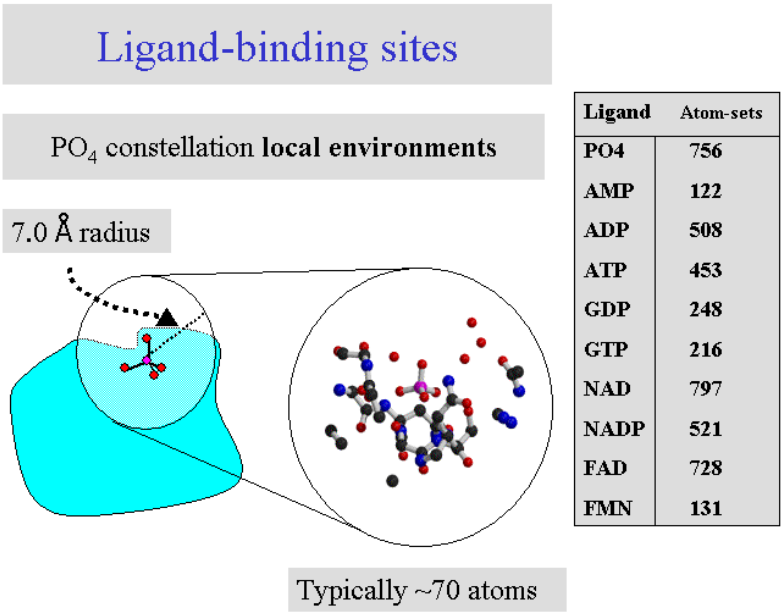
A new method for comparing ligand binding sites in biomolecules

Andreas Brakoulias, Gareth Stockwell and Richard Jackson

Analysis of completed genomes from a number of organisms reveal that about half of all gene products are classified as functionally unknown hypothetical proteins, and structures are already being solved prior to any knowledge of function. This will make the structure-based prediction of molecular function of increasing importance in the future. The ability to detect and classify local atomic level similarity will allow the development of rules for the prediction of function directly from structure (the premise of Structural Genomics).

Initially, proteins that bind ligands containing the phosphate moiety are being studied. In particular, the large class of nucleotide ligands (ATP/ADP, GTP/GDP, NAD, FAD etc.).



Ligand	Atom-sets
PO4	756
AMP	122
ADP	508
ATP	453
GDP	248
GTP	216
NAD	797
NADP	521
FAD	728
FMN	131

Figure 1. The local protein-ligand atomic environments are extracted from the PDB for proteins that bind phosphate containing ligands.

We have developed a method that identifies the extent of the similarity between the structure of different protein-ligand binding sites allowing their structural classification.

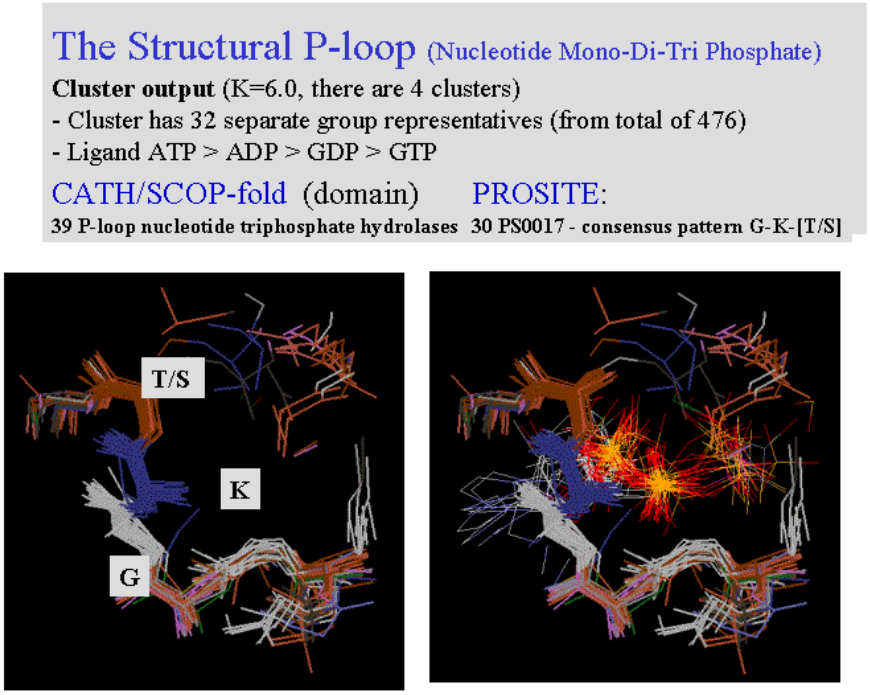


Figure 2. Following an all-against-all comparison of nucleotide phosphate binding sites group representatives are clustered to identify similar sites. This figure shows one such cluster (main-chain atoms only) which corresponds to the structural P-loop. The residues are coloured according to residue type. It can be seen that members of the cluster contain an identifiable sequence pattern identified in the PROSITE database (GK-[T/S]). The nucleotide moieties are also shown in the right hand picture.

The classification procedure allows the generation of structural binding site "templates". These could be used to identify similarity between the binding site of proteins in the database (for which an enzyme mechanism or binding site is well determined) and a newly determined protein of unknown function. Clearly, knowledge that a protein has a nucleotide binding site as well as information on the function of other proteins in the database that share this site will be valuable.

Collaborators

Prof. Janet Thornton (EBI)

References

Brakoulias A. & Jackson R. M. (2002) A new method for comparing ligand binding sites in biomolecules. *Manuscript in preparation.*

Funding

We wish to acknowledge support of The Wellcome Trust, BBSRC and Onasis Public Benefit Foundation.