Downloaded from rsif.royalsocietypublishing.org

JOURNAL OF THE ROYAL SOCIETY Interface

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

Neelan J Marianayagam and Sophie E Jackson

J. R. Soc. Interface 2005 **2**, 47-54
doi: 10.1098/rsif.2004.0025

References This article cites 23 articles, 1 of which can be accessed free
http://rsif.royalsocietypublishing.org/content/2/2/47.full.html#ref-list-1

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *J. R. Soc. Interface* go to: **http://rsif.royalsocietypublishing.org/subscriptions**

This journal is © 2005 The Royal Society

Downloaded from rsif.royalsocietypublishing.org

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

Downloaded from rsif.royalsocietypublishing.org

JOURNAL OF THE ROYAL SOCIETY Interface

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