comments

Response to Paoli

Paoli's analysis of archaeal and eukaryotic translation initiation factor 6 (aIF6 and eIF6, PDB codes 1G61, 1G62)[1] and a pair of functionally distinct enzymes, L-arginine:glycine amidinotransferase (PDB code 1JDW)[2] and L-arginine:inosamine-phosphate amidinotransferase (PDB code 1BWD)[3], underscores the difficulties inherent in trying to establish a conceptual framework for protein fold classification. To paraphrase Holm and Sander[4]: Are the IF6s amidinotransferase α/β-propellers or are they like the amidinotransferase α/β-propellers? Attempts to answer this question will be further complicated by the fact that the IF6s correspond to only a portion of the amidinotransferases (184/360 residues). Moreover, the relative spatial locations of N- and C-termini within the IF6s are not the same in the amidinotransferases.

In the Groft et al.[1] paper, we utilized a pragmatic definition of protein structure similarity based on a widely used automated structure-structure comparison tool (DALI[5]). At the time of publication of our original paper[1], searching DALI with full-length proteins and each one of the five individual subdomains of either IF6 failed to detect either 1JDW or 1BWD. However, recent modifications to the DALI server now allow the relationship between eIF6 (or aIF6) and 1JDW to be identified (no relationship is returned for 1BWD because the DALI server considers 1JDW to be representative of both structures). Comparative protein structure modeling, which relies on primary sequence similarity, also failed to detect the fact that 1G61, 1G62, 1JDW, and 1BWD are similar. Subsequent pair-wise comparisons revealed the following root-mean-square-deviations: 1G62 versus 1JDW = 3.5 Å (184/360 α carbons, 8% identity) and 1G62 versus 1BWD = 3.5 Å (181/348 α carbons, 8% identity). For reference, 1G62 versus 1G61 = 1.5 Å (221/225 α carbons, 33% identity).

Paoli's commentary implicitly questions the validity of including the IF6s in a structural genomics pilot study. One long-term goal of the New York Structural Genomics Research Consortium is one experimentally-determined structure for each 30% identity protein sequence family[6]; a generally accepted compromise between the accuracy of comparative protein structure modeling[7] and the number of experimental structures required for reasonable coverage[8]. Our aIF6 structure yielded high-quality homology models of all known aIF6s and our eIF6 structure permitted modeling of all known eIF6s. Modeling of aIF6 with eIF6 (33% identical) and vice versa, gave models that were largely correct but contained non-trivial local errors reflecting structural variation between IF6s from distinct branches of the evolutionary tree. Both IF6 modeling with the amidinotransferase structures and amidinotransferase modeling with the IF6 structures are effectively precluded by low amino acid sequence identity (8%), whereas homology modeling within the amidinotransferase family itself is possible (1JDW versus 1BWD, 39% identical)

Paoli's final point regarding the need for more and better bioinformatic tools is well taken, and we believe that structural genomics will play an important role in the development of these much needed new technologies.

Caroline M. Groft, Roland Beckmann, Andrej Sali and Stephen K. Burley

Rockefeller University, Molecular Biophysics and the Howard Hughes Medical Institute, 1230 York Avenue, New York, New York 10021-6399, USA. Correspondence should be addressed to S.K.B. email: burley@ rockvax.rockefeller.edu

1. Groft, C., Beckmann, R., Sali, A. & Burley, S.K. Nature Struct. Biol. **7**, 1156–1164 (2000).
2. Humm, A., Fritsch, E., Steinbacher, S. & Huber, R. EMBO J. **16**, 3373–3385 (1997).
3. Fritsche, E., Bergner, A., Humm, A., Piepersberg, W. & Huber, R. Biochemistry **37**, 17664–17672 (1998).
4. Holm, L. & Sander, C. Structure **5**, 165–171 (1997).
5. Holm, L. & Sander, C. Science **273**, 595–602 (1996).
6. Burley, S.K. et al. Nature Genet. **23**, 151–157 (1999).
7. Sanchez, R. & Sali, A. J. Comp. Phys. **151**, 388–401 (1999).
8. Vitkup, D., Melamud, E., Moult, J. & Sander, C. Nature Struct. Biol. **8**, 559–566 (2001).