MetalMine: A database of transition metal-binding sites in proteins

Kensuke Nakamura
kensuke-nm@is.naist.jp

Aki Hirai
hirai@gtc.naist.jp

Hiroki Takahashi
hi-takah@bs.naist.jp

Md. Altaf-Ul-Amin
amin-m@is.naist.jp

Shigehiko Kanaya
skanaya@is.naist.jp

Graduate School of Information Science, Nara Institute of Science and Technology (NAIST)
(Keihanna Science City) 8916-5, Takayama-cho, Ikoma, Nara, 630-0192, Japan

Keywords: metallocene, metallomics, metal-binding protein, database

1 Introduction

We present a new database system called MetalMine (http://metalmine.naist.jp/) that contains information about metal-ion-binding sites in proteins.

Transition metal ions such as iron, zinc, copper, manganese, nickel and several others are known to be essential for life. In most cases, transition metal ions exhibit their function in coordination with proteins, typically as a cofactor of enzymes. The functional and structural aspects of transition metal-associated enzymes have been extensively studied in the field of bioinorganic chemistry for a long time. Approximately, one-third of all structurally characterized proteins contain metal ions [1]. The recent emergence of terms such as metallocene and metallomics [4] indicates the increasing interest toward the study of these metallocproteins.

Despite the importance of the field, to our knowledge, there have been no organized list of metal binding proteins publicly available, and therefore we intended to start this MetalMine project. A similar effort has been made by a group at Scripps Research Institute, namely, the Metalloprotein Database and Browser (MDB) [2], which unfortunately has not been updated recently. There are some books available which covers metallocproteins such as the anthology edited by Messerschmidt et al. [3].

2 Data collection and classification of metal binding sites

The binding sites for metal ions in the protein structures in the PDB database are searched automatically using a computer program. A metal-ion-binding site is defined as a cluster of metal ions coordinated by several protein amino acid residues, and heteromolecules such as cofactors, substrates, and water.

Metal-binding sites are then classified according to the structural domain that the metal-coordinating residues belong to. We use Structural Classification of Proteins (SCOP) for the definition of structural domain [3]. Some proteins contain more than one metal binding sites in one domain, and they are distinguished.

In the PDB database, there are many metal ions artificially involved, for instance, due to the additive during the crystallization process. These artificial sites are excluded by manual curation.

3 Database structure and search capabilities

The webpage is designed to be simple and self-explanatory. It includes a top menu and a sidebar, with which users can access the information they need (Figure 1). By choosing a kind of metal ion in the sidebar, it shows a list of metal binding sites including the specified metal ion. The list contains information such as the name of proteins, SCOP fold id, PDB id of the representative structure, EC id and a link to Wikipedia,
whenever available. The top page includes a text field and a search button to enable MetalMine to be searched using a PDB id as a query. It also has a Google search box, with which users can search MetalMine by keywords such as protein names. BLAST search is also available from the top menu. With this function, residues homologous to known metal-binding sites in the query amino acid sequence are indicated.

![Image of MetalMine front page]

Figure 1: Front page of the MetalMine

4 Conclusion and future work

As of September 2009, MetalMine contains 386 entries of metal binding sites (14, 31, 76, 25, 14, 1, 4, 88 and 133 for Co, Cu, Mn, Mo, Ni, V, W, Fe and iron-sulfur clusters, respectively), with 6308 instances of binding site structures from the PDB database. We are now curating entries for zinc and hem-iron, which will be available shortly. We believe that MetalMine contains the largest number of reliable entries among publicly available databases for metalloproteins.

References


