

**蛋白質立体構造に対するパターン認識：**

**全体の形状、局所構造、表面構造の分類と類似性検索**

**Daron M. Standley<sup>1)</sup>, 金城 玲<sup>2)</sup>, 木下 賢吾<sup>3)</sup>, 中村 春木<sup>2)</sup>**

<sup>1)</sup>阪大免疫学フロンティア, <sup>2)</sup>PDBj, 阪大蛋白研, <sup>3)</sup>東大医科研

(harukin@protein.osaka-u.ac.jp, <http://www.pdbj.org>)

日本蛋白質構造データバンク (PDBj) では、蛋白質や核酸などの生体高分子の三次元立体構造を、国際蛋白質構造データバンク (wwPDB) のメンバーの一員として、品質管理を行いながら登録・編集業務を行っている[1,2]。その一方で、生体高分子の立体構造データ[3]という、テキストやDNA配列のような discrete でない情報 (アナログ的な情報) に対して、その分類と検索を行う仕組みを開発し公開している[4,5]。講演では、特に、(i) 蛋白質の全体の構造 (フォールド) の分類と類似性検索[6,7]、(ii) 蛋白質の局所的な機能部位の分類と類似性検索[8]、(iii)蛋白質の分子表面の類似性検索[9-11]、について、その手法の原理とサービスおよび研究への応用を紹介する[12]。

**Pattern recognition for Protein three-dimensional structures:**

**Similarity search of protein folds, local atomic arrangements, and molecular surfaces**

**Daron M. Standley<sup>1)</sup>, Akira R. Kinjo<sup>2)</sup>, Kengo Kinoshita<sup>3)</sup>, Haruki Nakamura<sup>2)</sup>**

<sup>1)</sup>Immunology Frontier Research Center, Osaka University, <sup>2)</sup>PDBj, Institute for Protein Research, Osaka University, <sup>3)</sup>Institute of Medical Science, The University of Tokyo

(harukin@protein.osaka-u.ac.jp, <http://www.pdbj.org>)

The Protein Data Bank Japan (PDBj) curates, edits and distributes protein three-dimensional (3D) structural data as a member of the worldwide Protein Data Bank (wwPDB) and currently processes about 25-30% of all deposited data in the world [1, 2]. In addition to prepare the description formats and the viewers of those structural data [3], several tools and services have been developed for functional annotations from those protein 3D structures [4,5]. In particular, we have developed the algorithms of (i) rapid and correct similarity search of protein folds [6,7], (ii) very rapid comprehensive atomic structural alignment of functional sites [8], and (iii) molecular surface comparison with physicochemical properties [9-11]. Several examples of the applications to annotate the protein unknown functions will be shown [12].

## 【References】

- [1] H. Berman, K. Henrick, H. Nakamura, Announcing the worldwide Protein Data Bank. *Nature Struct. Biol.*, **10**, 980-980 (2003).
- [2] H. Berman, K. Henrick, H. Nakamura, J. L. Markley, The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucl. Acids Res.*, **35**, D301-D303 (2007).
- [3] J. Westbrook, N. Ito, H. Nakamura, K. Henrick, H. M. Berman, PDBML: The representation of archival macromolecular structure data in XML. *Bioinformatics*, **21**, 988-992 (2005).
- [4] K. Kinoshita, H. Nakamura, Protein informatics towards function identification. *Curr. Opin. Struct. Biol.*, **13**, 396-400 (2003).
- [5] D. M. Standley, A. R. Kinjo, K. Kinoshita, H. Nakamura, Protein structure databases with new Web services for structural biology and biomedical research. *Briefings in Bioinformatics*, **9**, 276-285 (2008).
- [6] D. M. Standley, H. Toh, H. Nakamura, Detecting local structural similarity in proteins by maximizing number of equivalent residues. *PROTEINS*, **57**, 381-391 (2004).
- [7] D. M. Standley, H. Toh, H. Nakamura, GASH: An improved algorithm for maximizing the number of equivalent residues between two protein structures. *BMC Bioinformatics*, **6**, 221 (2005).
- [8] A. R. Kinjo, H. Nakamura, Similarity search for local protein structures at atomic resolution by exploiting a database management system. *Biophysics*, **3**, 75-84 (2007).
- [9] K. Kinoshita, H. Nakamura, Identification of protein biochemical functions by similarity search using the molecular surface database, eF-site. *Protein Science*, **12**, 1589-1595 (2003).
- [10] K. Kinoshita, H. Nakamura, eF-site and PDBjViewer: database and viewer for protein functional sites. *Bioinformatics*, **20**, 1329-1330 (2004).
- [11] K. Kinoshita, Y. Murakami, H. Nakamura, eF-seek: prediction of the functional sites of proteins by searching for similar electrostatic potential and molecular surface shape. *Nucl. Acids Res.*, **35**, W398-W402 (2007).
- [12] D. M. Standley, H. Toh, H. Nakamura, Functional annotation by sequence-weighted structure alignments: Statistical analysis and case studies from the Protein 3000 structural genomics project in Japan. *PROTEINS*, **72**, 1333-1351 (2008).