Protein Dynamics Derived from Clusters of Crystal Structures

D. M. F. van Aalten,* D. A. Conn,* B. L. de Groot,# H. J. C. Berendsen,# J. B. C. Findlay,* and A. Amadei#

*Department of Biochemistry and Molecular Biology, University of Leeds, Leeds LS2 9JT, England, and #Groningen Biomolecular Sciences and Biotechnology Institute, Department of Biophysical Chemistry, the University of Groningen, 9747 AG Groningen, The Netherlands

ABSTRACT   A method is presented to mathematically extract concerted structural transitions in proteins from collections of crystal structures. The "essential dynamics" procedure is used to filter out small-amplitude fluctuations from such a set of structures; the remaining large conformational changes describe motions such as those important for the uptake/release of substrate/ligand and in catalytic reactions. The method is applied to sets of x-ray structures for a number of proteins, and the results are compared with the results from essential dynamics as applied to molecular dynamics simulations of those proteins. A significant degree of similarity is found, thereby providing a direct experimental basis for the application of such simulations to the description of large concerted motions in proteins.

INTRODUCTION

A relationship between protein function and flexibility/dynamic characteristics was postulated before the first protein structure was even elucidated (Pauling, 1948). Using NMR and other forms of spectroscopy, it is possible to obtain some information on the motions of proteins (Nicholson et al., 1995; Hage et al., 1996). However, there is no experimental method for following the motion of every atom in a protein as a function of time, although with the advent of time-resolved crystallography a step has been made in this direction (Moffat, 1989; Bolduc et al., 1995; Genick et al., 1996). In most cases, computer simulation methods are used to obtain rough impressions of the motions that are possible in a given structure. Although these simulation methods have come of age (Berendsen, 1996), there are still a few key problems: 1) computer-generated data are treated with suspicion by many scientists, because they are not hard experimental observations; 2) simulations have the tendency to be restricted to small areas of the full configurational space of the protein (Clarage et al., 1995; Balsera et al., 1996); 3) interpretations of simulations are prone to subjectivity, as it is always possible to observe something that fits a particular hypothesis; and 4) large amounts of computer time are needed to properly simulate even small proteins (i.e., <25 kDa); the larger proteins are, at present, virtually impossible to simulate.

The recently introduced essential dynamics (ED) method (Amadei et al., 1993) (similar to principal components/multivariant analysis; Diamond, 1974; Garcia, 1992; Clarage et al., 1995) is able to extract the large (biologically significant) concerted motions from a molecular dynamics (MD) simulation. All relevant conformational states can be

described by only a few degrees of freedom. These essential degrees of freedom allow us to focus on the motions important for protein function, facilitating targeted mutagenesis studies aimed at affecting protein dynamics (van Aalten et al., 1996c), more direct comparison with experimental results (van Aalten et al., 1995, 1996b), and a full exploration of the relevant protein configurational space (Amadei et al., 1996; de Groot et al., 1996a, b). However, concerted motions revealed by this approach are calculated from MD simulations, and as such require experimental verification.

In the past few years, the Protein Data Bank (PDB) protein structure database (Bernstein et al., 1977) has been expanded rapidly by the addition of crystal structures of new proteins, and of different crystal structures of proteins already in the database. Proteins are crystallized in mutated forms, with different ligands, or under different conditions, all leading to slight conformational changes, restricted by the mechanical framework defined by the protein structure (and by the crystal packing environment; Phillips, 1990). Here we show that it is possible to combine the structural variations in thermally accessible conformations in a crystal environment into a formal description of large concerted movements of atoms by using the essential dynamics method. The results reveal a pattern of mobility similar to that derived from MD simulations started from a single crystal structure.

METHODS

Essential dynamics (similar to the single value decomposition method; Garcia, 1992; Romo et al., 1995) is based on the diagonalization of the covariance matrix, built from atomic fluctuations relative to their average positions:

$$C_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle \tag{1}$$

where x are the x, y, z coordinates of the atoms, $\langle x \rangle$ are the average positions of the coordinates, and the average is calculated over all structures, after they have been superimposed on a reference structure to remove overall transla-

tional and rotational motion. Here only Cα atoms are used, as it has been shown that this subset of atoms captures most of the conformational changes in the protein (Amadei et al., 1993; van Aalten et al., 1995). This covariance matrix is then diagonalized, yielding a set of eigenvalues and eigenvectors. The eigenvectors are directions in a $3N$-dimensional space (where N is the number of atoms), and motions along single eigenvectors correspond to concerted fluctuations of atoms. The eigenvalues represent the total mean square fluctuation of the system along the corresponding eigenvectors. The eigenvectors are sorted by the size of their corresponding eigenvalues, the "first" eigenvector being the one with the largest eigenvalue. In the case of proteins, there are always only a few ("essential") eigenvectors with large eigenvalues. Therefore the overall internal motion of the protein can be adequately described using only a few degrees of freedom (Amadei et al., 1993; van Aalten et al., 1995, 1996a, b).

The position of a structure along an eigenvector may be obtained by projection

$$q_l = (\vec{x} - \langle \vec{x} \rangle) \cdot \vec{\eta}_l \tag{2}$$

where $\vec{x}$ is a structure, $\langle \vec{x} \rangle$ is the average structure, $\vec{\eta}_l$ is an eigenvector, and q_l is the displacement of the structure along the lth eigenvector with respect to the average structure. From the definition of the eigenvectors, it is also possible to obtain the 3D structure corresponding to a displacement along a single eigenvector:

$$\vec{x} = q_l \vec{\eta}_l + \langle \vec{x} \rangle \tag{3}$$

X-ray structures were extracted from the PDB database for a test set of seven protein families (see Table 1). MD simulations of one protein from each family were performed

TABLE 1 List of PDB codes for the protein families included in the analysis

Family	PDB entries
FABP	1crb 1cbq 1cbr 1cbs 1opa 1opb 1adl 1alb 1hmr 1hms 1hmt 1icm 1ifb 1lib 1lic 1lid 1lie 1lif 2hmb 2ifb 1ftp
PL	1ayp 1bp2 1p2p 1poa 1pob 1pod 1poe 1psh 2bpp 3bp2 4p2p
MYO	1spe 1vxa 1vxb 1vxc 1vxd 1vxe 1vxf 1vxg 1vxh 1emy 1mbc 1mbd 1mbi 1mbn 1mbo 1myg 1pmb 1swm 2cmm 2mya 2myb 2myc 2myd 2mye 4mbn 5mbn
RAS	121p 1agp 1gnp 1gnq 1gnr 1q21 221p 2q21 421p 4q21 521p 5p21 621p 6q21 721p 821p
LYS	132l 135l 1ghl 1hel 1hew 1lma 1lsa 1lsb 1lsc 1lsd 1lse 1lsf 1lys 1lyz 1lz3 1lza 1lzb 1lzc 1lzh 1lzt 1lzy 1tew 2lym 2lyz 2lz2 2lzh 2lzt 3lyz 3lz2 4lyt 4lyz 5lyt 5lyz 6lyt 6lyz 7lyz 8lyz
BARN	1ban 1bao 1bgs 1bne 1bnf 1bng 1bni 1bnj 1bns 1brg 1brh 1bri 1brj 1brk 1brn 1brs 1bsa 1bsb 1bsc 1bsd 1bse 1rnb
HIV	1cpi 1dif 1hpx 1hsg 1pro 1aaq 1hbv 1hih 1hiv 1hos 1hps 1hpv 1hte 1htf 1htg 1hvi 1hvj 1hvk 1hvl 1hvr 1sbg 4hvp 4phv 5hvp 7hvp 8hvp 9hvp
AAT	1arg 1oxo 1oxp 1aaw 1ama 1amq 1amr 1ams 1ars 1art 1asa 1asd 1ase 1asl 1asm 1asn 1map 1maq 1tar 1tas 1tat 7aat 8aat 9aat

For proteins that crystallized in nonbiologically relevant multimeric forms, each chain was included as a separate structure.

if they were not already available from previous studies. The test set consisted of the families of fatty acid binding proteins (FABP), phospholipase A2 (PL), myoglobin (MYO), ras-p21 (RAS), egg white lysozyme (LYS), and barnase (BARN). Two additional protein families were used for which no MD simulation was performed: the HIV proteases (HIV) and aspartate aminotransferases (AAT). PDB entries were selected based on the following criteria: reasonable sequence similarity (i.e., good alignment with the rest of the proteins in the same set), the existence of more than 15 structures, X-ray structures only, mutant structures not included (except for RAS). The test set of proteins was chosen such that various causes of structural changes were included, i.e., structural variation due to sequence divergence (e.g., FABP), different crystallisation conditions like pH and temperature (e.g., MYO), different crystal forms (e.g., LYS), different ligands (e.g., FABP, HIV, AAT), and mutated forms (RAS, BARN). Crystal structures that were very different from the average structure in each set were deleted to prevent bias toward large motions based on a single outlier: the RMSD (root mean square deviation) of each structure with respect to the average structure of the whole set was calculated. The mean and standard deviation (σ) of all of these RMSDs was computed. All structures (but never more than two) that deviated more than 2.5σ from the average RMSD was deleted from the set. The same procedure was repeated with the resulting set, until all structures were within the 2.5σ limit, or until $\sigma < 0.1$ Å, in which case deleting further structures would result in a set with almost no structural variation. For each test protein, sequences from the selected structures were aligned using the Clustal-W (Thompson et al., 1994) program. Gaps in the sequence (indicating local structural divergence) were removed by deleting the inserted residues plus 2 on either side of the gap from the relevant protein structures. This resulted in an ensemble of superposable structural fragments, which were then directly used for ED analysis (CRY-ED).

MD simulations for comparison with CRY-ED were performed with the GROMOS (van Gunsteren and Berendsen, 1987) and GROMACS (van der Spoel et al., 1995; Berendsen et al., 1995) suite of programs. Simulations were performed in full solvent (water) with periodic boundary conditions. The simulations of FABP (300 ps) (van Aalten et al., 1996b), LYS (1 ns) (Smith et al., 1995), and RAS (300 ps) (Mello et al., 1997) have been described before. The simulations of PL (400 ps), MYO (1.5 ns), and BARN (300 ps) were performed with similar parameters (for details, see van Aalten et al., 1995, 1996b).

RESULTS

The results of CRY-ED on the 27 FABP crystal structures are illustrated in Fig. 1, and a list of eigenvalues is given in Table 2. The superposition of the crystal structures shows a rather noisy cloud of conformations. By studying the concerted atomic displacements described by the first eigen-
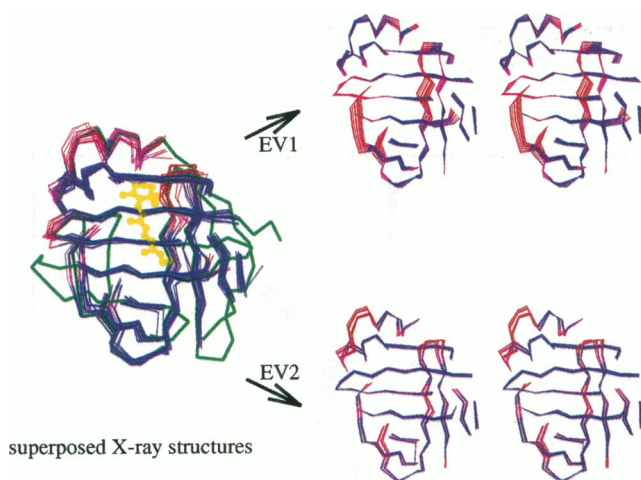
superposed X-ray structures

FIGURE 1   The results of the essential dynamics procedure. Superposition of all crystal structures in the FABP set (*left*), colored by average normalized B-factors (low B-factor = blue; high B-factor = red), and stereo pictures representing the projection of the same structures onto eigenvector 1 (*upper right*) and 2 (*lower right*) calculated by CRY-ED, colored by mobility as described by the eigenvector (i.e., regions with low mobility are colored blue; highly mobile regions are colored red).

TABLE 2   List of eigenvalues resulting from the application of ED on the crystal structures in the test sets

Family	EV1	EV2	EV3	EV5	EV10	EV15
FABP	0.28	0.14	0.088	0.034	0.0036	0.00098
PL	0.17	0.099	0.022	0.011	0.0036	0.0015
MYO	0.035	0.013	0.0098	0.0068	0.0017	0.00065
RAS	0.39	0.20	0.12	0.046	0.0074	0.0021
LYS	0.032	0.026	0.019	0.0076	0.0035	0.0021
BARN	0.018	0.013	0.0067	0.0038	0.0012	0.00056
HIV	0.11	0.082	0.026	0.0099	0.0046	0.0029
AAT	1.8	0.77	0.047	0.018	0.0073	0.0045

Eigenvalue 1 (EV1 in the table) corresponds to the largest concerted motion of atoms (described by eigenvector 1). Eigenvalues are in nm^2.

vectors, we filtered out the small uncorrelated structural variations. The structural fluctuations seem to correlate with the average B-factors determined from the crystal structures. The first two eigenvectors (i.e., the eigenvectors with the largest eigenvalues, thus the most dominant concerted fluctuations of atoms) describe two distinct movements. In eigenvector 1, one of the two helices at the entrance to the binding pocket moves away from the rest of the protein. In the second eigenvector a correlated displacement of the D and E strands is observed. The displacements along the two eigenvectors together seem to create an opening toward the ligand binding pocket. Similar motions were observed in two molecular dynamics simulations, where the ligand was seen to move toward this opening (van Aalten et al., 1996b; Zanotti et al., 1994).

Comparable results were obtained for the other proteins in the test set. Application of CRY-ED provided a formal description of large concerted structural rearrangements, captured in a few eigenvectors. Such structural rearrange-

ments, which were observed to be centered around the substrate or ligand-binding site, are likely to be linked to the functional properties of the proteins. Internal motions have been demonstrated for PL by fluorescence studies (Kuipers et al., 1991), involving regions containing engineered tryptophans, which are also highly mobile in the CRY-ED analysis. CRY-ED was able to detect rigid body motions of secondary structure elements, centered around the heme-binding pocket; concerted motions have also been found experimentally in myoglobin (Srajer et al., 1996; Richard et al., 1992). Fluctuations of two loops covering the nucleotide-binding site in RAS, detected by NMR (Kraulis et al., 1994), were also found by CRY-ED. Similarly, NMR experiments predicted flexible regions for LYS (Buck et al., 1995) and BARN (Meiering et al., 1993), which are reproduced by the CRY-ED experiments. For HIV, concerted motions of the well-characterized flaps and other regions close to the substrate binding site were observed, in agreement with previously published NMR (Nicholson et al., 1995) and simulation data (Collins et al., 1995). AAT, a protein too big for long simulations on present-day computers, shows large hinge-bending of domains around the substrate-binding site, as previously observed (Moser et al., 1994).

Unfortunately, for most proteins whose structures have been solved, too few structures are available to obtain the essential eigenvectors directly by the ED approach described above. However, MD simulations analyzed with ED (MD-ED) can be used to obtain a rough approximation of these vectors. Here we quantitatively compare the ED eigenvectors revealed by the analysis of crystallographic data (CRY-ED) with those obtained from MD (MD-ED).

Fig. 2 *A* demonstrates the method of comparison: MD-ED eigenvectors are projected onto the first three CRY-ED eigenvectors by calculating the cumulative square inner product. If these two sets of eigenvectors are very different, a nearly straight line will be obtained. The figure shows that the largest part of the overlap is concentrated in the first few MD-ED eigenvectors, which together form the "essential subspace" (Amadei et al., 1993). Thus the large concerted motions derived from CRY-ED are similar to those found from MD-ED. This is further quantified by inner products representing the overlap of the first few MD-ED eigenvectors (5% of the total number) of the protein with the first three CRY-ED eigenvectors, listed in Table 3. It appears that on average, the first three CRY-ED eigenvectors are contained for ~50% (overlap of 0.5) in the first 5% of MD-ED eigenvectors. There are four reasons why this overlap is not 100%: 1) Even a 1-ns simulation does not provide a complete sampling of the essential subspace (Clarage et al., 1995; Balsera et al., 1996). 2) An MD simulation represents the protein in solvent without any contacts with neighboring proteins; such contacts may play a role in a protein crystal (Zhang et al., 1995). 3) The covariance matrix is built from an ensemble of thousands of structures in the case of MD-ED, whereas for CRY-ED only a few tens of crystal structures are used. 4) Structural
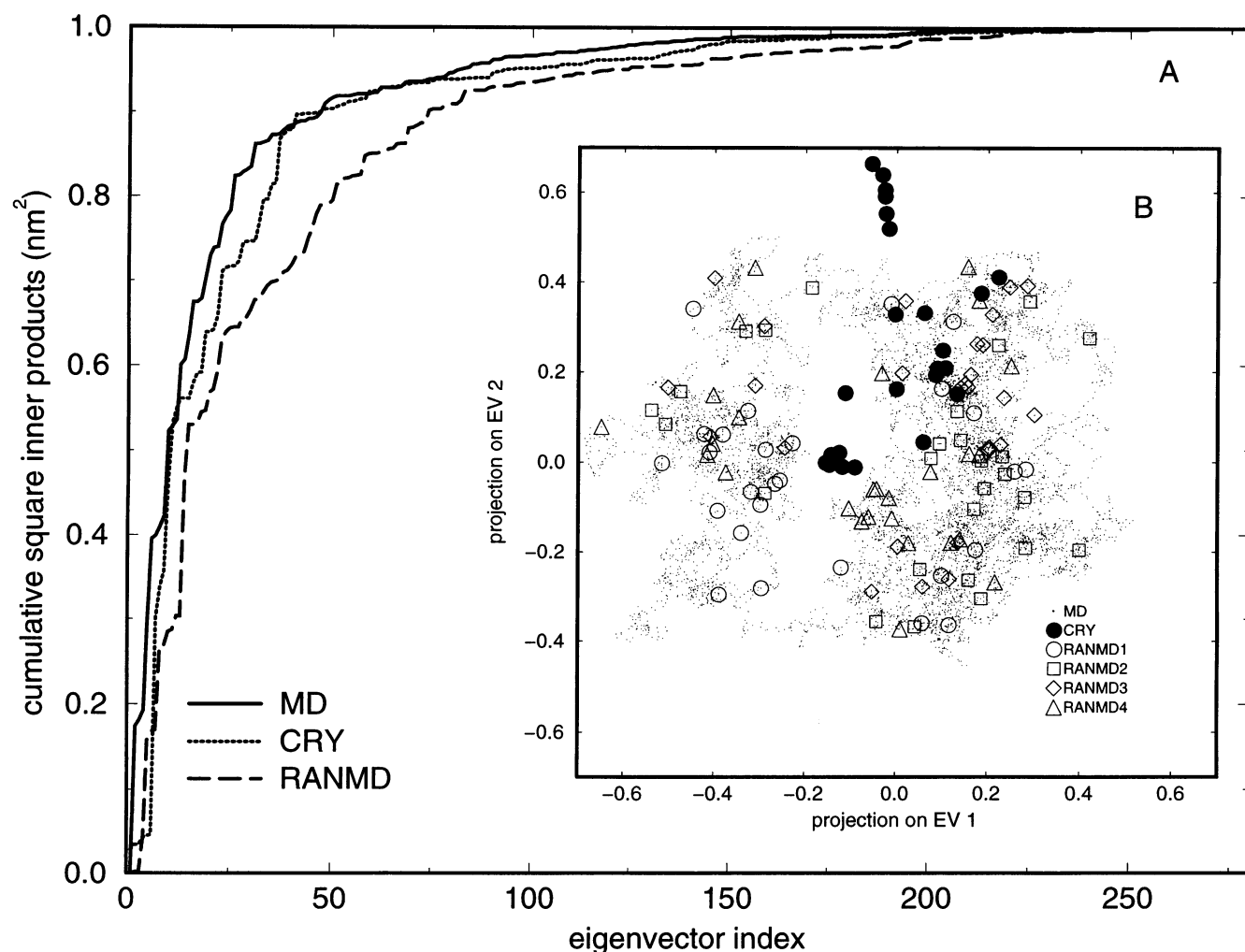
FIGURE 2   Projection of eigenvectors and trajectories. (A) Cumulative inner products from the projection of the MD-ED eigenvectors onto the first three CRY-ED eigenvectors, for FABP. EV1, EV2, EV3 are CRY-ED eigenvectors 1, 2, and 3, respectively. (B) Projection of the MD structures (MD), the crystal structures (CRY), and the four "random" (see Table 1) sets of MD structures (RANMD1-4) of FABP onto the plane defined by MD-ED eigenvectors 1 and 2.

variation in the ensemble of crystal structures is rather limited compared to that in the MD ensemble. The importance of the last two effects was investigated by randomly selecting structures (the same number and as much as possible the same spread as for the crystal structures) from the MD simulation (see Table 3). ED was then performed on these structures, and the resulting eigenvectors were compared to those calculated from the full MD simulation (Table 3). It appears that reducing the structural variability and the number of structures used reduces the overlap considerably, to ~70%. Thus considering the possible additional effect of crystal contacts, and the MD sampling problem, 50% indicates a significant similarity.

Fig. 2 B shows the projection of the three sets of structures (normal MD, crystal structures, and the randomly selected MD structures, for FABP) onto the first two MD-ED eigenvectors. The crystal structures show a considerable spread in projection onto these eigenvectors. This

implies that the structural repertoire of the experimental set is not restricted to a limited number of conformations.

DISCUSSION

Summarizing, it is possible to derive a formal description of concerted structural fluctuations of atoms in a protein from just a few tens of crystal structures. The "essential" motions found appear to have a significant similarity to those obtained from MD simulations of the same proteins, thus validating the results coming from such simulation techniques. Interestingly, the amount of overlap between the CRY-ED and MD-ED eigenvectors does not seem to depend on the cause of structural variation in the cluster of crystal structures. Both RAS and BARN contain mainly mutant crystal structures, whereas structural variation in FABP, for instance, is mainly caused by sequence diversity.

TABLE 3 Quantitative analysis of similarities between the eigenvectors derived from MD and CRY-ED

	RMS deviation (Å)					MD on CRY			MD on RANMD		
	$\langle$MD$\rangle$	$\langle$CRY$\rangle$	σCRY	#dim	#CRY	ev1	ev2	ev3	ev1	ev2	ev3
FABP	0.68	0.82	0.18	14	27	0.61	0.56	0.47	0.93 (0.01)	0.91 (0.02)	0.89 (0.03)
PL	0.86	0.63	0.09	14	20	0.41	0.41	0.26	0.68 (0.19)	0.58 (0.06)	0.41 (0.03)
MYO	0.87	0.23	0.11	22	28	0.59	0.46	0.39	0.47 (0.04)	0.38 (0.12)	0.29 (0.10)
RAS	1.59	0.67	0.34	24	19	0.46	0.49	0.41	0.86 (0.02)	0.78 (0.09)	0.57 (0.20)
LYS	0.85	0.33	0.12	19	41	0.49	0.36	0.43	0.61 (0.06)	0.58 (0.07)	0.43 (0.07)
BARN	0.93	0.26	0.07	14	63	0.61	0.69	0.33	0.42 (0.12)	0.34 (0.10)	0.25 (0.05)
Mean	0.96	0.49	0.15	18	33	0.53	0.49	0.38	0.66 (0.07)	0.60 (0.08)	0.47 (0.08)

For each protein in the test set, the following are listed: average root mean square deviation (RMSD) with respect to the average structure for the MD and crystal structures; σCRY, standard deviation in this RMSD for the crystal structures, #dim, number of eigenvectors (5% of the total) taken to represent the essential subspace of the MD simulation; #CRY, = number of crystal structures in the test set; MD on CRY, cumulative squared inner products between the first #dim eigenvectors of the MD simulation and eigenvectors 1, 2, and 3 (ev1, ev2, ev3, respectively) from CRY-ED; MD on RANMD, cumulative squared inner products between the first #dim eigenvectors of the MD simulation and eigenvectors 1, 2, and 3 (ev1, ev2, ev3, respectively) derived from a set of randomly selected structures from the MD simulation. These structures were chosen by using a reference frame from the MD simulation and selecting MD frames such that the spread in the RMSDs was the same as for the set of crystal structures. This procedure was repeated four times, with different MD structures as reference (equally spread out over the simulation) to increase statistics. The standard deviation in the overlap between the eigenvectors from these four RANMD sets and the MD simulation is listed between brackets. The last line in the table (Mean) lists the means of the columns.

Collections of crystal structures of the same protein have been used before in many ways to investigate biologically important structural changes. The Diamond plot (Diamond, 1974) shows variation in a set of structures along the main axes of structural displacement, in a way similar to that presented here. In general, there have been many reports of comparison of a few crystal structures by conventional structural superposition (e.g., Sondek et al., 1994; Moser et al., 1994; Zhang et al., 1995) and domain-searching algorithms (Nichols et al., 1995). The hinge-bending motion in a mutant bacteriophage T4 lysozyme (Faber and Matthews, 1990) is a well-known example of this kind. Furthermore, crystal structures have been sorted visually to yield a movie of conformational change depicting a reaction cycle (Vonrhein et al., 1995). The new approach presented here has many advantages: it is able to extract the large concerted conformational changes, thus eliminating small irrelevant structural changes; it provides a mathematical model for protein conformational changes, which enables us to describe biologically relevant conformational states by specifying only a few variables (the displacement along the essential eigenvectors). This opens up new directions in targeted site-directed mutagenesis (e.g., van Aalten et al., 1996c), or even in automatic docking and folding algorithms, which can now be based on eigenvectors derived from experimental rather than simulated structures.

REFERENCES

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

Amadei, A., A. B. M. Linssen, B. L. de Groot, D. M. F. van Aalten, and H. J. C. Berendsen. 1996. An efficient method for sampling the essential subspace of proteins. *J. Biomol. Struct. Dyn.* 13:615–625.

Balsera, M. A., W. Wriggers, Y. Oono, and K. Schulten. 1996. Principal component analysis and long-time protein dynamics. *J. Phys. Chem.* 100:2567–2572.

Berendsen, H. J. C. 1996. Bio-molecular dynamics comes of age. *Science.* 271:954–955.

Berendsen, H. J. C., D. van der Spoel, and R. van Drunen. 1995. Gromacs—a message-passing parallel molecular-dynamics implementation. *Comp. Phys. Comm.* 91:43–56.

Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, E. F. Meyer, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. 1977. The protein data bank: a computer based archival file for macromolecular structures. *J. Mol. Biol.* 112:535–542.

Bolduc, J. M., D. H. Dyer, W. G. Scott, P. Singer, R. M. Sweet, Jr., and B. L. Stoddard. 1995. Mutagenesis and laue structures of enzyme intermediates: isocitrate dehydrogenase. *Science.* 268:1312–1318.

Buck, M., J. Boyd, C. Redfield, D. A. MacKenzie, D. J. Jeenes, D. B. Archer, and C. M. Dobson. 1995. Structural determinants of protein dynamics: analysis of ^{15}N NMR relaxation measurements for mainchain and side-chain nuclei of hen egg white lysozyme. *Biochemistry.* 34:4041–4055.

Clarage, J. B., T. Romo, B. K. Andrews, and B. M. Pettitt. 1995. A sampling problem in molecular dynamics simulations of macromolecules. *Proc. Natl. Acad. Sci. USA.* 92:3288–3292.

Collins, J. R., S. K. Burt, and J. W. Erickson. 1995. Flap opening in hiv-1 protease simulated by "actived" molecular dynamics. *Nature Struct. Biol.* 4:334–338.

de Groot, B., A. Amadei, R. Scheek, N. van Nuland, and H. Berendsen. 1996a. An extended sampling of the configurational space of hpr from *E. coli. Proteins Struct. Funct. Genet.* 26:314–322.

de Groot, B. L., A. Amadei, D. M. F. van Aalten, and H. J. C. Berendsen. 1996b. Towards an exhaustive sampling of the configurational spaces of the two forms of the peptide hormone guanylin. *J. Biomol. Struct. Dyn.* 13:741–751.

Diamond, R. 1974. Real-space refinement of the structure of hen egg-white lysozyme. *J. Mol. Biol.* 82:371–391.

Faber, H. R., and B. W. Matthews. 1990. A mutant t4 lysozyme displays 5 different crystal conformations. *Nature.* 348:263–266.

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

Genick, U. K., G. E. O. Borgstahl, N. Kingman, Z. Ren, C. Pradervand, P. Burke, V. Srajer, T. Teng, W. Schildkamp, D. E. McRee, K. Moffat, and E. D. Getzoff. 1997. Millisecond time-resolved laue crystallography: structure of a protein photocycle intermediate. *Science.* 275:1471–1475.

Hage, W., M. Kim, H. Frei, and R. A. Mathies. 1996. Protein dynamics in the bacteriorhodopsin photocycle—a nanosecond step-scan FTIR investigation of the kl to l transition. *J. Phys. Chem.* 100:16026–16033.

Kraulis, P. J., P. J. Domaille, S. L. Campbell-Burk, T. van Aken, and E. D. Laue. 1994. Solution structure and dynamics of ras-p21-gdp determined by heteronuclear three- and four-dimensional NMR spectroscopy. *Biochemistry.* 33:3515–3531.

Kuipers, O. P., M. Vincent, J.-C. Brochon, H. M. Verheij, G. H. de Haas, and J. Gallay. 1991. Insight into the conformational dynamics of specific regions of porcine pancreatic phospholipase a2 from a time-resolved fluorescence study of a genetically inserted single tryptophan. *Biochemistry.* 30:8771–8785.

Meiering, E. M., M. Bycroft, M. J. Lubienski, and A. R. Fersht. 1993. Structure and dynamics of barnase complexed with 3'-GMP studied by NMR spectroscopy. *Biochemistry.* 32:10975–10987.

Mello, L. V., D. M. F. van Aalten, and J. B. C. Findlay. 1997. Comparison of RAS-p21 bound to GDP and GTP: differences in protein and ligand dynamics. *Prot. Eng.* 10:381–387.

Moffat, K. 1989. Time-resolved macromolecular crystallography. *Annu. Rev. Biophys. Biophys. Chem.* 18:309–332.

Moser, M., U. Sauder, and J. N. Jansonius. 1994. Crystal structures of *E. coli* aspartate aminotransferase in two conformations. *J. Mol. Biol.* 239:285–305.

Nichols, W. L., G. D. Rose, L. F. T. Eyck, and B. H. Zimm. 1995. Rigid domains in proteins: an algorithmic approach to their identification. *Proteins Struct. Funct. Genet.* 23:38–48.

Nicholson, L. K., T. Yamazaki, D. A. Torchia, S. Grzesiek, A. Bax, S. J. Stahl, J. D. Kaufman, P. T. Wingfield, P. Y. S. Lam, P. K. Jadhav, N. Hodge, P. J. Domaille, and C.-H. Chang. 1995. Flexibility and function in hiv-1 protease. *Nature Struct. Biol.* 4:274–280.

Pauling, L. 1948. Nature of forces between large molecules of biological interest. *Nature.* 161:707–709.

Phillips, G. N. 1990. Comparison of the dynamics of myoglobin in different crystal forms. *Biophys. J.* 57:381–383.

Richard, L., L. Genberg, J. Deak, H.-L. Chiu, and R. J. D. Miller. 1992. Picosecond phase grating spectroscopy of hemoglobin and myoglobin: energetics and dynamics of global protein motion. *Biochemistry.* 31:10703–10715.

Romo, T. D., J. B. Clarage, D. C. Sorensen, and G. N. Phillips. 1995. Automatic identification of discrete substrates in proteins: singular value decomposition analysis of time-averaged crystallographic refinements. *Proteins Struct. Funct. Genet.* 22:311–321.

Smith, L. J., A. E. Mark, C. M. Dobson, and W. F. van Gunsteren. 1995. Comparison of MD simulations and NMR experiments from hen lysozyme—analysis of local fluctuations, cooperative motions and global changes. *Biochemistry.* 34:10918–10931.

Sondek, J., D. G. Lambright, J. P. Noel, H. E. Hamm, and P. B. Sigler. 1994. GTPase mechanism of g-proteins from the 1.7 angstrom crystal structure of transducin α-GDP. *Nature.* 372:276–279.

Srajer, V., T. Y. Teng, T. Ursby, C. Pradervand, Z. Ren, S. Adachi, W. Schildkamp, D. Bourgeois, M. Wulff, and K. Moffat. 1996. Photolysis of the carbon monoxide complex of myoglobin—nanosecond time-resolved crystallography. *Science.* 274:1726–1729.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. Clustal-w—improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673–4680.

van Aalten, D. M. F., A. Amadei, R. Bywater, J. B. C. Findlay, H. J. C. Berendsen, C. Sander, and P. F. W. Stouten. 1996a. A comparison of structural and dynamic properties of different simulation methods applied to sh3. *Biophys. J.* 70:684–692.

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

van Aalten, D. M. F., J. B. C. Findlay, A. Amadei, and H. J. C. Berendsen. 1996b. Essential dynamics of the cellular retinol binding protein—evidence for ligand induced conformational changes. *Protein Eng.* 8:1129–1135.

van Aalten, D. M. F., P. C. Jones, M. de Sousa, and J. B. C. Findlay. 1996c. Engineering protein mechanics: inhibition of concerted motions of the cellular retinol binding protein by site directed mutagenesis. *Protein Eng.* (in press).

van der Spoel, D., H. J. C. Berendsen, A. R. van Buuren, E. Apol, P. J. Meulenhoff, A. L. T. M. Sijbers, and R. van Drunen. 1995. Gromacs User Manual. Nijenborgh 4, 9747 AG Groningen, The Netherlands (Internet: http://rugmd0.chem.rug.nl/~gmx).

van Gunsteren, W. F., and H. J. C. Berendsen. 1987. Gromos Manual. BIOMOS, Biomolecular Software, Laboratory of Physical Chemistry, University of Groningen, Groningen, The Netherlands.

Vonrhein, C., G. J. Schlauderer, and G. E. Schulz. 1995. Movie of the structural changes during a catalytic cycle of nucleoside monophosphate kinases. *Structure.* 3:483–490.

Zanotti, G., L. Feltre, and P. Spadon. 1994. A possible route for the release of fatty acid from fatty acid binding protein. *Biochem. J.* 301:459–463.

Zhang, X., J. A. Wozniak, and B. W. Matthews. 1995. Protein flexibility and adaptability seen in 25 crystal forms of t4 lysozyme. *J. Mol. Biol.* 250:527–552.