BIOINFORMATICS APPLICATIONS NOTE Vol. 23 no. 7 2007, pages 901–902
doi:10.1093/bioinformatics/btm014

Structural bioinformatics

DFprot: a webtool for predicting local chain deformability

José Ignacio Garzón[1], Julio Kovacs[2], Ruben Abagyan[2] and Pablo Chacón[1,*]

[1]Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain and
[2]Department of Molecular Biology, The Scripps Research Institute La Jolla, California 92037, USA

Received on November 6, 2006; revised on January 11, 2007; accepted on January 14, 2007

Advance Access publication February 3, 2007

Associate Editor: Alfonso Valencia

ABSTRACT

Summary: DFprot is a web-based server for predicting main-chain deformability from a single protein conformation. The server automatically performs a normal-mode analysis (NMA) of the uploaded structure and calculates its capability to deform at each of its residues. Non-specialists can easily and rapidly obtain a quantitative first approximation of the flexibility of their structures with a simple and efficient interface.

Availability: http://sbg.cib.csic.es/Software/DFprot

Contact: pablo@cib.csic.es

1 INTRODUCTION

Protein structural flexibility is closely coupled to function, as evidenced by many conformational changes observed on key cellular process. Thus, inferring the intrinsic molecular flexibility from a single conformation could offer a direct link to understand likely large-scale rearrangements. Standard protein dynamics simulations are too time consuming for studying the large timescale changes often associated to macromolecular function. There are, however, faster approximations based on Brownian dynamics with protein reduced representations that achieve good results (Sacquin-Mora and Lavery, 2006). In this context, the use of coarse-grained models and normal-mode analysis (NMA) is an interesting alternative, especially for relative large systems (Bahar and Rader, 2005; Ma, 2005). Tirion (Tirion, 1996) was the pioneer in the combination of NMA with a simplified protein representation (the so-called 'elastic network model') to study protein movements, but the idea was further refined, validated and extended by several research groups, including Bahar (Bahar *et al*., 1997), Hinsen (Hinsen, 1998), Sanejouand (Tama and Sanejouand, 2001), Jernigan (Song and Jernigan, 2006) and Brooks (Zheng *et al*., 2006). These groups and some others gathered a wealth of evidence that support the NMA coarse-grain approximation as a successful tool to simulate deformational motions of macromolecular complexes at extended length scales, even from low-resolution structures (Chacon *et al*., 2003; Kong *et al*., 2003). Web-based database systems such as MolmovDB (Alexandrov *et al*., 2005), ProMode (Wako *et al*., 2004) or iGNM (Yang *et al*., 2005) give access to numerous examples of the good correlation between low-frequency normal modes and the collective, large-amplitude motions observed experimentally. These tools are complemented with other web servers such as Elnémo (Suhre and Sanejouand, 2004), Webnm@ (Hollup *et al*., 2005), AD-ENM (Zheng and Doniach, 2003), NOMAD-ref (Lindahl *et al*., 2006) or MoViES (Cao *et al*., 2004) which also provide online normal-mode calculation with a variety of functionalities for the analysis of the results.

The location of the chain 'hinges' or high deformability areas can be derived from the normal modes. Here we report a web server to compute, analyze and visualize main-chain deformability, which is a measure of the capability of a given molecule to deform at each of its residues. This new measure is deduced by treating the normal modes as vector fields over the molecule, and applying the conformal field theory (Kovacs *et al*., 2004). Deformability predictions correlate well with the observed experimental flexibility, as it has been demonstrated previously by comparing such predictions with the observed flexibility of a set of kinases solved in at least two different conformations (Kovacs *et al*., 2004). In contrast with a recently proposed methodology based on graph theory (Jacobs *et al*., 2001), which gives a qualitative distinction between rigid and non-rigid residues, this method gives a quantitative first-order measure of flexibility. Moreover, these flexibility predictions have been used to define a 'relevance measure' that has been successfully applied in the generation of multiple receptor conformations from low-frequency NMA information, thereby providing an efficient approach to include intrinsic receptor flexibility in ligand docking and virtual screening (Cavasotto *et al*., 2005).

The purpose of the web server presented here, called DFprot, is to make deformability calculations accessible to all structural biologists and other researchers worldwide. This server combines a simple input interface with an efficient implementation that enhances interactivity with a suitable 2D/3D display of the results.

2 IMPLEMENTATION

Starting from a structure file provided by the user in PDB format, the server calculates the local chain deformability. This is done by automatically performing the normal-mode analysis on the C_α atoms of the uploaded structure, setting the spring strengths to $C_{ij}=(r_0/r_{ij})^6$, where i,j denote residue numbers, r_{ij} is the distance between the C_α atoms of residues i and j and r_0 was set to 3.8 Å, which is approximately the mean distance between consecutive C_α atoms. For simplicity, we have removed the normalized residue contact area term from the previous formulation. Using the residue masses and the spring strengths given above, the Hessian matrix is diagonalized, yielding $3N-6$ eigenvalues λ_n and their corresponding eigenvectors $\mathbf{u}^n$ (the normal modes of vibration).

*To whom correspondence should be addressed.

© The Author 2007. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org

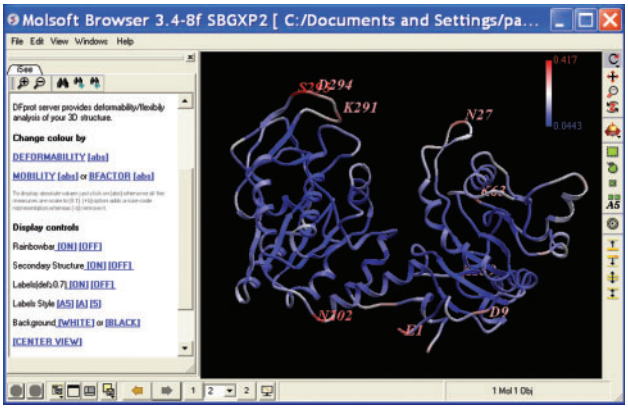

Fig. 1. Snapshot of deformability results using iSee methodology.

Considering each mode as a vector field over the molecule, the deformability is calculated using the following formula (Cavasotto *et al.*, 2005; Kovacs *et al.*, 2004):

$$d_M^2(i) = \sum_{n=7}^{3N} \left(\left\| S_{\mathbf{u}^n}(i) \right\| / \lambda_n \right)^2 \quad S_{\mathbf{u}^n} = \frac{1}{2}\left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k} \right) - \frac{1}{3}\delta_{kl} \operatorname{div} \mathbf{u}^n$$

where $S_{\mathbf{u}^n}$ describes how the vector field $\mathbf{u}^n$ affects (locally) the shape of the molecule. Also, k, l are indices of spatial coordinates ($1 \leq k, l \leq 3$), δ denotes the Kronecker delta function and div the *divergence*. The details for the numerical computation of deformability are described elsewhere (Kovacs *et al.*, 2004).

The server presents the deformability results in a simple 2D plot. Here the user can readily identify the relative flexibility/rigidity of residues and regions of the protein. The visualization is extended to 3D by either using the simple Jmol viewer (www.jmol.org) or the rich ICM-Browser (free at www.molsoft.com). The generated data files are fully interactive and use the iSee technology (Abagyan *et al.*, 2006).

The user can view the query structures colored by the predicted deformability. Since hinge points are highly flexible regions, a simple way to detect them is by locating the most deformable regions (colored in red) and checking its structural reliability with the help of an interactive visualization (Fig. 1). In addition, the user can analyze the mobility calculated from the vibrations observed by NMA and experimental B-factors. Note that mobility and deformability are complementary and different measures. The relation between them is akin to a relation between function and its derivative. The deformability is a measure of likelihood to form a hinge in protein chain, whereas the protein mobility reflects amplitude of the atomic fluctuations.

Finally, for users who prefer to analyze the results with their favorite software, raw results can be downloaded. This includes the deformability/mobility-tabulated data, the eigenvalues and eigenvectors of the NMA and the displayed structures in PDB format. Like other NMA webtools, DFprot offers online animations of the normal modes. To better display the mode-perturbed structures, the user also can vary, interactively, the vibrational amplitude.

3 CONCLUSIONS

The web-based server described here offers an effectual way of computing the first-order flexibility measure of user supplied protein structures. It provides a quantitative measure of local chain deformability, along with a simple and convenient framework to analyze and display such measure on its 3D structural context. We believe that this web tool will provide a straightforward link between protein structures and their potential dynamics.

ACKNOWLEDGEMENTS

This work was supported by funds from BFU2004-01282 and Fundación BBVA (P.C.) and NIH 1-R01-GM071872-01 (R.A.).

Conflict of Interest: none declared.

REFERENCES

Abagyan,R. *et al.* (2006) Disseminating structural genomics data to the public: from a data dump to an animated story. *Trends Biochem. Sci.*, **31**, 76–78.

Alexandrov,V. *et al.* (2005) Normal modes for predicting protein motions: a comprehensive database assessment and associated Web tool. *Protein Sci.*, **14**, 633–643.

Bahar,I. *et al.* (1997) Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold. Des.*, **2**, 173–181.

Bahar,I. and Rader,A.J. (2005) Coarse-grained normal mode analysis in structural biology. *Curr. Opin. Struct. Biol.*, **15**, 586–592.

Cao,Z.W. *et al.* (2004) MoViES: molecular vibrations evaluation server for analysis of fluctuational dynamics of proteins and nucleic acids. *Nucleic Acids Res.*, **32**, W679–W685.

Cavasotto,C.N. *et al.* (2005) Representing receptor flexibility in ligand docking through relevant normal modes. *J. Am. Chem. Soc.*, **127**, 9632–9640.

Chacon,P. *et al.* (2003) Mega-Dalton biomolecular motion captured from electron microscopy reconstructions. *J. Mol. Biol.*, **326**, 485–492.

Hinsen,K. (1998) Analysis of domain motions by approximate normal mode calculations. *Proteins*, **33**, 417–429.

Hollup,S.M. *et al.* (2005) WEBnm@: a web application for normal mode analyses of proteins. *BMC Bioinformatics*, **6**, 52.

Jacobs,D.J. *et al.* (2001) Protein flexibility predictions using graph theory. *Proteins*, **44**, 150–165.

Kong,Y. *et al.* (2003) Conformational flexibility of pyruvate dehydrogenase complexes: a computational analysis by quantized elastic deformational model. *J. Mol. Biol.*, **330**, 129–135.

Kovacs,J.A. *et al.* (2004) Predictions of protein flexibility: first-order measures. *Proteins*, **56**, 661–668.

Lindahl,E. *et al.* (2006) NOMAD-Ref: visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis. *Nucleic Acids Res.*, **34**, W52–W56.

Ma,J. (2005) Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. *Structure*, **13**, 373–380.

Sacquin-Mora,S. and Lavery,R. (2006) Investigating the local flexibility of functional residues in hemoproteins. *Biophys. J.*, **90**, 2706–2717.

Song,G. and Jernigan,R.L. (2006) An enhanced elastic network model to represent the motions of domain-swapped proteins. *Proteins*, **63**, 197–209.

Suhre,K. and Sanejouand,Y.H. (2004) ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. *Nucleic Acids Res.*, **32**, W610–W614.

Tama,F. and Sanejouand,Y.H. (2001) Conformational change of proteins arising from normal mode calculations. *Protein Eng.*, **14**, 1–6.

Tirion,M.M. (1996) Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.*, **77**, 1905–1908.

Wako,H. *et al.* (2004) ProMode: a database of normal mode analyses on protein molecules with a full-atom model. *Bioinformatics*, **20**, 2035–2043.

Yang,L.W. *et al.* (2005) iGNM: a database of protein functional motions based on Gaussian Network Model. *Bioinformatics*, **21**, 2978–2987.

Zheng,W. *et al.* (2006) Low-frequency normal modes that describe allosteric transitions in biological nanomachines are robust to sequence variations. *Proc. Natl Acad. Sci. USA*, **103**, 7664–7669.

Zheng,W. and Doniach,S. (2003) A comparative study of motor-protein motions by using a simple elastic-network model. *Proc. Natl Acad. Sci. USA*, **100**, 13253–13258.