Beyond Text-based Queries at the Protein Data Bank, Japan

Daron M. Standley^{a,b}, Haruki Nakamura^a, ^aProtein Research Institute, Osaka University, Osaka, Japan. ^bJapan Science and Technology Agency. E-mail: standley@protein.osaka-u.ac.jp

The Protein Data Bank has traditionally offered only text-based query services. These tools are very powerful in the hands of experts when the data entries have been well annotated. As the database grows through structural genomics projects, however, annotation will likely become limited. Here we introduce a suite of query tools that do not require complex textual input. Starting from a particular entry one may find sequence homologs (Sequence Navigator[1]), structural neighbors (Structure Navigator[2]), or, if the entry is a protein complex, structurally similar protein-protein interfaces (PISup[3]). In addition, alignments may be further optimized and refined using our powerful structural alignment program GASH[4]. All of the above programs utilize the Number of Equivalent Residues (NER[5]), a novel scoring function that detects similarities rather than differences between structures. In this way, even local similarities (i.e., domains, active sites, etc.) can be detected.

[1] http://www.pdbj.org/cgi-bin/run_seq_hom.cgi [2] http:// www.pdbj.org/cgibin/run_algn_struc.cgi [3] http://www.pdbj.org/cgi-bin/run_pisup.cgi [4] http://www.pdbj.org/cgi-bin/run_gash.cgi [5] Standley D.M., Toh H., Nakamura H., *Proteins*, 2004, **57**(2), 381-391.

Keywords: protein structure comparison, databases, proteinprotein interactions