JOURNAL OF THE ROYAL SOCIETY Interface

J. R. Soc. Interface (2005) **2**, 47–54
doi:10.1098/rsif.2004.0025
Published online 2 February 2005

Native-state dynamics of the ubiquitin family: implications for function and evolution

Neelan J. Marianayagam[†] and Sophie E. Jackson

Centre for Protein Engineering, Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK

Protein dynamics are integral to protein function. In recent years, the use of computer simulation to understand the molecular motions of proteins has become widespread. However, there are few such studies which compare the dynamics of proteins that are structurally and functionally related. In this study, we present native-state molecular dynamic simulations of four proteins which possess a ubiquitin-like fold. Three of these proteins are thought to have evolved from a common ancestral ubiquitin-like protein and have similarities in their function. A fourth protein, which is structurally homologous but which appears to have a different function, is also studied. Local fluctuations in the native state simulations are analysed, and conserved motions of the C-α backbone atoms are identified in residues which are important for function. In addition, the global dynamics of the proteins are analysed using the essential-dynamics method. This analysis reveals a slightly higher degree of conservation in dynamics for the three proteins which are functionally related. Both the global and local analyses illustrate how nature has optimized and conserved protein motions for specific biological activity within the ubiquitin family.

Keywords: essential dynamics; native-state simulations; ubiquitin family; protein function

1. INTRODUCTION

The dynamic properties of proteins are now well known to play important roles in protein function. Many different aspects of protein function can be affected by protein dynamics. For example, protein–protein recognition (Gohlke *et al.* 2004), protein–DNA interactions (Kalodimos *et al.* 2004) and enzyme–substrate binding and enzyme activity (Rasmussen *et al.* 1992; Vitagliano *et al.* 2002; Cui *et al.* 2004) are all determined, in part, by the conformational flexibility of the protein backbone as well as specific side chains. It is, therefore, important to characterize not only the structure of a protein but also its dynamic properties as well. While X-ray crystallography provides an excellent method for the determination of high-resolution structures, it generates a static picture of a protein and, in general, provides little information on protein dynamics. Experimentally, a number of different nuclear magnetic resonance (NMR) techniques have been used to obtain information on the molecular motions within proteins on several different time-scales. The number of proteins on which such studies can be performed, however, is limited. The use of computer simulations to probe protein motions, using existing structural information, is, therefore, proving extremely fruitful. However, in order to extract useful information about the dynamics observed during the course of these simulations, mathematical models must be employed. Two of these models are normal mode analysis (NMA; Brooks *et al.* 1995) and essential dynamics (Amadei *et al.* 1993).

In NMA, the motion of the protein is assumed to be harmonic. The technique looks mainly at vibrational motion while ignoring all other types of motion. In NMA, the potential energy function is approximated as a sum of quadratic terms, which describe atomic displacement (Brooks *et al.* 1995). The coefficients of these terms represent force constants, which can be put into a matrix. If one adds the atomic masses to this matrix, one can set up a matrix equation to calculate the vibrational modes of the molecule. This then becomes an eigenvalue problem. For a system with N atoms, there are $3N-6$ eigenvalues and eigenvectors, which specify the normal modes of the system. The advantage of NMA is that it gives us insight into macromolecular motion without the need of a molecular-dynamics (MD) trajectory. This makes it less computationally intensive than looking at native state dynamics using MD simulations.

The method of essential dynamics looks at the positional fluctuations of atoms rather than motion confined to a harmonic potential (Amadei *et al.* 1993). It is used in conjunction with MD simulations. This method divides the conformational space of a protein into two subspaces, an essential subspace and a physically constrained subspace (Amadei *et al.* 1993).

[†]Author for correspondence (neelanm@yahoo.com).

Received 11 August 2004
Accepted 1 December 2004

© 2005 The Royal Society

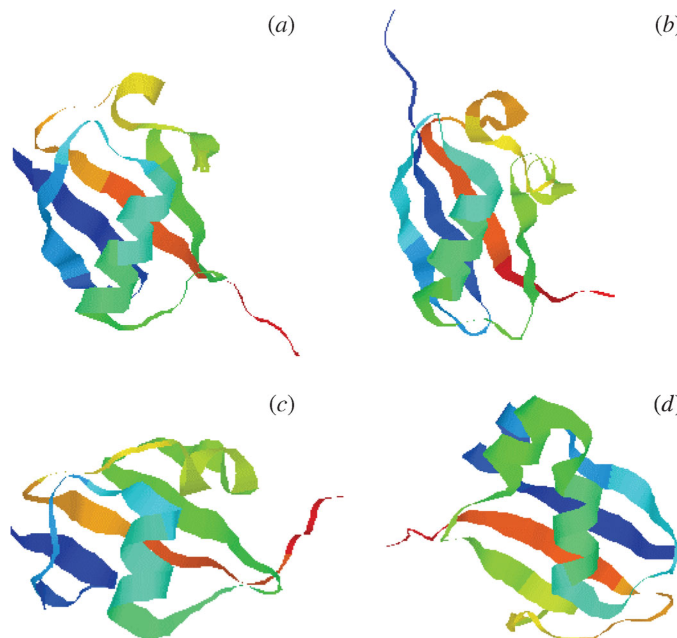

Figure 1. Structure of the four ubiquitin-like proteins: (*a*) ubiquitin, (*b*) UBX, (*c*) ThiS and (*d*) MoaD.

The essential subspace is described by the anharmonic motion of the positional fluctuations of the atoms. The motion in the remaining subspace is defined by a narrow Gaussian distribution.

The essential-dynamics method represents a principal-component analysis of the atomic fluctuations of the protein. The first step is the generation of non-mass weighted coordinate matrix. For an N-atom system, this will have $3N$ columns and at least $3N+1$ rows. This matrix, which we will call $\mathbf{A}$, represents the movement of atomic positions from an average value throughout the course of the simulation. The covariance matrix of $\mathbf{A}$, which we will call $\mathbf{C}$, is defined by the following equation:

$$\mathbf{C} = \mathbf{A}^{\mathrm{T}}\mathbf{A}, \tag{1.1}$$

where T is the transpose of the matrix. The transpose is found by exchanging the rows and columns of a matrix. The eigenvectors of the covariance matrix are the principal components. This then turns into an eigenvalue problem:

$$\mathbf{C}\boldsymbol{x} = \lambda\boldsymbol{x}, \tag{1.2}$$

where λ is the eigenvalue associated with the eigenvector $\boldsymbol{x}$. For an N-atom system, there are $3N$ eigenvectors and associated eigenvalues. Equation (1.2) can be simplified to the following:

$$(\mathbf{C} - \lambda\mathbf{I})\boldsymbol{x} = 0, \tag{1.3}$$

where $\mathbf{I}$ is the identity matrix. The solution to equation (1.3) can be obtained by diagonalizing the covariance matrix. The diagonal matrix, $\mathbf{D}$, of the covariance matrix is defined by the following:

$$\mathbf{D} = \mathbf{U}^{-1}\mathbf{C}\mathbf{U}. \tag{1.4}$$

The matrix $\mathbf{U}$ contains the eigenvectors, and $\mathbf{D}$ is a matrix of the corresponding eigenvalues. The eigenvector with the highest eigenvalue is considered the first principal component, the eigenvector with the second highest eigenvalue is considered the second principal component and so on. The eigenvectors represent the direction of motion, and the eigenvalues represent the amount of motion along the eigenvectors. The dynamics of a protein can thus be analysed by projecting its atomic motion during a MD simulation onto its first two to three principal components (Amadei *et al.* 1999*a*). Essential dynamics is a powerful tool for monitoring protein dynamics in phase space since the observed motion is unconstrained and represents the atomic fluctuations of the protein. Essential dynamics has been used to look at the native-state fluctuations of proteins (Ceruso *et al.* 1999; Merlino *et al.* 2003; Merlino *et al.* 2004) as well as thermal denaturation trajectories (Roccatano *et al.* 2003). It has also proven useful in the identification of protein folding transition state ensembles (Marianayagam & Jackson 2004).

In this paper, the native-state dynamics of four proteins with a ubiquitin-like fold are analysed using all-atom molecular dynamic simulations. The structures of the four proteins—ubiquitin (Vijay-Kumar *et al.* 1987), UBX (Buchberger *et al.* 2001), ThiS (Wang *et al.* 2001) and MoaD (Rudolph *et al.* 2001)—are all shown in figure 1. They all adopt the ubiquitin-like β-grasp fold, in which a highly curved mixed β-sheet packs against an α-helix to form the hydrophobic core of the protein. The four proteins come from different organisms—ubiquitin and UBX are mammalian proteins, whereas MoaD and ThiS are bacterial proteins. Despite the fact that they have relatively little sequence homology, it has been proposed that ubiquitin, ThiS and MoaD are evolutionarily related, having evolved from a common ubiquitin-like ancestor (Rudolph *et al.* 2001; Wang *et al.* 2001).

Ubiquitin is involved in tagging proteins for degradation by the proteasome by forming a covalent link through its C-terminus with the target protein. The two C-terminal glycine residues are essential for this function and for the activation of ubiquitin by

cofactor enzymes (Jentsch & Pyrowolakis 2000). ThiS and MoaD also have a conserved C-terminal double glycine motif. ThiS is a sulphur carrier protein that is involved in thiamine biosynthesis in prokaryotes (Wang *et al.* 2001). It is also enzymatically activated through its C-terminal region (Wang *et al.* 2001). MoaD is part of a complex that is involved in molybdenum cofactor biosynthesis in various organisms including humans (Rudolph *et al.* 2001). It binds to the complex through its C-terminal double glycine motif in a similar manner to the way in which ubiquitin tags proteins targeted for degradation (Rudolph *et al.* 2001). In contrast, although the UBX domain of human FAF1 is structurally similar to ubiquitin (Buchberger *et al.* 2001), it lacks the double glycine motif at its C-terminus and is, therefore, thought to have a different function. In addition, whereas ubiquitin, ThiS and MoaD are full-length proteins, UBX is a domain from a larger protein. The UBX domain has now been found in a number of proteins, some of which have been implicated in ubiquitin-mediated degradation pathways. However, the exact function of these proteins and the UBX domain remains unknown (Buchberger *et al.* 2001).

In this study, we use MD simulations in conjunction with an essential dynamics analysis to look for conserved dynamics within the family. The backbone dynamics are observed by monitoring average C-α root mean square deviation (r.m.s.d.) over time, and individual C-α root mean square fluctuations (r.m.s.f.) per residue. Chain compactness is also probed by monitoring the radius of gyration for each of the proteins over time. The motions of the four proteins in phase space are compared by carrying out an essential dynamics analysis on the atomic motion. This motion is then visualized by projection onto the first two principal components determined from essential dynamics. The similarity in the motions between the four proteins can be determined by calculating the root mean square inner product (r.m.s.i.p.) of the first 10 eigenvectors (see §2). The analysis reveals implications for the role of evolution and function in protein dynamics.

2. METHODS

All simulations were carried out using the GROMOS 96 forcefield (van Gunsteren *et al.* 1996) within the GROMACS software package (Berendsen *et al.* 1995). The native-state simulations were run for 2 ns at a constant temperature of 298 K and a constant pressure of 1 atm. The temperature and pressure were regulated by weak coupling to an external bath (Berendsen *et al.* 1984). Electrostatics were dealt with using a 8 Å coloumbic cut-off. The simulations were run from the crystal structures of MoaD (Rudolph *et al.* 2001) and UBQ (Vijay-Kumar *et al.* 1987) and from an average structure of the 20 lowest energy NMR structures of UBX (Buchberger *et al.* 2001) and ThiS (Wang *et al.* 2001). Each structure was fully solvated with SPC water (Berendsen *et al.* 1981) in a cubic box with a 20 Å edge length. The structures were equilibrated in two steps: first, the structures were energy minimized using

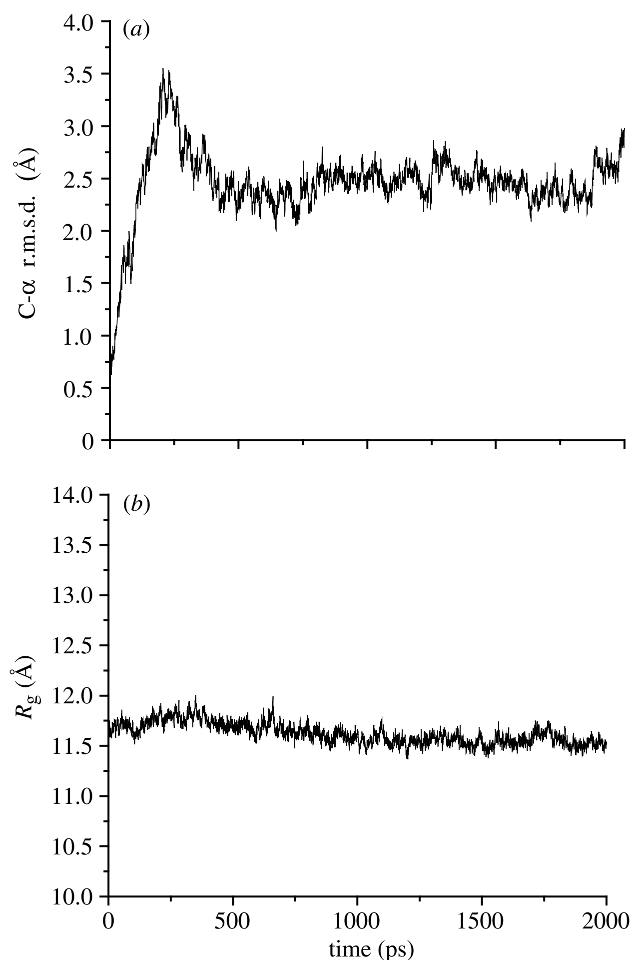

Figure 2. Evolution of structural properties over time for ubiquitin: (*a*) C-α r.m.s.d. and (*b*) radius of gyration, R_g.

steepest descents for 1000 steps and, second, a position-restrained MD run was carried out which holds the protein atoms fixed and allows the solvent to equilibrate around the solute. The atoms in the system were given initial velocities according to a Maxwellian distribution. The system was allowed to evolve according to Newton's equations of motion, with the equations being integrated every 2 fs. The progress of the simulations was monitored by calculating several structural parameters over time: C-α r.m.s.d., radius of gyration (R_g) and C-α r.m.s.f. per residue for the native-state simulations only. Each of these structural properties were calculated within GROMACS.

The essential dynamics method was described in the introduction and it is also described in detail elsewhere (Amadei *et al.* 1993). The essential dynamics analysis was carried out, using programs within GROMACS, on the atomic fluctuations of the four proteins during the course of the simulation. Comparison of the motions spanned by the first 10 eigenvectors from each simulation can be carried out by calculating the r.m.s.i.p. between the first 10 eigenvectors of each trajectory (Amadei *et al.* 1999*b*). The r.m.s.i.p. is defined as

$$\text{r.m.s.i.p.} = \left(\frac{1}{10} \sum_{i=1}^{10} \sum_{j=1}^{10} \nu_{Ai} \cdot \nu_{Bj} \right)^{\frac{1}{2}}, \tag{2.1}$$

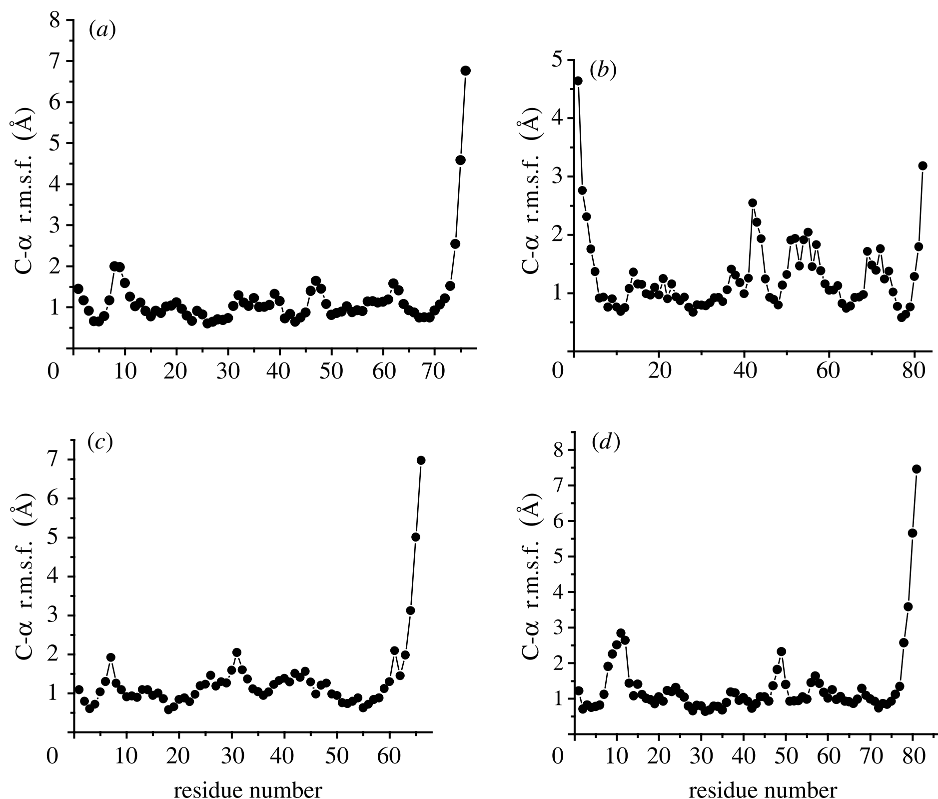

Figure 3. The C-α r.m.s.f. per residue for the four proteins: (*a*) ubiquitin, (*b*) UBX, (*c*) ThiS and (*d*) MoaD.

Table 1. Structural properties for the four ubiquitin-like proteins.
(Columns three, four and five represent average values. Columns six and seven are the net displacements along each eigenvector.)

protein	number of residues	C-α r.m.s.d. (Å)	R_g (Å)	C-α r.m.s.f. (Å)	eig. 1 net (Å)	eig. 2 net (Å)
ubiquitin	76	2.5	11.6	1.2	−16	5
UBX	82	2.9	12.1	1.3	−53	−28
ThiS	66	2.8	11.0	1.3	−23	−7
MoaD	81	2.7	11.6	1.3	−14	1

where ν_{Ai} and ν_{Bj} are the ith and jth eigenvectors from a set of eigenvectors A and B. The r.m.s.i.p. measures the overlap of the motions for each protein in the subspace spanned by the first 10 eigenvectors. The r.m.s.i.p. was calculated between each of the proteins used in this study. The r.m.s.i.p. was also computed for the individual proteins by splitting the trajectories in half and comparing the first 10 eigenvectors from each half of the trajectory.

3. RESULTS AND DISCUSSION

To check the stability of the simulations, the C-α r.m.s.d. and R_g were calculated and monitored for all four proteins over the course of the simulation. The data obtained for ubiquitin was typical of all four proteins and are shown in figure 2. The plot of C-α r.m.s.d. versus time (figure 2*a*) clearly shows that there is an initial spike, after which the system appears to reach an equilibrium value which does not change with time. This profile is similar to that obtained for the other proteins (data not shown). In other native-state simulations of ubiquitin, similar results have been reported (Alonso & Daggett 1998). In our case, the average C-α r.m.s.d. is somewhat higher (2.5 Å) than previous simulations (1.8 Å), which we attribute to the fact that these simulations used a shortened construct of ubiquitin (residues 1–72), which does not contain the flexible C-terminal region included in our studies. The R_g remains constant throughout the simulation for ubiquitin (figure 2*b*) and the three other proteins (data not shown), indicating that the relative compactness of the protein is maintained. Average values for C-α r.m.s.d. and R_g for all four proteins are given in table 1. The R_g values vary between 11 and 12 Å, typical for proteins of this size assuming a spherical conformation (Creighton 1984), and are in good agreement with the R_g reported for other native-state simulations of ubiquitin (Alonso & Daggett 1998). These results demonstrate that no major structural change is taking place during the simulation, confirming that the protein remains within a native-state ensemble.

Figure 3 shows the C-α r.m.s.f. for each residue for the four native-state simulations. The data show that the majority of the amino acids stay within 2–3 Å of their positions in the crystal or NMR structures. As can

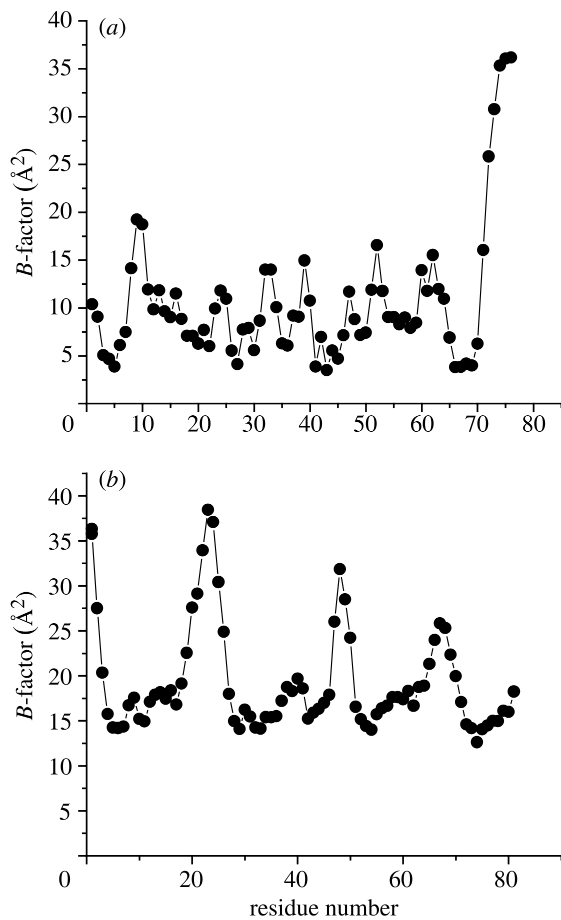

Figure 4. The average *B*-factors per residue taken from the crystal structures of (*a*) ubiquitin and (*b*) MoaD.

be seen for ubiquitin (figure 3*a*), the major fluctuations occur in the last two residues comprising the conserved double glycine motif. This is also the case for ThiS (figure 3*c*) and MoaD (figure 3*d*). The high degree of conformational freedom exhibited by these residues is most probably necessary for forming the protein–protein interactions essential for protein function. This is similar to MD results on the cell-cycle regulatory proteins Ras and Raf, which form an active complex. In this case, flexibility in both the unbound proteins was shown to be necessary for mediating the interaction between them (Gohlke *et al.* 2004).

The inherent flexibility of the C-terminal glycines in ubiquitin can be gleaned from the crystallographic *B*-factors (figure 4*a*), which are high for these two residues. Interestingly though, the C-terminal glycines of MoaD do not have high *B*-factors (figure 4*b*); this maybe a result of the fact that the structure of MoaD was solved as part of a larger complex (Rudolph *et al.* 2001). The inherent flexibility of these residues is only revealed by the native-state simulations reported here, demonstrating the importance of such simulations in interpreting structural data in terms of function.

In addition to the C-terminal residues, another putative protein interaction site has been identified in ThiS (Wang *et al.* 2001). This consists of a solvent exposed hydrophobic patch comprising residues Ile40, Leu58 and Phe60 (Wang *et al.* 2001). A similar patch, consisting of residues Leu8, Ile44 and Val70, is found in ubiquitin and is also thought to be necessary for protein interactions, in this case with the proteasome (Wang *et al.* 2001). An analysis of the dynamics in this area

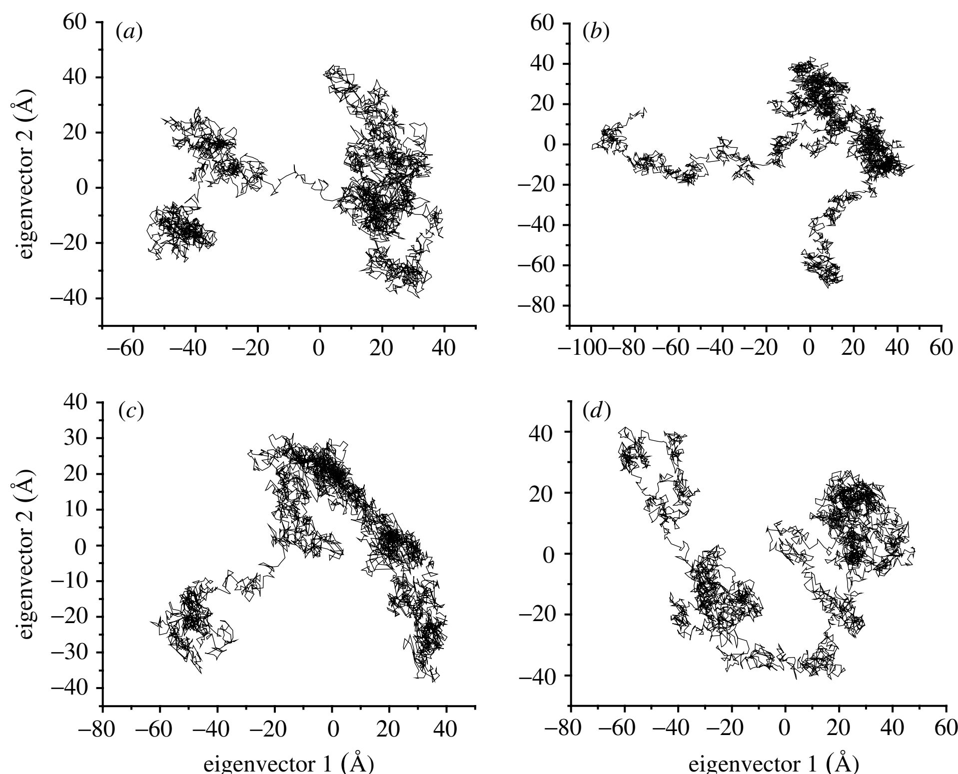

Figure 5. The motion of the four proteins in phase space projected along their first two principal eigenvectors: (*a*) ubiquitin, (*b*) UBX, (*c*) ThiS and (*d*) MoaD.

shows a moderate amount of flexibility. It is interesting to note, however, that they show the same C-α r.m.s.f. of 1.2 Å. Whether this is functionally significant, and therefore demonstrates an optimized mobility for binding surfaces, remains to be shown.

In contrast to the data shown for ubiquitin, ThiS and MoaD, the UBX domain shows different behaviour. Although fluctuations are observed at both the N- and C-termini for the UBX domain (figure 3*b*), these are significantly less than those observed for ubiquitin, ThiS or MoaD. The UBX domain does not contain the double glycine motif nor has it yet been implicated directly in any conjugation events in which the C-terminus of the protein is covalently modified. Thus, the UBX domain is functionally distinct from the other proteins and shows different dynamic behaviour. Thus, there is a correlation between the observed dynamics of these proteins and their function.

So far, we have seen locally correlated motion that is linked to the function of three of the four ubiquitin-like proteins. In order to see whether there is globally correlated motion among the four proteins, a principal-component or essential-dynamics analysis was undertaken on the atomic motion of the proteins during the course of the simulation. The motion of the four proteins projected along their principal components or eigenvectors is shown in figure 5. To see if there is any correlated motion between the proteins, the r.m.s.i.p. of the first 10 eigenvectors from each of the simulations was computed. These values are shown in table 2. The r.m.s.i.p. measures the degree of overlap or similarity of eigenvector sets (see §2). A r.m.s.i.p. of 1 indicates the sets are identical, while a value of 0 indicates that the eigenvectors are orthogonal (Amadei *et al.* 1999*b*). As expected, the highest degree of overlap occurs within the individual trajectories for each protein. The average overlap in the dynamics of proteins in the ubiquitin family is 0.48. The greatest overlap is between functionally related proteins with ubiquitin and ThiS having an overlap of 0.55, and MoaD and ThiS having an overlap of 0.54. It is interesting to note that MoaD and ThiS have the highest sequence identity out of the four proteins studied (29.5%). This is reflected in the fact that ThiS and MoaD are thought to have diverged from the common ancestral ubiquitin protein at roughly the same time (Wang *et al.* 2001). These results suggest that the dynamic properties of the proteins have been conserved along with sequence. Our results are similar to those of other studies which have also used essential-dynamics analysis to compare the molecular motions of two members of the pancreatic-like superfamily (Merlino *et al.* 2003). In this case, the native-state dynamics of RNase A was compared with human angiogenin, with which it shares 33% sequence identity. Similar motions were observed and a r.m.s.i.p. of 0.6 reported (Merlino *et al.* 2003). The proteins of the ubiquitin family studied here show slightly less correlated motion despite structural and functional similarity; this is probably owing to their lower sequence homology.

Table 2. The calculated r.m.s.i.p. between each of the proteins.
(The values on the diagonal are calculated by splitting the individual trajectories in half and calculating the overlap between each half.)

	ubiquitin	UBX	ThiS	MoaD
ubiquitin	0.71	0.43	0.55	0.49
UBX		0.73	0.46	0.42
ThiS			0.70	0.54
MoaD				0.74

UBX has almost no sequence homology to ubiquitin, while the sequence identity of ThiS is very low (14%). It is interesting to note that UBX, which is not functionally related to the other three proteins, shows the least overlap in terms of its molecular motions with the other three (table 2). These results support the idea that there is a strong link between protein dynamics and protein function.

The native-state simulations represent how a macromolecule behaves at its free-energy minimum. Owing to the structural complexity of proteins, there are many conformations or sub-minima that make up this global minimum. The essential dynamics analysis shown in figure 5 shows the amplitudes of the motions of each protein at its energy minimum. The net displacement of each of the proteins along their first two eigenvectors is shown in table 1. Ubiquitin, ThiS and MoaD all have similar displacements along their two principal eigenvectors, whereas the displacement for UBX is quite different. In figure 6, backbone representations of the four proteins are shown; the shaded regions are those sections of the protein that contribute most to motion along the first eigenvector. The regions shaded in red and magenta have the highest contribution, while those shaded in blue contribute to a lesser degree. It is interesting to note that the C-terminal region of ubiquitin, ThiS and MoaD contribute the most to motion along the first eigenvector, whereas in UBX, the N-terminus shows the largest contribution. The essential-dynamics analysis reinforces the idea that the UBX domain demonstrates quite different dynamic behaviour to the other three proteins. In other cases, it has been shown that proteins with similar native structures display similar dynamics (Keskin *et al.* 2000). Here, however, we have clearly demonstrated for four ubiquitin-like proteins that global dynamics are influenced by function as well as by structure.

In this study, we have investigated the dynamics of four proteins in phase space. It is important to take into account that these simulations are run for 2 ns. Ideally, a more complete picture of the dynamics can be obtained if the simulations are run for longer. However, figure 6 helps to rectify the problem by showing any correlated motions between the proteins. Another consideration is how the use of a coloumbic cut-off might affect the results of the essential dynamics analysis. However, it has been shown that using either a cut-off or a particle mesh Ewald summation to deal with electrostatics produces similar

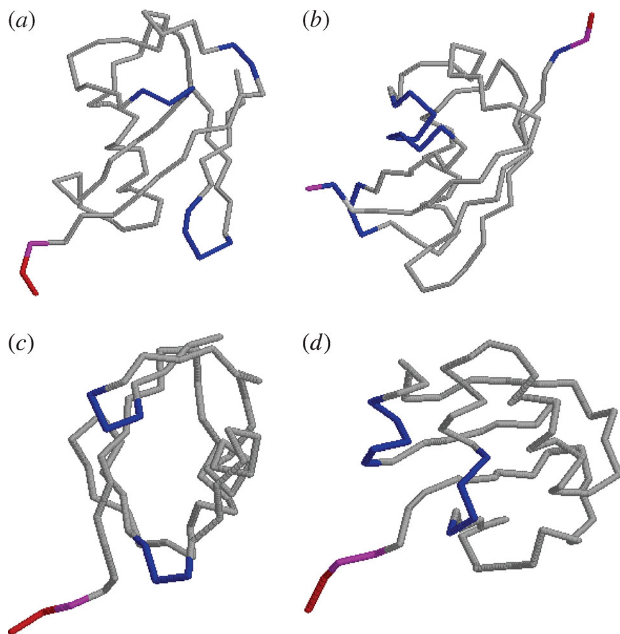

Figure 6. Backbone representations of the four proteins: (*a*) ubiquitin, (*b*) UBX, (*c*) ThiS and (*d*) MoaD. The shaded regions of the structure correspond to regions, which contribute most to motion along the first eigenvector. Red and magenta indicate the largest contributions, blue indicates a smaller contribution, and the unshaded regions indicate no significant contribution to motion along the first eigenvector.

results in the essential dynamics analysis (Merlino *et al.* 2003).

4. CONCLUSIONS

The data presented here show a clear correlation in the dynamics of three proteins in the ubiquitin family. There is a high degree of correlated motion in both the local and global dynamics of the functionally related proteins: ubiquitin, ThiS and MoaD. In contrast, a lower degree of correlation is observed with the UBX domain, a structural but not a functional homologue of ubiquitin. Together, our results suggest that protein dynamics have been optimized during evolution for protein function and that structural similarity between proteins does not necessarily imply dynamic similarity.

The authors would like to thank the Welton Foundation for funding during the course of this work. We would also like to acknowledge The Unilever Centre for Molecular Informatics, Cambridge, for computational time. We are grateful to Dr. Charlotte Bolton and Dr. David Wales for help in setting up GROMACS. N.J.M. would like to thank Andrew Brown and Abhirami Ratnakumar for useful discussions on protein dynamics.

REFERENCES

Alonso, D. O. & Daggett, V. 1998 Molecular dynamics simulations of hydrophobic collapse of ubiquitin. *Protein Sci.* **7**, 860–874.

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

Amadei, A., Ceruso, M. A. & Di Nola, A. 1999*a* On the convergence of the conformational coordinates basis set obtained by the essential dynamics analysis of proteins' molecular dynamics simulations. *Proteins* **36**, 419–424.

Amadei, A., de Groot, B. L., Ceruso, M. A., Paci, M., Di Nola, A. & Berendsen, H. J. 1999*b* A kinetic model for the internal motions of proteins: diffusion between multiple harmonic wells. *Proteins* **35**, 283–292.

Berendsen, H. J. C., Postman, J. D. M., van Gunsteren, W. F. & Hermans, J. 1981 *Intermolecular forces. Interaction models for water in relation to protein hydration*, pp. 331–342. Dordrecht: Reidel.

Berendsen, H. J. C., Postma, J. P. M., DiNola, A. & Haak, J. R. 1984 Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.

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

Brooks, B. R., Dusanka, J. & Karplus, M. 1995 Harmonic analysis of large systems I. Methodology. *J. Comp. Chem.* **16**, 1522–1542.

Buchberger, A., Howard, M. J., Proctor, M. & Bycroft, M. 2001 The UBX domain: a widespread ubiquitin-like module. *J. Mol. Biol.* **307**, 17–24.

Ceruso, M. A., Amadei, A. & Di Nola, A. 1999 Mechanics and dynamics of B1 domain of protein G: role of packing and surface hydrophobic residues. *Protein Sci.* **8**, 147–160.

Creighton, T. E. 1984 *Proteins: structure and molecular properties.* New York: Freeman.

Cui, Q., Li, G., Ma, J. & Karplus, M. 2004 A normal mode analysis of structural plasticity in the biomolecular motor F(1)-ATPase. *J. Mol. Biol.* **340**, 345–372.

Gohlke, H., Kuhn, L. A. & Case, D. A. 2004 Change in protein flexibility upon complex formation: analysis of Ras–Raf using molecular dynamics and a molecular framework approach. *Proteins* **56**, 322–337.

Jentsch, S. & Pyrowolakis, G. 2000 Ubiquitin and its kin: how close are the family ties?. *Trends Cell Biol.* **10**, 335–342.

Kalodimos, C. G., Biris, N., Bonvin, A. M., Levandoski, M. M., Guennuegues, M., Boelens, R. & Kaptein, R. 2004 Structure and flexibility adaptation in nonspecific and specific protein–DNA complexes. *Science* **305**, 386–389.

Keskin, O., Jernigan, R. L. & Bahar, I. 2000 Proteins with similar architecture exhibit similar large-scale dynamic behavior. *Biophys. J.* **78**, 2093–2106.

Marianayagam, N. J. & Jackson, S. E. 2004 The folding pathway of ubiquitin from all-atom molecular dynamics simulations. *Biophys. Chem.* **111**, 159–171.

Merlino, A., Vitagliano, L., Ceruso, M. A. & Mazzarella, L. 2003 Subtle functional collective motions in pancreatic-like ribonucleases: from ribonuclease A to angiogenin. *Proteins* **53**, 101–110.

Merlino, A., Vitagliano, L., Ceruso, M. A. & Mazzarella, L. 2004 Dynamic properties of the N-terminal swapped dimer of ribonuclease A. *Biophys. J.* **86**, 2383–2391.

Rasmussen, B. F., Stock, A. M., Ringe, D. & Petsko, G. A. 1992 Crystalline ribonuclease A loses function below the dynamical transition at 220 K. *Nature* **357**, 423–424.

Roccatano, D., Daidone, I., Ceruso, M. A., Bossa, C. & Nola, A. D. 2003 Selective excitation of native fluctuations during thermal unfolding simulations: horse heart cytochrome c as a case study. *Biophys. J.* **84**, 1876–1883.

Rudolph, M. J., Wuebbens, M. M., Rajagopalan, K. V. & Schindelin, H. 2001 Crystal structure of molybdopterin synthase and its evolutionary relationship to ubiquitin activation. *Nat. Struct. Biol.* **8**, 42–46.

van Gunstren, W. F., Billeter, S. R., Eising, A. A., Humberger, P. H., Kruger, P., Mark, A. E., Scott,

W. R. P. & Tironi, I. G. 1996 *Biomolecular simulation: the GROMOS96 manual and user guide*. Zurich: Hochschulverlag AG an der ETH.

Vijay-Kumar, S., Bugg, C. E. & Cook, W. J. 1987 Structure of ubiquitin refined at 1.8 Å resolution. *J. Mol. Biol.* **194**, 531–544.

Vitagliano, L., Merlino, A., Zagari, A. & Mazzarella, L. 2002 Reversible substrate-induced domain motions in ribonuclease A. *Proteins* **46**, 97–104.

Wang, C., Xi, J., Begley, T. P. & Nicholson, L. K. 2001 Solution structure of ThiS and implications for the evolutionary roots of ubiquitin. *Nat. Struct. Biol.* **8**, 47–51.