

Fold independent structural comparisons of ligand binding sites in protein structures for function prediction

Nicola D. Gold* and Richard M. Jackson

University of Leeds

The rapid growth in protein structural data and the emergence of structural genomics projects have increased the need for automatic structure analysis and tools for function prediction. Small molecule recognition is critical to the function of many proteins, therefore determination of ligand binding site similarity is important for understanding ligand interactions and may allow their functional classification. Here, we present a Binding Sites DataBase (BSDB) that given a known protein-ligand binding site allows rapid retrieval of other binding sites with similar structure independent of overall sequence or fold similarity. However, each match is also annotated with sequence similarity and fold information to aid interpretation of structure and functional similarity. Similarity in ligand binding sites can indicate common binding modes and recognition of similar molecules, allowing potential inference of function for an uncharacterised protein or providing additional evidence of common function where sequence or fold similarity is already known. Alternatively, the resource can provide valuable information for detailed studies of molecular recognition including structure-based ligand design and in understanding ligand cross-reactivity. Here we show examples of atomic similarity between superfamily or more distant fold relatives as well as between seemingly unrelated proteins. Assignment of unclassified proteins to structural superfamilies is also undertaken and in most cases substantiates assignments made using sequence similarity. Correct assignment is also possible where sequence similarity fails to find significant matches, illustrating the potential use of binding site comparisons for newly determined proteins.

References

- Brakoulias, A., Jackson, R.M. (2004). Towards a Structural Classification of Phosphate binding sites in protein-nucleotide complexes: an automated all-against-all structural comparison using geometric matching *Proteins: Structure, Function, and Bioinformatics*, **56**, 250-260.
- Stark, A., Sunyaev, S. and Russell, R.B. (2003). A model for statistical significance of local similarities in structure. *JMB*, **326**, 1307-1316.