Computational Prediction of Metal Binding Sites in Lysyl aminopeptidase in *Pyrococcus furiosus* (strain ATCC 43587)

Mohsen Doostmohammadi¹, Shahin Gavanji²*, Behrouz Larki³, Amin Mojiri⁴

¹Department of Biotechnology, Faculty of Advanced Sciences and Technologies, University of Isfahan, Isfahan, Iran
²Young Researcher Club, Khorasgan (Isfahan) Branch, Islamic Azad University, Isfahan, Iran
³Department of Plant protection, Khorasgan (Isfahan) Branch, Islamic Azad University, Isfahan, Iran
⁴School of Civil Engineering, Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia

*Corresponding Author: shahin.gavanji@tahoo.com

Received 13 September 2013; Accepted 18 October 2013

**Abstract.** More than 70,000 protein structures are currently found in the Protein Data Bank, and approximately one-third contain metal ions essential for function. Identifying and characterizing metal ion–binding sites experimentally is time-consuming and costly. Recently, the three-dimensional structure of two aminopeptidases, the methionine aminopeptidase from *Escherichia coli* and the leucine aminopeptidase from *Aeromonas proteolytica*, have been elucidated by crystallographic studies. Aminopeptidases play a role in several important physiological processes. It is noteworthy that some of them take part in the catabolism of exogenously supplied peptides and are necessary for the final steps of protein turnover. The object of the present study is to characterize the metal binding sites of lysine aminopeptidase (KAP) from the hyperthermophilic archaeon, *P. furiosus*. It is shown that the *P. furiosus* enzyme, while having the same narrow substrate specificity as the KAPs from *S. cerevisiae* and *A. niger*, contains conserved sequences that are homologous to those of members of the M18 rather than M1 family of peptidases. In contrast to the large M1 family, only two members of the M18 family have been characterized. These are a leucyl aminopeptidase from yeast and an aspartyl aminopeptidase from rabbit. Herein, we therefore report characterization of the first KAP from an archaeon, in the form of the *P. furiosus* enzyme, which is also the first prokaryotic member of the M18 family.

**Key words:** Amino peptidase, Prediction, Metal, Bioinformatics, *Pyrococcus furiosus*

1. **INTRODUCTION**

Many metal ions are essential as trace elements for plants growth, but at higher concentrations they become toxic. Heavy metals are difficult to remove from the environment and unlike many other pollutants cannot be chemically or biologically degraded and are ultimately indestructible. Today, many heavy metals constitute a global environmental hazard. For example, environmental pollution by Cd, arising mainly from mining and smelting, dispersal of sewage sludge and the use of phosphate fertilizers, is increasing. Thus, the use of microorganisms and plants for the decontamination of heavy metals has attracted growing attention because of several problems associated with pollutant removal using conventional methods (Mejáre and Bülow, 2001). Throughout evolution, properties of metals have been harnessed by proteins for performing functions such as redox reactions which cannot be performed by using functional groups found amino acids (Messerschmidt, 2001).

These metalloproteins have many different functions in cells such as, enzymatic functions, transport, storage, and signal transduction. It is now known that the new zinc and copper-binding proteins are only found in Eukaryotes, not in the Bacteria and Archaea. The nucleus houses most of the new zinc binding proteins and this unique utilization of zinc is one of the defining features of all Eukaryotes. A possible hypothesis is that zinc concentrations in the ancient ocean were too low to allow for the evolution of the Eukaryotes, at least until global changes in oxygen occurred (Dupont et al., 2010). Since evolution of metalloproteins has happened multiple times under different selection pressures and for different metals, it can be hypothesized that any overarching signature common to all metalloproteins will be very weakly predictable. Finally it can be safely argued that each metalloprotein shares a common ancestor with a non-metalbinding protein and that an alignment free distance metric such as Euclidean distance between kmer frequency vectors (Edgar, 2004) could approximate the actual evolutionary distance between a metalloprotein and its nearest non-metal-binding relative.

This fact will be useful for creating the training and test sets. *Pyrococcus furiosus* is noted for its rapid doubling time of 37 minutes under optimal conditions, meaning that every 37 minutes, the number of
Such a system could be composed of amino acids that are used to predict metal binding sites. For example, the protein Lysyl aminopeptidase in Pyrococcus furiosus is 38,214 Da in weight. In this case, H₂S can be produced through its metabolic processes, although no energy seems to be derived from this series of reactions. Interestingly to note is that, while many other hyperthermophiles depend on sulfur for growth, P. furiosus does not. Pyrococcus furiosus is also notable for an unusual and intriguingly simple respiratory system, which obtains energy by reducing protons to hydrogen gas and uses this energy to create an electrochemical gradient across its cell membrane, thereby driving ATP synthesis. Such a system could be a very early evolutionary precursor of respiratory systems in all higher organisms (Fiala and Stetter, 1986). The bacterial Lysyl aminopeptidase is a broad-specificity metallo aminopeptidase. Peptidase N (PepN) is the most commonly used name for the enzyme belonging the family M1, which includes bacterial, fungal and mammalian enzymes. The name originally used to describe aminopeptidase hydrolizing aminoacyl β-nanhydramides. The databanks currently assign the sequence of bacterial Lysyl aminopeptidase to EC 3.4.11.2, for which recommended name membrane alanyl aminopeptidase (Poolman et al., 1995; Niven et al., 1995).

The goal of this study was computational prediction of metal binding sites in Lysyl aminopeptidase in Pyrococcus furiosus (strain ATCC 43587) using the 'CHED' algorithm to predict 3D metal binding sites using with metal binding sites using 'CHED' algorithm. This site uses the 'CHED' algorithm to predict 3D intra-chain protein binding sites for transition metals (Zn, Fe, Mn, Cu, Ni, Co, and Ca and Mg sites that can be replaced by a transition metal). The algorithm searches for a triad of amino acids composed of 4 residue types (Cys, His, Glu, Asp) having ligand atoms within specific distances. It allows one target residue to rotate in rotamer space, taking into account structural rearrangements that may occur upon metal binding. A binding site is considered to be true if one or more correct amino acid ligands have been predicted. Machine learning algorithms are used to filter out false positives. MILD FILTER - based on the frequency of hits, yields high sensitivity.

2. MATERIALS AND METHODS

2.1. Preparing 3 dimensional structures of Chitinase Lysyl aminopeptidase

In the first step, amino acid sequences of Lysyl aminopeptidase with accession number Q8TZW4 was taken from NCBI website (www.ncbi.nlm.nih.gov) (Figure 1). Then the Lysyl aminopeptidase enzyme with the number of 2PE3 was taken from Protein Data Bank website (www.rcsb.com) (Figure 2).

2.2. Studying metal binding site

Studies on metal binding site were done using Metal Detector Predicts v2.0 software (http://metaldetector.dsi.unifi.it). Zinc-binding site prediction is by PredZinc version 1.4, 2011-10-14 (c) Shu (http://casio.fos.su.se), UCL-CS Bioinformatics Web Servers (http://bioinf.cs.ucl.ac.uk/structure) and CHED - predicting soft metal binding sites (http://ligin.weizmann.ac.il) (Figure 3).

3. RESULTS AND DISCUSSIONS

3.1. Protein Structure Analysis

The Lysyl aminopeptidase consists of 346 amino acids, and its molecular weight is 38,214 Da. In this structure, there are Zinc binding sites at amino acid positions: 207, 63, 177, 208, 230 and 314 (http://www.uniprot.org).

3.2 Metal binding site

3.2.1 Predicting 3D interaction of Lysyl aminopeptidase with metal binding sites using 'CHED' algorithm

This site uses the 'CHED' algorithm to predict 3D intra-chain protein binding sites for transition metals (Zn, Fe, Mn, Cu, Ni, Co, and Ca and Mg sites that can be replaced by a transition metal). The algorithm searches for a triad of amino acids composed of 4 residue types (Cys, His, Glu, Asp) having ligand atoms within specific distances. It allows one target residue to rotate in rotamer space, taking into account structural rearrangements that may occur upon metal binding. A binding site is considered to be true if one or more correct amino acid ligands have been predicted. Machine learning algorithms are used to filter out false positives. MILD FILTER - based on the frequency of hits, yields high sensitivity.
(depending on metal type, 70-100% of apo sites are captured, with predictions statistically accurate in 50-70% of cases). STRINGENT FILTER - using decision tree and support vector machine technology, yields high selectivity (depending on metal type, 45-85% of apo sites are captured with predictions statistically accurate in 85-100% of cases). Lysyl aminopeptidase 3D structure analysis showed that 3 positions in this protein interact with metals which are ordinarily: 11 amino acid include HIS 71, ASP 73, ASP 185, ASP 186, CYC 189, GLU 215, GLU 216, ASP 238, ASP 298, HIS 322 and GLU 326 at position 1, 3 amino acids include: ASP45, GLU 46 and HIS 58 at position 2 and finally amino acids HIS 89, ASP 83 and HIS 81 at position 3 (Figure 4).

3.2.1 Predicting 3D interactions of Lysyl aminopeptidase with metal binding sites using UCL-CS Bioinformatics Web Server (http://bioinf.cs.ucl.ac.uk/structure)

The interaction of Lysyl amino peptidase with metals studied using UCL-CS Bioinformatics Web Server and the results showed that 11 amino acids are included in binding to metals (Figure 5). According to the results GLN 234, GLY 248 and ASP 40 had the highest scores (Table 1).

Table 1: Metsite Prediction scores by using UCL-CS Bioinformatics Web Server

<table>
<thead>
<tr>
<th>No</th>
<th>Residue</th>
<th>Raw neural network score</th>
<th>Residue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14</td>
<td>0.644152</td>
<td>ASP40</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>0.584224</td>
<td>GLN94</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>0.509245</td>
<td>GLY107</td>
</tr>
<tr>
<td>4</td>
<td>36</td>
<td>0.706971</td>
<td>GLY248</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>0.545927</td>
<td>ASP135</td>
</tr>
<tr>
<td>6</td>
<td>41</td>
<td>0.564214</td>
<td>GLY137</td>
</tr>
<tr>
<td>7</td>
<td>69</td>
<td>0.574616</td>
<td>GLU205</td>
</tr>
<tr>
<td>8</td>
<td>79</td>
<td>0.904547</td>
<td>GLN234</td>
</tr>
<tr>
<td>9</td>
<td>84</td>
<td>0.699156</td>
<td>GLY248</td>
</tr>
<tr>
<td>10</td>
<td>88</td>
<td>0.586593</td>
<td>ASN258</td>
</tr>
<tr>
<td>11</td>
<td>100</td>
<td>0.488412</td>
<td>GLN293</td>
</tr>
</tbody>
</table>


3.2.2 Zinc-binding site prediction by PredZinc:

Analyzing the sequence of Lysyl aminopeptidase with PredZinc server showed that in four amino acids, Zinc interact with Lysyl aminopeptidase. The probability score of the interactions was more than 0.450 (Table 2).
Table 2: Results of binding Zinc to Lysyl aminopeptidase

<table>
<thead>
<tr>
<th>No</th>
<th>AA</th>
<th>Sequence Index</th>
<th>Zn Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HIS</td>
<td>63</td>
<td>0.933</td>
</tr>
<tr>
<td>2</td>
<td>HIS</td>
<td>314</td>
<td>0.925</td>
</tr>
<tr>
<td>3</td>
<td>ASP</td>
<td>65</td>
<td>0.633</td>
</tr>
<tr>
<td>4</td>
<td>HIS</td>
<td>117</td>
<td>0.522</td>
</tr>
</tbody>
</table>

Predicted 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

4. CONCLUSION

The Bioinformatic checking shows that Lysyl aminopeptidase is capable to absorb heavy metals, so the predictions using UCL-CS Bioinformatics Web Servers showed that 11 amino acids existed in this protein are related to Glutamine number 293. Also results obtained by using PredZinc revealed that, 4 amino acids have metal absorbance property. Generally, obtained results discovered that the protein has several sites for heavy metal absorbance. Hence, Bioinformatics is involved in prediction of the bounds, exact experimentation is necessary in laboratory conditions.
Fig. 4: A: Prediction of interaction between metal and Lysyl aminopeptidase in three positions B: Interaction between metal and Lysyl aminopeptidase in position one (11aa) C: Interaction between metal and Lysyl aminopeptidase in position two (3aa) D: Interaction between metal and Lysyl aminopeptidase in position three (3aa)

Fig. 5: Prediction of interaction between metal and Lysyl aminopeptidase using UCL-CS Bioinformatics Web Server
REFERENCES


Fiala G, Stetter KO (1986). Pyrococcus furiosus sp. nov. represents a novel genus of marine heterotrophic archaeabacteria growing optimally at 100°C. Archives of Microbiology, 145: 56–61


Mohsen Doostmohammadi graduated in Biotechnology at MSc at the Department of Biotechnology, Faculty of Advanced Sciences and Technologies, University of Isfahan, Isfahan, Iran.

Shahin Gavanji graduated in Biotechnology at MSc at the Department of Biotechnology, Faculty of Advanced Sciences and Technologies, University of Isfahan, Isfahan, Iran. He has over 10 international medals in invention. Shahin Gavanji’s research has focused on Pharmacy and Pharmacology, Nano Biotechnology, Bioinformatics, Biotechnology - Medical Biotechnology. He is editor in chief of International Journal of Scientific Research in Inventions and New Ideas.

Behrouz Larki is MSc student at Isfahan (Khorasgan) Branch, Islamic Azad University. Behrouz Larki’s research has focused on Biology, Pharmacy, Linguistic and Translation. He is managing editor of International Journal of Scientific Research in Inventions and New Ideas.

Amin Mojiri is a PhD candidate in Environmental Engineering, School of Civil Engineering, Universiti Sains Malaysia (USM), Pulau Pinang. He is fellowship holder and research assistant at the School of Civil Engineering (USM). He has more than 18 articles in the international journals. He is editor and reviewer of some international journals. His area of specialization is waste management, waste recycling, wastewater treatment, wastewater recycling, and soil pollutions.