An elusive propeller-like fold

Determination of protein structures has shown that often distinct sequences are compatible with the same fold. It is believed that the number of possible folds is limited relative to the number of, say, genes in the human genome. Some research efforts therefore aim to structurally characterize new proteins which have no detectable homologs, and thus eventually elucidate all folding topologies[1]. With this structural genomics initiative, the New York Structural Genomics Research Consortium set out a number of targets for structure determination, one of which was the ribosome anti-association factor IF6. Reported in the December 2000 issue of *Nature Structural Biology*, IF6 was said to have a novel fold[2]. However, this is not the case since the same fold was first revealed by the structure of the enzyme L-arginine:glycine amidino transferase[3]. The fold has a modular construction of five units packed into a circular assembly with a five-fold pseudo symmetry. Each module consists of a ββαβ motif. Fig. 1*a* highlights the overall fold of amidino transferase and IF6, and Fig. 1*b* shows one ββαβ module from each of the two proteins. The repeating units are topologically identical.

The fold adopted by IF6 and amidino transferase is reminiscent of the β-propeller architecture which has been found in numerous proteins[4,5]. For of this reason, it was included in the SCOP[6] structural classification under the name of β/α propeller fold. Indeed, the modular ββαβ motif is equivalent to the four-stranded β-sheet motif that is the basic building block of the canonical β-propeller structure, with the difference that the third strand is substituted with a helical segment. In the β-propeller fold the ends of the polypeptide chain are generally 'sealed' in the circular array. The arrangement of joining the termini in one of the modular β-sheets is called 'Velcro'[7]. Interestingly, both amidino transferase and IF6 have a system of ring closure[5], just like many propeller domains. The name of β/α propeller fold is therefore appropriate.

β-propeller proteins are characterized by extreme diversity in sequence, function and phylogenetic origin. This diversity is shared by β/α propellers. In fact, despite their structural similarity, amidino transferase and IF6 have no detectable sequence identity, and are functionally unrelated. Although searches for structural homologs of IF6 with DALI[8] did not give hits, submission of the coordinates to the TOPS[9] server (*www3.ebi.ac.uk/tops*) results in clustered matches (with significance scores 18–20) including exclusively L-Arg:Gly amidino transferase (PDB code 1JDW) and the homologous enzyme inosamine-phosphate amidino transferase[10] (PDB code 1BWD). Thus, it is possible to detect the similar scaffold of these proteins.

The compatibility of diverse sequences with the same folding topology is not a new finding, as β-propeller proteins have shown[5]. Structural genomics, then, has two difficulties to face. Firstly, proteins without sequence homology to any structurally characterized protein can still turn out to have a known fold. Secondly, fold recognition can be hindered by sequence insertions and structure irregularities within the same scaffold. Prediction of folds from sequences and detection of, in some elusive cases, fold similarity, remain a major challenge in structural biology.

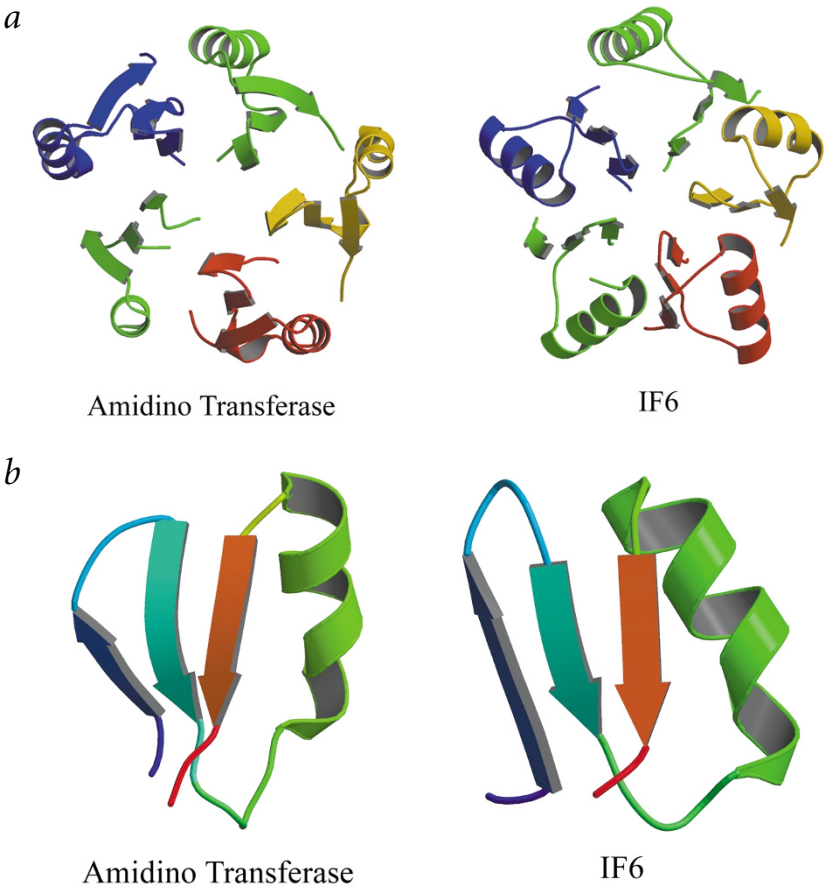

Fig. 1 The structures of the unrelated proteins L-Arg:Gly amidino transferase and ribosome anti-association factor IF6 share the same fold. ***a***, Schematic diagrams showing the overall fold of amidino transferase (PDB code 1JDW) and IF6 (PDB code1G61). The colors highlight the modular construction of the structures. For ease of viewing, loop regions and connecting peptides have been omitted. ***b***, The ββαβ modules that build the β/α propeller scaffold in these proteins have the same topology. The third module of each protein is shown.

Massimo Paoli

Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QW, England, UK. email: max@cryst.bioc.cam.ac.uk

1. Burley, S.K. *et al. Nature Genet.* **23**, 151–157 (1999).
2. Groft, C.M., Beckmann, R., Sali, A. & Burley, S.K. *Nature Struct. Biol.* **7**, 1156–1164 (2000).
3. Humm, A., Fritsche, E., Steinbacher, S. & Huber, R. *EMBO J.* **16**, 3373–3385 (1997).
4. Murzin, A.G. *Proteins* **14**, 191–201 (1992).
5. Paoli, M. *Prog. Biophys. Mol. Biol.* **76**, 103–130 (2001).
6. Murzin, A.G., Brenner, S.E., Hubbard, T. & Chothia, C. *J. Mol. Biol.* **247**, 536–540 (1995).
7. Neer, E.J. & Smith, T.F. *Cell* **84**, 175–178 (1996).
8. Holm, L. & Sanders, C. *Science* **273**, 595–602 (1996).
9. David, G. *et al. Bioinformatics* **15**, 317–326 (1999).
10. Fritsche, E. *et al. Biochemistry* **37**, 17664–17672 (1998).