WALIA journal 30(S3): 33-42, 2014
Available online at www.Waliaj.com
ISSN 1026-3861

A review on using bioinformatics web servers, softwares and databases for prediction and analysis of metal binding sites in proteins

Zahra Golestannejad [1], Shahin Gavanji [2,*], Mohsen Doostmohammadi [2]

[1]*Department of Oral Medicine and Torabinejad Dental Research Center, School of Dentistry, Isfahan University of Medical Sciences, Isfahan, Iran*
[2]*Young Researchers and Elite Club, Khorasgan Branch, Islamic Azad University, Khorasgan, Iran.*

Abstract: Bioinformatics, as an interdisciplinary science, seeks to find various biological issues in molecular level using available techniques in computer science, mathematics, chemistry, physics and other relevant sciences. One part of this science concerns the sequences and proteins absorbing metals in which many softwares and servers are widely used for bioinformatics prediction. More than half of proteins in Protein Data Bank possess metals and thus it expected that most of these proteins are metalloproteins. Metals are necessary for proteins function and structure. Since metal ions are essential in too many biological functions, mainly in metalloproteins, finding metalloproteins is of great importance in the biological and medical fields. Metal binding site prediction is considered as a step in function assignment for new proteins. In this review, we tried to study softwares and servers which are found to be pioneer in finding metal binding sites in proteins and introduce them briefly. In this way, the importance of metal ions is explained and then the binding sites of proteins like CHDEs have been partly introduced.

Key words: *Prediction; Metal; Bioinformatics; Servers; Softwares*

1. Introduction

Many different metal ions like Cu^{2+}, Zn^{2+}, Mn^{2+}, Fe^{2+}, Ni^{2+} and Co^{2+} are micronutrients that are found to be essential for growth and metabolism of plants (Ideker *et al.,* 2001; Frausto *et al.,* 2001). High soil existential concentration of these ions is extremely toxic (Ideker *et al.,* 2001; Sharma *et al.,* 2009). Because of strength, accumulation and persistence of the ions in ecosystems, heavy metals have expected to receive special consideration (Ahluwalia et *al.,* 2007; Machado *et al.,* 2008; Volesky, 2001). Heavy metal referred to a chemical element with an atomic mass greater than 22 and a density greater than 5 g/mL. 69 elements include in this classification, of which 16 are synthetic. Some of them even at very low concentrations are toxic to human beings (Ashok et al., 2006; Wang and Chen, 2006)copper (Cu), zinc (Zn), silver (Ag), lead (Pb), mercury (Hg), arsenic (As), cadmium (Cd), chromium (Cr), strontium (Sr), cesium (Cs), cobalt (Co), nickel (Ni), thallium (Tl), tin (Sn) and vanadium (V) are the most important heavy metals associated with environmental contamination (Byrne *et al.,* 1980; Shore *et al.,* 1987; Tubek *et al.,* 2008),and are shown to be dangerous for ecosystem (Wang and Chen, 2006). Generally metal ions classified in three groups include: essential, toxic, radioactive and semi metal ions. Na, K, Mg, Ca, V, Mn, Fe, Co, Ni, Cu, Zn, Mo and W are known essential ions that are important for plants metabolism (Williams, 2001; Ashok *et al.,* 2009). Toxic ions (Hg, Cr, Pb, Cd, As, Sr, Ag, Si, Al, Tl) show no biological functions. Radioactive ions are also toxic for cells including U, Rn, Th, Ra, Am, Tc and semi metal ions (B, Si, Ge, As, Sb, Te, Po, At, Se) have distinct functions in plant metabolism (Ashok *et al.,* 200; Nobuaki *et al.,* 2002; Rhee *et al.,* 1998) (Rhee et al., 1998). The main feature which has determined metals bioavailability and their destination is their ionic form. Most of heavy metals are categorized as cationic ions and this caused to their sorption to negative groups like cells (Saedi *et al.,* 2013). Heavy metal ions possess high electrostatic attraction and their high binding affinities to sites that essential ions usually bind in cellular structures and caused to bio molecules structures disability leading replication defects, cancer and hereditary genetic disorders. Arsenate, for example, competes with phosphate and cadmium competes with zinc (Mehrasa et al., 2014). According to microarray analysis (Kawata *et al.,* 2007; Passerini *et al.,* 2006), six heavy metals (arsenic, cadmium, nickel, antimony, mercury and chromium) induce gene expression patterns similar to pattern induced by DMNQ (2.3-dimethoxy-1, 4-naphthoquinone), and caused oxidative damage through producing ROS and inactivating the cellular antioxidant system (Liu *et al.,* 2005; Mannazzu *et al.,* 2000). Metalloproteins are diverse category of proteins which contain one or more metal ions in their conformation. Metal ions are fundamental to protein function and play many different regulatory, structural or catalytic roles (Christopher *et al.,* 2002; Chance and Shi, 2008; Bertini and Cavallaro, 2010). For instance, zinc ions stabilize the structure of transcription factors like

* Corresponding Author.

zinc fingers. In enzymes, metal ions are normally cofactors that increase catalytic activity (Christianson, 1991). There are also many different processes including apoptosis, aging and carcinogenesis that severely depend upon metal binding proteins. Finding metal binding sites in novel proteins considerably lead to recognition their functional characteristics and will clear metal related malfunctions. Recently, many different laboratory techniques have been proposed for identifying metalloproteins. X-ray absorption spectroscopy (HT-XAS) for example (Shi *et al.,* 2011), is able of discovering metalloproteinase with high reliability (Shi *et al.,* 2005; Shi and Chance, 2011, Passerini *et al.,* 2011). However, ligands which are related in binding the metal ions cannot be identified by this technique. Bioinformatics tools can considerably contribute to a detailed explanation of metal binding sites, as well as in scaling-up to proteome-wide analyses (Passerini *et al.,* 2011; Degtyarenko, 2000; Sodhi *et al.,* 2009). Motif based approaches, based on regular expression patterns or Pfam probabilistic models and applied for sequence based predictions on proteins(Passerini *et al .,* 2011; Shi *et al .,* 2011). These methods cannot identify new sites while regular expression patterns tend to be wholly specific with low coverage, and Pfam models are incomplete to known metal binding sites. Recently, some supervised learning techniques have been proposed for predicting the metal binding state of all residues in a sequence (Passerini *et al.,* 2011; Gavanji *et al.,* 2013). In this paper, we tried to introduce some useful servers and software's for prediction and analysis of metal binding sites.

2. Methodology

There are many different servers and sofwares used to predict metal binding sites with high accuracy. Only running a program would not lead to prediction of the exact ions which we are looking for. There are many different methods by which these programs used to predict metal binding sites. In order to use any of them, it is necessary to be aware of the method of prediction and their results. Here, we studied some of this software's and servers which are more useful and tried to explain them in brief. CHDE algorithm, for instance, is explained because of its importance in metal ion binding proteins and some more useful servers such as MetaRouter, Metalmine, Discovery Studion and metalDetector have been presented (Gavanji *et al.,* 2013).

2.1. MetaRouter

MetaRouter is applied for identifying heterogeneous information concerning biodegradation which permits to administrating and extracting new data (Pazos *et al.,* 2005). It is a practical application for laboratories which need to keeping data and extracting information out of it. This program is realized in Postgre SQL referred to standard language for relational databases and used client/server architecture, so it can be accessed quite easily only by having a web browser (Kanehisa *et al.,* 2004). MetaRouter is an application for laboratories operation in biodegradation and bioremediation which necessitate to maintain and consult public and private data, linked internally and with external databases, and also to extract new information from it. The system can be available and administrated through a web interface. The full featured system, except administration facilities, is freely available at (http://pdg.cnb.uam.es/MetaRouter) (Fulekar and Jaya, 2008; Pazos *et al.,* 2005).

2.2. MetalMine: a database of functional metal-binding sites in proteins

MetalMine is a very helpful database for discovery of metal binding sites. The significance of metal ions in life system through the emergence of terms such as Metallome and Metallomics is considerably clear (Nakamura et al., 2009; Jernigan et al., 1999). However, there is no complete list of metal binding proteins. A metal binding site is distinct, in this search, as a collection of metal ions coordinated by several amino acid residues with heteromolecules like cofactors, substrates and water. Structural Classification of Proteins, or SCOP for short (Murzin *et al.,* 1995), is used for structural domains definition. A metal binding site family is distinct as homologous metal binding sites (Messerschmidt *et al.,* 2001). In some cases, metal binding site may contain different amino acid residues from different structural folds. Therefore, metal binding site can be positioned at boundary of two domains in a single chain, or at the boundary of two protein subunits. The purpose of this program is to set up an organized list of functional metal binding sites (Nakamura *et al*., 2009). Routinely, extraction of metal binding sites from PDB database will result in many metal binding sites with no biological function which are typically the result of experimental conditions. MetalMine seeks to exclude artificial coordination by manual curation (Andreini *et al.,* 2009; Gasteiger *et al.,* 2003).

2.2.1. Method of use

Each page has a top menu and a sidebar. From the top menu, one can contact pages for a description of the database, a BLAST search, and tutorial. The top page includes a text field and a search key to permit MetalMine to be searched using a PDB ID as a query. From the sidebar, users can look through the information of metal ions contained in MetalMine. Once one of the metal ions located in the sidebar is chosen; a table of metal-binding sites will be shown. There are columns inside the table that list the names used to identify the metal-binding site, SCOP IDs of residues concerning the metal-binding site, the type and number of metal ions and residues, the representative PDB IDs, number of PDB files recognized for this site, whole number of sites found

in the entire PDB, links to Wikipedia and Nice Zyme entries at ExPASy when available, and so on. By clicking on the name of a site one can see a list of all instances of the metalbinding site (Gasteiger *et al.,* 2003). This list contains columns that explain the PDB IDs, the residues and the heteromolecules like cofactors and Ligands (Herraez, 2006). Moreover, a Jmol window is existed to show the local metal-binding structure when an instance of the metal-binding site is distinct by clicking on the PDB ID (Punta and Ofran, 2008; Bernstein *et al.,* 1977). On the right side of the Jmol window, more options are accessible for looking over the metal binding sites, such as a button for displaying the second layer of coordinating residues (Nakamura *et al.,* 2009).

2.2.2. Search by amino acid sequence

In MetalMine, an amino acid sequence search can be applied via BLAST. For this purpose, a series of the amino acid sequences from the PDB structures is put in MetalMine. Determining an amino acid sequence as a query, BLAST performs a sequence looking of the database for a match with the metal-coordinating residues in MetalMine (Altschul *et al.,* 1990). Consequently, only the exact matches will be shown, and poor matches that their E-value is greater than 0.0001, are distinct as low reliability matches. The output of a BLAST search includes the amino acid sequence used as the query, the hit residue highlighted with magenta or light cyan for a regular or low-reliability match, respectively, followed by a list of matching residues with links to the metal-binding sites contained in MetalMine (Babor *et al.,* 2008). The BLAST search function in MetalMine can be applied as a device to predict metal-binding residues in amino acid sequences. One can use threading model and force field model to predict metal binding residues based on structural information (Goyal and Mande, 2008; Sodhi *et al.,* 2004). An empirical method is based the comparing holo-apo pairs of known metal binding sites and its results have very high reliability. Also the *de novo* approach based on a sequence using a machine-learning method has been developed, (Lippi *et al.,* 2008) while its prediction ability, compared with the structure based predictions, appears to be limited (Gavanji *et al.,* 2013; Sodhi *et al.,* 2011).

2.3. CHDEs

The exact prediction of zinc-binding proteins and zinc-binding sites from sequences are of attention of researchers. Zinc is considered as the most common transition metals bound to proteins. Cys, His, Asp and Glu (CHDE) are four amino acids which account for almost 98% of all residues that bind to zinc. Among four mentioned amino acids, Cys and His (CH) are dominant, which account for almost 84% of all zinc-binding residues. One zinc atom binds to three or four amino acid residues. Zinc bound by three amino acid residues are often catalytic zinc and those bound by four are usually structural zinc (Shu *et al.,* 2008). SVM-based predictor and a homology-based predictor are two method of CHDE. In the SVM-based predictor, CHDEs could be selected in both training set and test set and encode into single-site vectors and pair-based vectors which display a window of residues in the center of each selected CHDE or a pair of selected CHDEs respectively. The available Gist SVM package (Sodhi *et al.,* 2011; Punta and Ofran, 2008) could be used for implementing SVM. SVM predictions on individual selected residues obtained through combination of the predictions by using single-site vectors and pair-based vectors through a gating network. Zinc binding sites can be predicted from amino acid sequences by combining SVM predictions and homology-based predictions. Shu and his coworkers (2008) applied this method and predicted Cys, His, Asp and Glu with 75% accuracy at 50% recall level, when they tested on a non-redundant set of PDB containing 2727 unique protein chains. The success rate could be even higher, if homologues were predicted: for Cys, His, Asp and Glu with 76% accuracy (90% for Cys and His) at the 70% recall level (Shu *et al.,* 2008). The predictions were so reliable so that some occasional presumed errors of PDB concerning zinc-binding had been found (Shu *et al.,* 2008; Pavlidis *et al.,* 2004).

2.3.1. Metal binding sites using 'CHED' algorithm

The 'CHED' algorithm is capable to predict 3D intra-chain protein binding sites in transition metals such as Zn, Fe, Mn, Cu, Ni, Co, and Ca and Mg sites that can be substituted by a transition metal. The algorithm looks for a triad of amino acids composed of 4 residue types (Cys, His, Glu, Asp) which possess ligand atoms in specific distances. It permits one target residue to rotate in rotamer space, considering structural rearrangements that may arise upon metal binding. A binding site is considered to be accurate if one or more correct amino acid ligands have been predicted (Shu *et al.,*2008; Sodhi *et al.,* 2004; Mika and Rost, 2003).Two algorithms are used to filter out false positives: MILD FILTER which is based on the frequency of hits, yields high sensitivity. STRINGENT FILTER which uses decision tree and support vector machine technology, yields high selectivity (Shu *et al.,* 2008; Nanjiang *et al.,* 2008),

2.4. Metal Detector Predicts v2.0 software

Metal detector 2.0 predicts binding to transition metals (Passerini et *al.,* 2011), which create about 66% of the PDB metallo-chains and consider iron and zinc as the two most basic ions in cellular functioning. Metal Detector uses protein sequences as input for identification of CYS and HIS which have role in transition metal protein binding sites. One of its main abilities is predicting which residues are mutually responsible in the coordination of the same metal ion. The server is available at http://metaldetector.dsi.unifi.it/v2.0/ (Frasconi and

Passerini, 2009; Passerini *et al.,* 2011). Focusing on CYS and HIS only, about 74% transition metal ligands are covered. For computational efficiency reasons, it is supposed that each ligand binds just one ion. This is almost always the case for CYS and HIS. It is not a full 3D description of the sites geometry, with angles and distances, but rather the prediction of the connection between ions and their ligands. Bonding condition of each CYS and HIS is predicted in two states including metal-bound or not (Lippi *et al.,* 2008; Shi *et al.,* 2011). When the chain is not belonged to a metalloprotein, bonding state for CYS is predicted in three states as disulfide-bound, metal-bound and free. Chains with both disulfide bridges and metal binding sites are very uncommon (less than 3%) (Fig. 1).

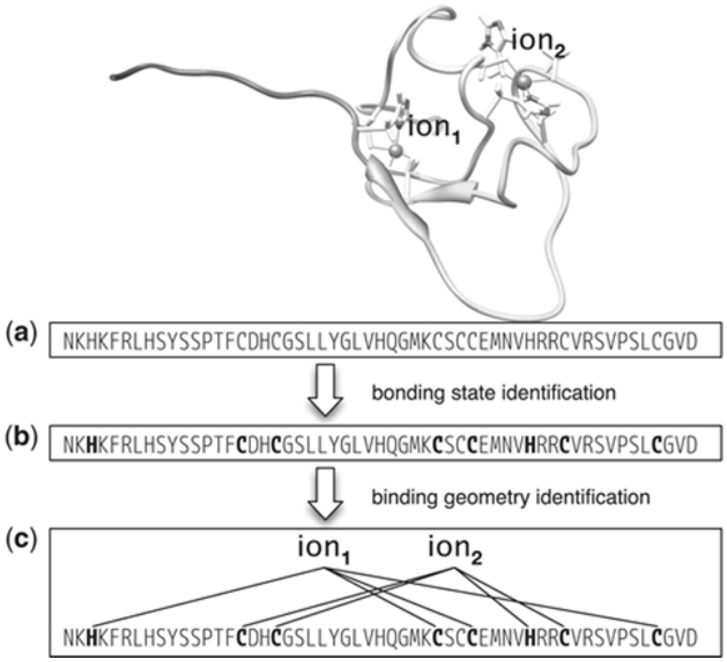


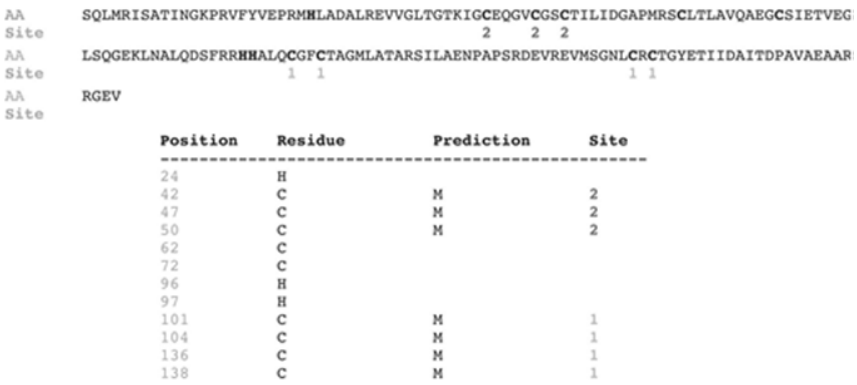
Fig. 1: Metal binding prediction subtasks. (**a**): given sequence; (**b**) candidate ligands (CYS and HIS) are assigned bonding state (boldface for metal binding); (**c**) metal-binding residues are grouped to form binding site configurations (Passerini *et al.,* 2001)

2.4.1. Method of use

In the first step, SVM-HMM is used to predict metal bonding state of each CYS and HIS. The algorithm uses dynamic programming to find the best general bonding state assignment to all CYS and HIS in the input chain. After that, residues predicted to be ligands are collected together to shape binding sites (Sodhi et al., 2004; Bakir et al., 2007). This could be achieved throughan ad-hoc algorithm whichexploits the assumption that each ligand binds exactly one ion. The web interface permits to decide between three different settings, corresponding to the three different paths represented in Figure 2: no prior knowledge (default operation mode); (ii) the chain is identified to belong to a metalloprotein; (iii) the chain is identified to belong to a metalloprotein, and the user can also provide the bonding state of each CYS and HIS. Output is either accessible on a separate web page or delivered by through e-mail. Residues predicted to match the same ion will share the same identifier (Passerini *et al.,* 2006). Every identifier is an integer ranging from 1 to 4 indicating maximum number of binding sites that can be predicted (Ceroni *et al.,* 2006). Its value shows no particular biochemical semantics but lower values relate to a higher level of confidence for the predictor. Figure 2 shows a web browser output for PDB entry 1t3qA (Shu *et al.,* 2008; Lippi *et al.,* 2008).

2.5. PredZinc

PredZinc is a program for predicting zinc-binding sites in proteins from their amino acid sequences. The program is written in c/c++ and bash shell scripts (Shu *et al.,* 2008; Vallee and Auld, 1990). Currently, PredZinc can be run on Linux and Windows (with Cygwin). PredZinc is copyrighted (c) to Nanjiang Shu, Structural Chemistry, Stockholm University, Sweden and is free for academic use (Shu *et al.,* 2008; Shore *et al.,*1978). The main focus is on predicting four types of amino acids, i.e. cysteines, histidines, asparates and glutamates (CHDE) (Kristen *et al.,* 2012; Maurer-Stroh *et al.,* 2013), in order to cover most residues of interest while enabling reasonable prediction accuracy. A five-fold cross-validation indicates that this server predicts zinc-binding Cys and His with 76% precision at 60% recall and on the chain level 60% precision at 60% recall (Menchetti *et al.,* 2006; Passerini *et al.,*2006).


AA SQLMRISATINGKPRVFYVEPRMHLADALREVVGLTGTKIGCEQGVCGSCTILIDGAPMRSCLTLAVQAEGCSIETVEG
Site 2 2 2
AA LSQGEKLNALQDSFRRHHALQCGFCTAGMLATARSILAENPAPSRDEVREVMSGNLCRCTGYETIIDAITDPAVAEAAR
Site 1 1 1 1
AA RGEV
Site

Position	Residue	Prediction	Site
24	H		
42	C	M	2
47	C	M	2
50	C	M	2
62	C		
72	C		
96	H		
97	H		
101	C	M	1
104	C	M	1
136	C	M	1
138	C	M	1

Fig. 2: Web browser output of MetalDetector for PDB entry 1t3qA

2.5.1. Method of use

Sequences in FASTA format or Bare sequence format (just lines of sequence data, without definition) or any sequence format supported by EMBOSS Seqret (EMBOSS Seqret reads and writes (returns) sequences), are used in this program. After submitting the sequence, the results will be ready for download in the coming webpage shortly (Fig. 3). Under Advanced options, you may also choose the thresholds of recall and precision as well as the training set (Shu *et al.*, 2004).

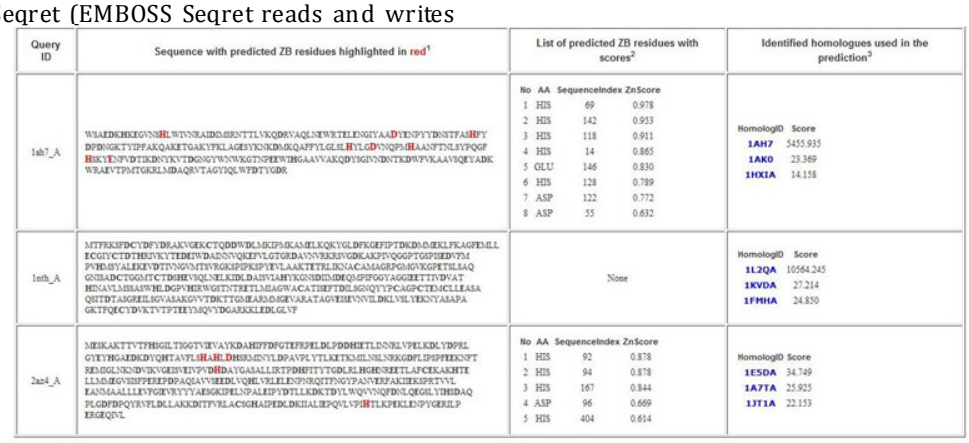


Fig. 3: Output of PrediZinc webserver 1Predicted zinc-binding (ZB) residues are highlighted in red and with larger font size. Cys, His, Asp and Glu are bolded. Residues are predicted as zinc-binding if the score is >= 0.450. Abbreviations: CYS: cystein, HIS: histidine, ASP: aspartate, GLU: glutamate (Shu *et al.*, 2004)

2.6. The PSIPRED structure analysis workbench

PSIPRED is a simple and accurate secondary structure prediction method, incorporating two feed-forward neural networks which perform an analysis on output obtained from PSI-BLAST (Position Specific Iterated - BLAST) (Jones and Swindells, 2002; McGuffin *et al .*, 2000; Buchan *et al.*, 2013). Using a very stringent cross validation method to evaluate the method's performance, PSIPRED 3.2 achieves an average Q3 score of 81.6% (Brylinski *et al.*, 2011). PSIPRED 2.0 achieved an average Q3 score of 80.6% across all 40 submitted target domains with no obvious sequence similarity to structures present in PDB, which ranked PSIPRED top out of 20 evaluated methods. UCL-CS Bioinformatics Web Servers (http://bioinf.cs.ucl.ac.uk/structure) (Sodhi et *al.*, 2004; McGuffin *et al.*, 2000).

2.7. FINDSITE-metal server

FINDSITE-metal is an extended version of FINDSITE that predicts metal-binding sites, residues and binding metal preferences from evolutionarily related templates discovered by threading (http://cssb.biology.gatech.edu/findsite-metal) (Brylinski *et al.*, 2011). Moreover, it uses machine learning to increase the prediction accuracy,mainly against low-to-moderate quality protein structures. So, crystal structures as well as protein models can be used as the targets. The server accepts only single protein chains 40-400 residues in length. Only the ATOM part is parsed and unusual residues are removed. PROSPECTOR_3, SP3 and Sparks2 threading algorithms are used to recognize evolutionarily related templates. One can upload either the crystal structure or the predicted model of his target. Estimated TM-score must be in the range of 0.0 to 1.0 (Shu *et al.*, 2008; Citiulo *et al.*, 2012).

2.8. GRID

GRID is a computational process which verifies energetically favorable binding sites on proteins consisting known structure (Von Itzstein *et al.,* 1993; Kastenholz *et al.,* 1977). It can be used for studying individual molecules like drugs, molecular arrays such as membranes or crystals, and macromolecules including proteins, nucleic acids, glycoproteins or polysaccharides. Several different molecules can be processed one by one (Ortuso et al., 2006; Wade and Goodford, 1993; Carosati, et al., 2004).

2.9. Disulfide

The server is for predicting the disulfide bonding state of cysteines and their disulfide connectivity starting from sequence alone. Disulfide bridges play a key role in the stabilization of the folding process for some proteins. Prediction of disulfide bridges from sequence is very useful for studying structural and functional features of particular proteins (Frasconi and Passerini, 2002; Ceroni *et al.,* 2003). This server predicts disulfide patterns through two computational steps: (1) the disulfide bonding state of each cysteine is predicted by a BRNN-SVM binary classifier; (2) those cysteines that are recognized to contribute in the construction of bridges are paired by a Recursive Neural Network to obtain a "connectivity pattern" (Vullo and Frasconi, 2004; Frasconi and Passerini, 2009).

Table 1: Summary of softwares and servers presented in this study

No	Software/Server	Introduction	Link
1	MetaRouter	Identifying heterogenous information related to biodegradation	http://pdg.cnb.uam.es/MetaRouter
2	MetalMine	Finding metal binding sites	http://metalmine.naist.jp/metalmine009/index.html
3	Discovery Studio		http://accelrys.com/products/discovery-studio/
4	Metaldetector	predicting binding to transition metals	http://metaldetector.dsi.unifi.it/
5	PredZinc	Predicting zinc-binding sites in proteins from their amino acid sequences	http://casio.fos.su.se/server/predzinc/index.php?about=predzinc
6	PSIPRED	secondary structure prediction	http://bioinf.cs.ucl.ac.uk/index.php?id=779
7	FINDSITE-metal	Predicting metal-binding sites, residues and binding metal preferences from evolutionarily related templates	http://cssb.biology.gatech.edu/findsite-metal
8	GRID	Determining energetically favorable binding sites on proteins with known structure	https://www.personal.reading.ac.uk/~sas97sca/Metal%20binding%20sites.htm
9	Disulfide	Predicting the disulfide bonding state of cysteines and their disulfide connectivity	http://disulfind.dsi.unifi.it/

3. Discussion

Metal ions play a significant role in living organisms. About one third of proteins have to bind metal for their stability and/or function. In this review, current sequence based and structure based methods for metal binding site prediction with highlighting the CHED methods of prediction are proposed. SVM-based predictor and a homology-based prediction presented as two method of CHDE. MetaRouter is a suitable system for bioremediation studies which could identify heterogeneous information useful for biodegradation. Finding functional metal binding sites is the final aim of MetalMine. Extracting metal binding sites from PDB database will lead to numerous binding sites lacking biological function. Using MetalMine, we could set up a list of functional binding sites which is not as complex as the data obtained from PDB database. Metal detector predicts transition metals with high speed and is able to predict which residues are jointly concerned in the coordination of the same metal ions. PrediZinc is another server which focuses on CHDE and predicts zinc binding sites with high accuracy. GRID determines favorable binding sites of proteins with known structure. Disulfide bridges are extremely importance in folding process of proteins and for finding them. Disulfide is a good instance of a simple server which predicts bonding state of cycteins from protein sequence.

4. Conclusions

In natural proteins, metal ions play a variety of roles, including nucleophilic catalysis, electron transfer and the stabilization of protein structure (Andreini *et al.,* 2008; Ukaegbu *et al.,* 2006). In this review, we introduced some useful software's and websites for predicting metal ion binding sites in proteins. Conserved structures and sequences are used for identification of metal ion binding residues. Sequence based and structure based methods for metal binding site predictions had been reviewed in this study. The CHED method of prediction from protein structures and translated gene sequences were described in detail respectively, as well as their web server applications. There are so many software's which predict metal binding ions, but they are not as much as necessary and still many

functional sites in proteins are not identified and need to be designed and improved accordingly.

Acknowledgments

Support of this research by University of Isfahan and Azad University is highly acknowledged.

References

Sharma. A., Roy. S., Tripathi. K.P., Roy. P., Mishra M., Khan. F., Meena. A. 2009. Predicted metal binding sites for phytoremediation. Bioinformation., 4, pp. 66-70.

Ahluwalia. S.S. and Goyal. D.; 2007. Microbial and plant derived biomass for removal of heavy metals from wastewater. Bioresour Technol., 98, pp. 2243-2257.

Machado. M.D., Santos. M.S., Gouveia. C., Soares. H.M. and Soares. E.V.; 2008. Removal of heavy metals using a brewer's yeast strain of Saccharomyces cerevisiae: The flocculation as a separation process. Bioresour Technol., 99, pp. 2107-2115.

Volesky. B.; 2001. Detoxification of metal-bearing effluents: biosorption for the next century. Hydrometallurgy., 59, pp. 203-216.

Ashok. S., Sudeep. R., Kumar. P.T., Pratibha. R., Manoj. M., Feroz K. and Abha. M.; 2009. Predicted metal binding sites for phytoremediation. Bioinformation., 4, pp. 66-70.

Wang. J. and Chen. C.; 2006. Biosorption of heavy metals by Saccharomyces cerevisiae: Biotechnology Ad., 24, pp. 427-451.

Kawata. K., Yokoo. H., Shimazaki. R. and Okabe. S.; 2007. Classification of heavy-metal toxicity by human DNA microarray analysis. Environ Sci Technol., 41, pp. 3769-3774.

Liu. J., Zhang. Y., Huang. D. and Song. G.; 2005. Cadmium induced MTs synthesis via oxidative stress in yeast Saccharomyces cerevisiae. Mol cell biochem., 280, pp. 139-145.

Mannazzu. I., Guerra. E., Ferretti. R., Pediconi. D. and Fatichenti. F.; 2000. Vanadate and copper induce overlapping oxidative stress responses in the vanadate-tolerant yeast Hansenula polymorpha. BBA General Sub., 1475, pp. 151-156.

Christopher. C. and Peter. G.; 2002. Phytochelatins and metallothioneins: Roles in Heavy Metal Detoxification and Homeostasis. Annu. Rev. Plant Biol., 53, pp.159–82.

Passerini. A., Lippi. M. and Frasconi. P.; 2011. MetalDetector v2. 0: predicting the geometry of metal binding sites from protein sequence. Nucleic acids res., 39, pp. 288-292.

Nakamura. K., Hirai. A., Altaf-Ul-Amin. M. and Takahashi. H.; 2009. MetalMine: a database of functional metal-binding sites in proteins. Plant biotechnol., 26, pp. 517-521.

Sodhi. J.S., Bryson. K., McGuffin. L., Ward. J., Wernisch. L., Jones. D.; 2011. Predicting metal-binding sites from protein sequence. IEEE Trans Comput., 9, pp. 203-213.

Shu. N., Zhou. T. and Hovmöller. S.; 2008. Prediction of zinc-binding sites in proteins from sequence. Bioinformatics., 24, pp. 775-782.

Lippi. M., Passerini. A., Punta. M., Rost. B. and Frasconi. P.; 2008. MetalDetector: a web server for predicting metal-binding sites and disulfide bridges in proteins from sequence. Bioinformatics., 24, pp. 2094-2095.

Brylinski. M. and Skolnick. J. 2011. FINDSITE-metal: Integrating evolutionary information and machine learning for structure-based metal-binding site prediction at the proteome level. Proteins: Struc Fun Bioinf., 79, pp. 735-751.

Ortuso. F., Langer. T. and Alcaro. S.; 2006. GBPM: GRID-based pharmacophore model: concept and application studies to protein–protein recognition. Bioinformatics., 22, pp. 1449-1455.

Gavanji. S., Doostmohammadi. M., Mojiri. A. and Jalali Zand. A.; 2013. Studying the Relationship between Acyl-CoA-Binding Protein 2 and Lysophospholipase 2 in Arabidopsis thaliana and Their Importance in Recycling of Cadmium. IJSRES., 1(12) , pp. 351-356.

Nobuaki. S., Yube. Y., Nozomu. K. and Hiroshi. S.; 2002. Functional characterization of a heavy metal binding protein CdI19 from Arabidopsis. Plant J., 32, pp. 165–173.

Malin. M. and Leif. B.; 2001. Metal-binding proteins and peptides in bioremediation and phytoremediation of heavy metals. Tren in Biotech., 19(2) , pp.67-73.

Gavanji S, Abdul Aziz H, Larki B, Mojiri A.; 2013. Computational Prediction and Analysis of Interaction of Silver Nitrate with Chitinase Enzyme. IJSRES., 1(4) , pp. 50-62.

Gavanji. S., Abdul Aziz H., Larki. B. and Mojiri. A.; 2013. Bioinformatics Prediction of Interaction of Silver Nitrate and Nano Silver on Catalase and Nitrat Reductase. IJSRES., 1(2) , pp. 26-35.

Fulekar. M.H. and Jaya. S.; 2008. Bioinformatics Applied in Bioremediation. Innov Roma Food Biotech., 2(2) , pp.28-36.

Nanjiang. S., Tuping. Z. and Sven. H.; 2008. Prediction of zinc-binding sites in proteins from sequence. Bioinformatics., 24(6) , pp. 775–782.

Degtyarenko. K.; 2000. Bioinorganic motifs: towards functional classification of metalloproteins. Bioinformatics., 16, pp.851-864.

Chance. M.R. and Shi. W.; 2008. Metallomics and metalloproteomics. Cell Mol. Life Sci., 65, pp.3040-3048.

Bertini. I. and Cavallaro. G.; 2010. Bioinformatics in bioinorganic chemistry. Metallomics., 2, pp.39-51.

Passerini. A., Punta. M., Ceroni. A., Rost. B. and Frasconi. P.; 2006. Identifying cysteines and histidines in transition-metal-binding sites using support vector machines and neural networks. Proteins., 65, pp.305-316.

Sodhi. J.S., Bryson. K., McGuffin. L.J., Ward. J.J., Wernisch. L. and Jones. D.T.; 2004. Predicting metal-binding site residues in low-resolution structural models. J. Mol. Biol., 342, pp.307-320.

Shi. W. and Chance. M.R.; 2008. Metallomics and metalloproteomics. Cell Mol Life Sci, 65, pp.3040–3048.

Shi. W., Zhan. C., Ignatov. A., Manjasetty. B.A., Marinkovic. N., Sullivan. M., Huang. R. and Chance. M.R. 2005. Metalloproteomics: High-throughput structural and functional annotation of proteins in structural genomics. Structure., 13, pp. 1473–1486.

Shi. W., Punta. M., Bohon. J., Sauder. J.M., D'Mello. R., Sullivan. M., Toomey. J., Abel. D., Lipp.i M., Passerini. A., Frasconi. P., Burley. S.K., Rost. B. and Chance. M.R.; 2011. Characterization of metalloproteins by high-throughput x-ray absorption spectroscopy. Genome Res., 21(6) , pp.898-907.

Ideker. T., ThorssonV., Ranish. J.A., Christmas. R., Buhler. J., Eng. J.K., Bumgarner. R., Goodlett. D.R., Aebersold. R. and Hood. L.; 2001. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science., 292, pp.929–934.

Parales. R.E., Bruce. N.C., Schmid. A. and Wackett. L.P.; 2002. Biodegradation, biotransformation, and biocatalysis (b3). Appl. Environ. Microbiol., 68, pp.4699–4709.

Pazos. F., Guijas. D., Valencia. A. and De Lorenzo. V.; 2005. MetaRouter: bioinformatics for bioremediation. Nucleic Acids Res., 1(33) , pp. 88-92.

Kanehisa. M., Goto. S., Kawashima. S., Okuno. Y. and Hattori. M.; 2004. The KEGG resource for deciphering the genome. Nucleic Acids Res., 32, pp. 277–280.

Pazos. F., Valencia. A. and De Lorenzo. V.; 2003. The organization of the Microbial Biodegradation Network from a Systems-Biology perspective. EMBO Rep., 4(10) , pp. 994-999.

Messerschmidt. A., Huber. R. and Poulos. T.; 2001. Wieghardt K (eds.), Handbook of Metalloproteins Vol.1 and 2, John Wiley and Sons, Ltd., West Sussex, England.

Jernigan. R., Raghunathan. G. and Bahor. I.; 1994. Characterization of interactions and metal ion binding sites in proteins, Curr. Opin. Struct. Biol., 4, pp.256-263.

Murzin. A.G., Brenner. S.E., Hubbard. T. and Chothia. C.; 1995. SCOP: a structural classification of proteins database for the investigation of sequence and structures. J. Mol. Biol., 247, pp.536-540.

Williams. R.J.P.; 2001. Chemical selection of elements by cells. Coord. Chem. Rev, 217, pp.583-595.

Mehrasa. M., Zaker. S.R, Larki. B., Mehmandoust. M., Baghshahi. H., Sekhavati. M.H. and Gavanji. S.; 2014. Computational Prediction and Analysis of Interaction of Silver Nitrate with Peptidoglycan-Associated lipoprotein (Pal). IJSRES., 2(1) , pp. 8-13.

Frausto da Silva. J.J.R. and Williams. R.J.P.; 2001. The Biological Chemistry of the Elements: The Inorganic Chemistry of Life. Second edition. Oxford University Press, Oxford.

Rhee. K.H., Morris. E.P., Barber. J. and Kuhlbrandt. W.; 1998. Threedimensional structure of the plant photosystem II reaction center at 8 Å resolution. Nature., 396, pp. 283–286.

Andreini. C., Bertini. I., Cavallaro. G., Holliday. G.L. and Thornton. J.M.; 2008. Metal ions in biological catalysis: from enzyme databasesto general principles. J Biol Inorg Chem., 13, pp. 1205–1218.

Ukaegbu. U.E., Henery. S. and Rosenzweig. A.C.; 2006. Biochemical characterization of MmoS, a sensor protein involved in copperdependent regulation of soluble methane monooxygenase. Biochem., 45, pp.10191–10198.

Christianson. D.W.; 1991. Structural biology of zinc. Adv Protein Chem., 42, pp. 281–355.

Andreini. C., Bertini. I., Cavallaro. G., Najmanovich. R.J. and Thornton. J.M.; 2009. Structural analysis of metal sites in proteins: non-heme iron sites as a case study. J Mol Biol., 388, pp. 356–380.

Gasteiger. E., Gattiker. A., Hoogland. C., Ivanyi. I., Appel. R.D. and Bairoch. A.; 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucl Acids Res., 31, pp.3784–3788.

Herraez. A.; 2006. Biomolecules in the computer. Jmol to the rescue. Biochem Mol Biol Educ., 34, pp. 255–261.

Altschul. S.F., Gish. W., Miller. W., Myers. E.W. and Lipman. D.J.; 1990. Basic local alignment search tool. J Mol Biol., 215, pp.403–410.

Goyal. K. and Mande. S.C.; 2008. Exploiting 3D structural templates for detection of metal-binding sites in protein structures. Proteins., 70, pp. 1206–1218.

Babor. M., Gerzon. S., Raveh. B., Sobolev. V. and Edelman. M.; 2008. Prediction of transition metal-binding sites from apo protein structures. Proteins., 370, pp. 208–217.

Punta. M. and Ofran. Y.; 2008. The rough guide to in silico functional prediction, or how to use sequence and structure information to predict protein funciton. PLoS Comput Biol., 4(10) , pp. e1000160.

Bernstein. F.C., Koetzle. T.F., Williams. G.J., Meyer. E.F. J.r., Brice. M.D., Rodgers. J.R., Kennard. O., Shimanouchi. T. and Tasumi. M.; 1977. The Protein Data Bank: a computer-based archival file for macromolecular structures. J. Mol. Biol., 112, pp.535–542.

Altschul. S.F., Madden. T.L., Schäffer. A.A., Zhang. J., Zhang. Z., Miller. W. and Lipman. D.J.; 1997. Gapped BLAST and PSI-BLAST, pp. a new generation of protein database search programs. Nucleic Acids Res., 25: 3389–3402.

Menchetti. S., Passerini. A., Frasconi. P., Andreini. C. and and Rosato. A.; 2006. Improving prediction of zinc binding sites by modeling the linkage between residues close in sequence. Res Comp Mol Biol., 3909, pp.309–320.

Pavlidis. P., Wapinski. I. and Noble. W.S.; 2004. Support vector machine classification on the web.Bioinformatics., 20, pp. 586–587.

Mika. S. and Rost. B.; 2003. UniqueProt: creating representative protein sequence sets. Nucleic Acids Res., 31, pp. 3789–3791.

Frasconi. P. and Passerini. A.; 2009. Predicting the Geometry of Metal Binding Sites from Protein Sequence. NIPS., 21, pp. 465-472.

Passerini. A., Lippi. M. and Frasconi. P.; 2011. Predicting Metal-Binding Sites from Protein Sequence. IEEE/ACM., 9 (1) , pp. 203–213.

Ceroni. A., Passerini. A., Vullo. A. and Frasconi. P.; 2006. DISULFIND: a Disulfide Bonding State and Cysteine Connectivity Prediction Server. Nucleic Acids Res., 34, pp. 177-181.

Byrne. R., Amphlett. G.W. and Castellino. F.J.; 1980. The binding of metal ions to bovine factor IX. J Biol Chem., 255:1430–1435.

Shore. J.D., Day. D.E., Bock. P.E. and Olson. S.T.; 1987 . Acceleration of surface-dependent autocatalytic activation of blood coagulation factor XII by divalent metal ions. Biochemistry., 26(8), pp.2250-8.

Tubek. S., Grzanka. P. and Tubek. I.; 2008. Role of zinc in hemostasis: a review. Biol Trace Elem Res., 121(1) , pp.1-8.

Vallee. B.L. and Auld. D.S.; 1990. Zinc coordination, function, and structure of zinc enzymes and other proteins. Biochemistry., 29(24) , pp. 5647-59.

Sen. P., Sahoo. S., Pendurthi. U.R. and Rao. L.V.; 2010. Zinc modulates the interaction of protein C and activated protein C with endothelial cell protein C receptor. J Biol Chem., 285(26) , pp.20410-20420.

Kristen. L., Huber. and Jeanne. A.H.; 2012. Mechanism of zinc-mediated inhibition of caspase-9. Protein Sci., 21(7) , pp. 1056–1065.

Maurer-Stroh. S., Gao. H., Han. H., Baeten. L., Schymkowitz. J., Rousseau. F., Zhang. L. and Eisenhaber. F.; 2013. Motif discovery with data mining in 3D protein structure databases: discovery, validation and prediction of the U-shape zinc binding ("Huf-Zinc") motif. J Bioinform Comput Biol., 11(1) , pp.1340008.

Saedi. Z, Gavanji. S. and Davodi. S.; 2013. A Review of Genetic and Epigenetic Mechanisms in Heavy Metal Carcinogenesis: Nickel and Cadmium. IJSRES., 1 (8) , pp. 202-216.

Buchan. D.W.A., Minneci. F., Nugent. T.C.O., Bryson. K. and Jones. D.T.; 2013. Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Res., 41 (1) , pp. 340-348.

McGuffin. L.J., Bryson. K. and Jones. D.T.; 2000. The PSIPRED protein structure prediction server. Bioinformatics., 16, pp.404-405.

Jones. D.T. and Swindells. M.B.; 2002. Getting the most from PSI-BLAST. Trends Biochem Sci., 27, pp.161-164.

Citiulo. F., Jacobsen. I.D., Miramón. P., Schild. L., Brunke. S., Zipfel. P., Brock. M, Hube. B. and Wilson. D.; 2012. Candida albicans scavenges host zinc via Pra1 during endothelial invasion. PLoS Pathog., 8(6) , pp. e1002777.

Bakir. G., Hofmann. T., Schölkopf. B., Smola. A, Taskar. B. and Vishwanathan. S.V.N.; 2007. Predicting Structured Data. Cambridge, MA: The MIT Press.

Von Itzstein. M., Wu. W.Y., Kok. G.B., Pegg. M.S., Dyason. J.C., Jin. B., Van Phan. T., Smythe. M.L., White. H.F. and Oliver. S.W.; 1993. Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature., 363(6428), pp.418-23.

Kastenholz. M.A., Pastor. M., Cruciani. G., Haaksma. E.E. and Fox. T.; 2000. GRID/CPCA: a new computational tool to design selective ligands. J Med Chem., 43(16) , pp.3033-3044.

Pastor. M., Cruciani. G. and Watson. K.A.; 1997. A strategy for the incorporation of water molecules present in a ligand binding site into a three-dimensional quantitative structure--activity relationship analysis. J Med Chem., 40(25) , pp. 4089-102.

Wade. R.C. and Goodford PJ.; 1993. Further development of hydrogen bond functions for use in determining energetically favorable binding sites on molecules of known structure. 2. Ligand

probe groups with the ability to form more than two hydrogen bonds. J Med Chem., 36(1) , pp. 148-56.

Carosati. E., Sciabola. S. and Cruciani G.; 2004. Hydrogen bonding interactions of covalently bonded fluorine atoms: from crystallographic data to a new angular function in the GRID force field. J Med Chem.,47(21) , pp. 5114-5125.

Vullo. A. and Frasconi P.; 2004. Disulfide connectivity prediction using recursive neural networks and evolutionary information. Bioinformatics., 20(5) , pp. 653-659.

Frasconi. P., Passerini. A. and Vullo, A.; 2002. A two-stage SVM architecture for predicting the disulfide bonding state of cysteines. Proc. IEEE., 3, pp. 25-34.

Ceroni. A., Frasconi. P., Passerini. A. and Vullo. A.; 2003. Predicting the disulfide bonding state of cysteines with combinations of kernel machines. J VLSI Signal Pro., 35, pp. 287-295.