Reflection of Knowledge Information in *Pro*Mode
(A Database of Normal Mode Analyses on Proteins)

Hiroshi Wako,[1]   **Motohiro Otsuka,**[2]   **Yuhki Tomizawa,**[3]
wako@waseda.jp   ootsuka@mxs.nes.nec.co.jp   tomizawa@sciencehouse.co.jp

Masaki Kato,[4]   **Shigeru Endo**[5]
kato@tsurumi.yokohama-cu.ac.jp   endo@sci.kitasato-u.ac.jp

[1]   School of Social Sciences, Waseda University, 1-6-1, Nishi-Waseda, Shinjuku-ku, Tokyo, 169-8050, Japan
[2]   NEC Soft, Ltd., 1-18-6, Shinkiba, Koto-ku, Tokyo 136-8608, Japan
[3]   Sciencehouse, Inc., 1-4-17, Higashi-Tabata, Kita-ku, Tokyo 114-0013, Japan
[4]   Department of Science of Biological Supramolecular Systems, Graduate School of Integrated Science, Yokohama City University, 1-7-29, Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
[5]   Department of Physics, School of Science, Kitasato University, 1-15-1, Kitasato, Sagamihara 228-8555, Japan

1   Introduction

*Pro*Mode (http://promode.socs.waseda.ac.jp) is a database collecting results from normal mode analyses (NMA) of various protein molecules [5]. *Pro*Mode is designed so as to play an animation of a vibrating protein molecule with a free plug-in soft, Chime (MDL Information Systems, Inc.), and to graph out the results from NMA, such as root-mean-square fluctuations of atoms and dihedral angles. A user can learn the dynamic structure of the protein through the animation and the graphs.

In this paper we will report the recent improvements in *Pro*Mode.

2   Use of the SCOP classification in a protein search

The collection of the NMA results for various proteins in *Pro*Mode may be useful for comparative and statistical studies of dynamic structures of proteins. Although *Pro*Mode has collected the NMA data of more than 1,500 protein chains, it is still hard to carry out NMA for the entire PDB data, because of computation time required. In such a situation, for proteins not contained in *Pro*Mode, it is useful to provide some information about the relative proteins in *Pro*Mode. For this purpose we utilize the SCOP classification of protein conformations [4]. If a user searches *Pro*Mode for a protein with some PDB ID and no data is found there, the search tries to find an alternative protein in the same superfamily or family in *Pro*Mode. The user may obtain some information about the protein of interest through the alternative one.

Additionally, a user can also search the *Pro*Mode database along the SCOP tree structure like the original SCOP web page. At any level (class, fold, superfamily and family), the user is informed whether or not a given category, e.g. superfamily, has an entry data. With this information we can consider which proteins should be a target in the next NMA calculation. It may be also useful for a user to compare the NMA results of the proteins in a given superfamily or family.

3   Reflection of knowledge information in an animation

It has been shown by many studies that the results from NMA can provide a proper description of the functionally important motions of the proteins. Therefore, in *Pro*Mode, an animation can be manipulated by a control panel we developed to highlight residues in which a user is interested. As a result, we can observe

the relationship between the vibrational motion and the functionally important residues in the animation. However, information about, for example, which residues are functionally important, had to be provided by the user. It is inconvenient, especially for a user not familiar to the protein displayed. In the new version of *Pro*Mode described here, some knowledge information is prepared and can be reflected easily in the animation just by clicking on a button. The knowledge information from Catalytic Residue Dataset [1], PROSITE [2], and eF-site [3] is available at present. An example is shown in Figure 1. A user may be able to relate functionally important residues to the dynamic structure of the protein through the observation of the animation.
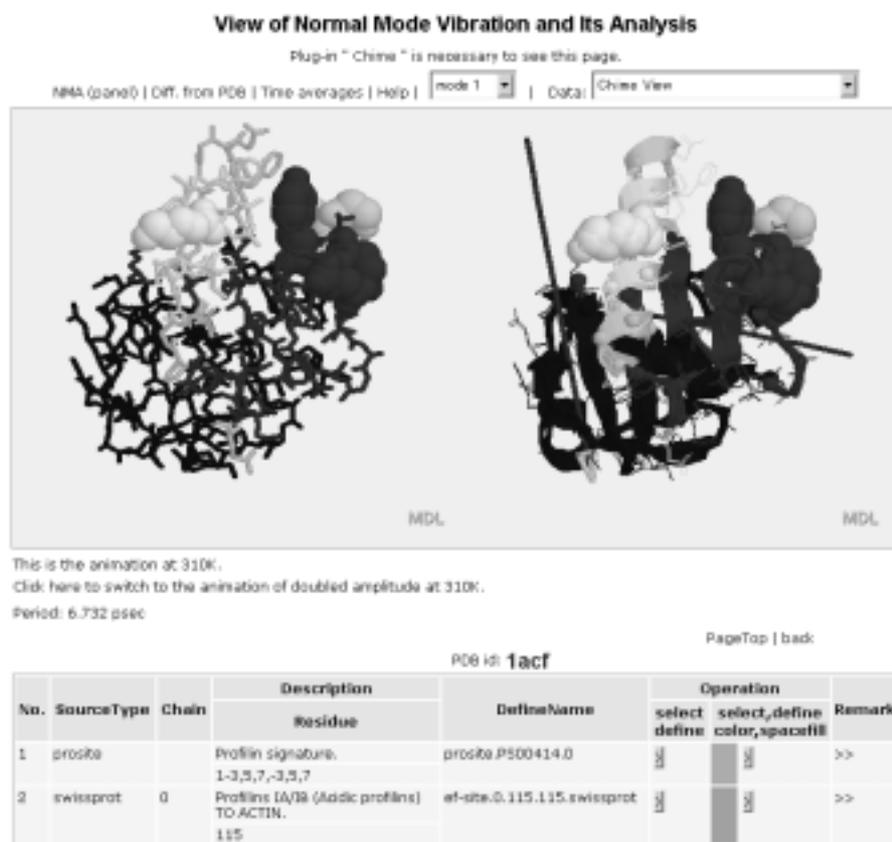

Figure 1. Example for the reflection of knowledge information in the Chime views (1acf: profilin I). The left Chime view is an animation and the right one is a static image displayed in cartoon. The residues are colored depending on dynamic domains they belong to. The arrows in the right view indicate a screw axis related to the screw motion between the dynamic domains. The knowledge information is given in the table below. The specified residues in the table (the PROSITE motif and the binding site to actin in this case) are displayed in a space-filling model in both Chime views.

No.	SourceType	Chain	Description / Residue	DefineName	Operation select define	select,define color,spacefill	Remark
1	prosite		Profilin signature. / 1-3,5,7,-3,5,7	prosite.PS00414.0	☑	☑	>>
2	swissprot	0	Profilins IA/IB (Acidic profilins) TO ACTIN. / 115	ef-site.0.115.115.swissprot	☑	☑	>>

References

[1] Bartlett, G. J., Porter, C. T., Borkakoti, N. and Thornton, J.M., Analysis of catalytic residues in enzyme active sites, *J. Mol. Biol.,* **324**:105-121, 2002.

[2] Hulo, N., Sigrist, C. J. et al., Recent improvements to the PROSITE database, *Nucleic Acids Res.,* **32 Database issue**:D134-137, 2004.

[3] Kinoshita, K. and Nakamura, H., eF-site and PDBjViewer: database and viewer for protein functional sites, *Bioinformatics,* **20**:1329-1330, 2004.

[4] Murzin A. G., Brenner S. E., Hubbard T. and Chothia C., SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.,* **247**:536-540, 1995.

[5] Wako, H., Kato, M. and Endo, S., ProMode: a database of normal mode analyses on protein molecules with a full-atom model, *Bioinformatics,* **20**:2035-2043, 2004.