van Nuland, N.A.J., Hangyi, I.W., van Schaik, R.C., Berendsen, H.J.C., van Gunsteren, W.F., Scheek, R.M., Robillard, G.T. The high-resolution structure of the histidine-containing phosphocarrier protein HPr from *Escherichia coli* determined by restrained molecular dynamics from NMR-NOE data. J. Mol. Biol. 237:544–559, 1994.
3. Jia, Z., Quail, J.W., Waygood, E.B., Delbaere, L.T.J. The 2.0 Å-resolution structure of Escherichia coli histidine-containing phosphocarrier protein HPr: A redetermination. J. Biol. Chem. 268:22490–22501, 1993.
4. Amadei, A., Linssen, A.B.M., Berendsen, H.J.C. Essential dynamics of proteins. Proteins 17:412–425, 1993.
5. Garcia, A.E. Large-amplitude nonlinear motions in proteins. Phys. Rev. Lett. 68:2696–2699, 1992.
6. van Aalten, D.M.F., Findlay, J.B.C., Amadei, A., Berendsen, H.J.C. Essential dynamics of the cellular retinol binding protein- evidence for ligand induced conformational changes. Prot. Eng. 8:1129–1136, 1995.
7. van Aalten, D.M.F., Amadei, A., Bywater, R., Findlay, J.B.C., Berendsen, H.J.C., Sander, C., Stouten, P.F.W. A comparison of structural and dynamic properties of different simulation methods applied to SH3. Biophys. J. 70: 684–692, 1996.
8. Clarage, J.B., Romo, T., Andrews, B.K., Pettitt, B.M. Jr., G.N.P. A sampling problem in molecular dynamics simulations of macromolecules. Proc. Natl. Acad. Sci. USA 92: 3288–3292, 1995.
9. Balsera, M.A., Wriggers, W., Oono, Y., Schulten, K. Principal component analysis and long time protein dynamics. J. Phys. Chem. 100:2567–2572, 1996.
10. Janezic, D., Venable, M., Brooks, B.R. Harmonic analysis of large systems. Comparison with molecular dynamics. J. Comp. Chem. 16:1554–1566, 1995.
11. Amadei, A., Linssen, A.B.M., de Groot, B.L., van Aalten, D.M.F., Berendsen, H.J.C. An efficient method for sampling the essential subspace of proteins. J. Biomol. Str. Dyn. 13:615–626, 1996.
12. de Groot, B.L., Amadei, A., van Aalten, D.M.F., Berendsen, H.J.C. Towards an exhaustive sampling of the configurational spaces of the two forms of the peptide hormone guanylin. J. Biomol. Str. Dyn. 13:741–751, 1996.
13. van Nuland, N.A.J., Boelens, R., Scheek, R.M., Robillard, G.T. High resolution structure of the phosphorylated form of he histidine-containing phosphocarrier protein HPr from *Escherichia coli* determined by restrained molecular dynamics from NMR-NOE data. J. Mol. Biol. 246:180–193, 1995.
14. Jia, Z., Vandonselaar, M., Quail, J.W., Delbaere, L.T.J. Active-center torsion angle strain revealed in 1.6 Å-resolution structure of histidine-containing phosphocarrier protein. Nature 361:94–97, 1993.
15. van Nuland, N.A.J., Wiersma, J.A., van der Spoel, D., de Groot, B.L., Scheek, R.M., Robillard, G.T. Phosphorylation-induced torsion-angle strain in the active center of HPr, detected by NMR and restrained molecular dynamics refinement. Prot. Sci. 5:442–446, 1996.
16. van Gunsteren, W.F., Berendsen, H.J.C. "Gromos Manual. BIOMOS, Biomolecular Software." The Netherlands: Laboratory of Physical Chemistry, University of Groningen, 1987.
17. Vriend, G. WHAT IF: A molecular modeling and drug design program. J. Mol. Graph. 8:52–56, 1990.
18. Kabsch, W., Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637, 1983.
19. Laskowski, R.A., MacArthur, M., Moss, D.S., Thornton, J.M. PROCHECK: A program to check the stereochemical quality of protein structures. J. Appl Crystalloqr 26:283–291, 1993.
20. van der Spoel, D., Berendsen, H.J.C., van Buuren, A.R., Apol, E., Meulenhoff, P.J., Sijbers, A.L.T.M., van Drunen, R. Gromacs User Manual. Nijenborgh 4, 9747 AG Groningen, The Netherlands. Internet: http: rugmdO.chem.rug.nl/~gmx 1995.
21. Scheek, R.M., van Nuland, N.A.J., de Groot, B.L., Amadei, A. Structure from NMR and molecular dynamics: distance restraining inhibits motion in the essential subspace. J. Biomol. NMR. 6:106–111, 1995.
22. Ramachandran, G.N., Ramakrishnan, C., Sasisekharan, V. Stereochemistry of polypeptide chain configurations. J. Mol. Biol. 7:95–99, 1963.