Detection of Similarity between Metal-Containing Protein Cavities

Simon Cottrell and Gerhard Klebe

Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35032 Marburg, Germany

Introduction

CavBase is a module of the Relibase+ database system for the storage of protein crystal structures [1]. It consists of a database of protein surface cavities, and enables searching of the Protein Data Bank (PDB) or in-house databases for cavities that are similar to a given query cavity. The CavBase similarity measure is based purely on the spatial arrangement of functional groups and surface patches. Therefore, it can detect similarities and potential cross-reactivity between cavities from proteins unrelated in fold or sequence [2].

Here we present a further development of CavBase, in which metal ions bound to the protein, which were previously ignored, are now taken into account and are effectively considered as part of the protein.

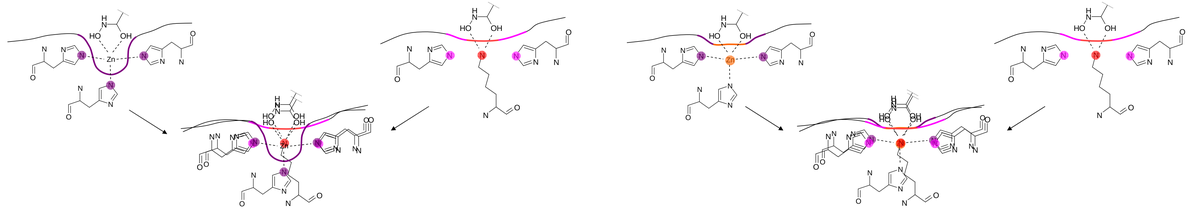
Introduction of Metal Pseudocentres

The generation of the CavBase data for a protein takes place in three stages. Firstly, the cavities themselves are detected using a grid-based method, and a surface is generated. Secondly, the functional groups exposed to the surface are represented by pharmacophore-like points known as “pseudocentres”, each with a type such as “hydrogen-bond donor”. Thirdly each pseudocentre is assigned a surface patch, consisting of those surface points to which it is nearest.

The cavity generation is not affected by the presence of ligands, although the ligands that lie in each cavity are stored in the database. In the original version of CavBase, metal ions were ignored during cavity generation, since they are generally treated as ligands within Relibase.

However, metal ions are often much more strongly bound to the protein than drug-like ligands. Therefore, when predicting the binding of such ligands, the metal ions should effectively be considered as part of the protein.

To account for this, the data generation algorithm has been modified so that the cavity surface now encloses metal ions as part of the protein. A new metal pseudocentre type, positioned on metal ions, has been introduced. Since metal ions are generally able to bind to the same types of ligand groups as hydrogen-bond donors in the protein, we allow metal pseudocentres to be matched to donor and donor-acceptor pseudocentres when performing a CavBase similarity search (Figure 1).

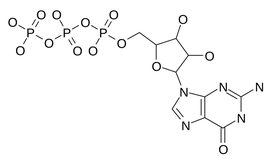


(a) without metal pseudocentres – poor surface overlap *(b) with metal pseudocentres – good surface overlap*

Figure 1 – Overlay of a metal- and a non-metal-containing cavity binding the same ligand, with and without metal pseudocentres

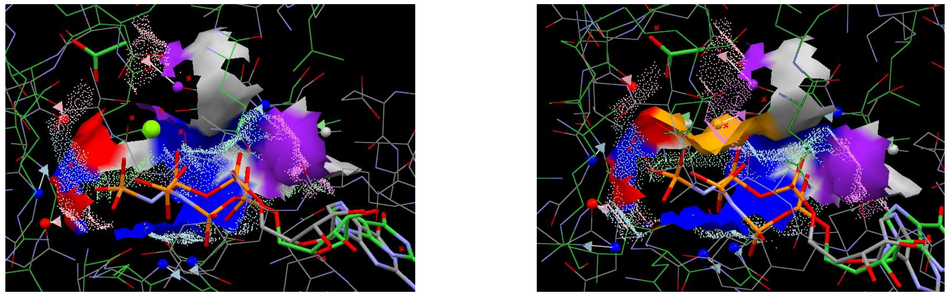
Results

The effects of introducing the metal pseudocentres are illustrated here using a dataset of 130 cavities from GTPases. The GTPases are a large family of hydrolase enzymes which bind GTP:



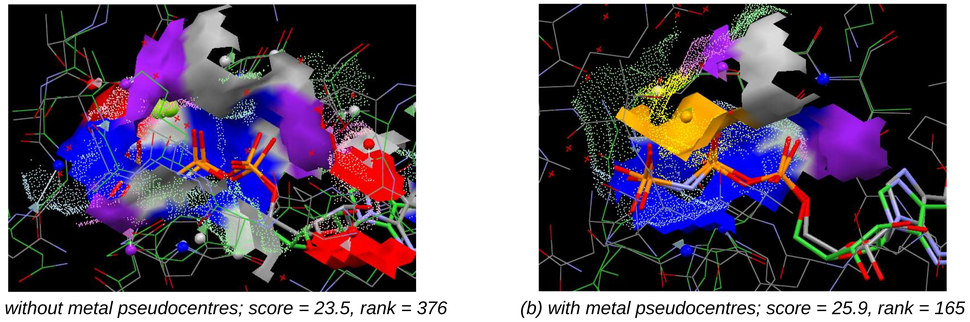
Most, but not all of the cavities bind GTP via a magnesium ion. CavBase similarity searches were carried out using an Mg^{2+}-containing GTPase cavity (PDB code 1k5d.38) as the query and a database consisting of the other GTPase cavities and approximately 1600 randomly selected cavities from the PDB. Two versions of the database, with and without metal pseudocentres, were compared.

In many cases CavBase was able to identify the similarity between the GTPases even without metal pseudocentres, due to the similarity of the rest of the binding sites. However, cases were also found where the introduction of metal pseudocentres led to a significant improvement in the similarity ranking of a GTPase cavity (Figures 2 and 3). In these cases, there was less similarity in the amino acid structure of the GTP binding sites.



(a) without metal pseudocentres; score = 27.6, rank = 172 *(b) with metal pseudocentres; score = 28.3, rank = 138*

Figure 2 – Overlay of the query cavity (grey carbons; solid surface) with the database cavity 1jpn.6 (green carbons; dotted surface). The latter contains no metal ion, but in overlay (b), the metal pseudocentre of the query (orange surface patch) has been matched with a donor-acceptor pseudocentre of the database cavity (pink surface patch)



(a) without metal pseudocentres; score = 23.5, rank = 376 *(b) with metal pseudocentres; score = 25.9, rank = 165*

Figure 3 – Overlay of the query cavity (grey carbons; solid surface) with the database cavity 1kk1.3 (green carbons; dotted surface). In overlay (b), the matched surface patches of the metal pseudocentres are shown in orange (query) and yellow (hit).

Conclusions

The introduction of metal pseudocentres provides a more accurate representation of the cavity surface, and in some cases leads to a significant increase in the similarity score of cavities binding the same ligand.

However, we have identified possibilities for further improvement. Metal ions are generally found in a “feature-dense” area of the binding site, and hence have quite small surface patches. By giving a greater weighting to metal pseudocentres than, for example, aliphatic pseudocentres, the importance of metal ions in ligand binding would be better recognised.

References and Acknowledgments

1. Schmidt, S., Kuhn, D., Klebe, G. A new method to detect related function among proteins independent of sequence and fold homology. *J. Mol. Biol.*, **2002**, *323*, 387-406.
2. Weber, A., Casini, A., Heine, A., Kuhn D., Supuran, C.T., Scozzafava, A., Klebe, G. Unexpected nanomolar inhibition of carbonic anhydrase by COX-2-selective celecoxib: new pharmacological opportunities due to related binding site recognition. *J. Med. Chem.*, **2004**, *47*, 550-557.

We thank Robin Taylor, Jason Cole and Greg Shields of the Cambridge Crystallographic Data Centre for helpful discussions and technical assistance. This work was funded by the Cambridge Crystallographic Data Centre.