Structural context around amino acid residues binding Mn^{+2} ions in bacterial proteins

T.A. Khrustaleva

Institute of Physiology of the National Academy of Sciences of Belarus, Minsk, Academicheskaya, 28, tanissia.lir@gmail.com

V.V. Khrustalev, E.V. Barkovsky

Belarussian State Medical University, Belarus, Minsk, Dzerzinskogo, 83, vvkhrustalev@mail.ru, barkovsky@hotmail.ru

There are three amino acid residues which are most frequently involved in coordination of manganese (II) ions: aspartic acid (Asp), histidine (His) and glutamic acid (Glu) [1]. It is known that those amino acids are distributed unequally among secondary structure elements. In general, Asp and His are overrepresented in unstructured regions of proteins (in random coil), while Glu is overrepresented in alpha helices [2]. On the other hand, Asp is overrepresented in first positions of alpha helices where it is thought to be involved in helix capping [2]. Moreover, Asp and Glu are responsible of the dipole moment of alpha helices formation: N-terminal part of helix is usually negatively charged, while C-terminal one is positively charged [2].

The aim of this work was to find out what secondary structure elements are often involved in Mn^{+2} ions binding and whether that preference can be explained by the unequal distribution of amino acids between secondary structure elements and within them.

As the material we used 149 PDB files from the Protein Data Bank containing description of bacterial proteins with Mn^{+2} ions. There were 39 proteins from GC-poor bacteria, 62 proteins from bacteria with average genomic GC-content and 48 proteins from GC-rich bacteria. We collected information on amino acid content of alpha helices, beta strands and 3/10 helices for those proteins. Moreover, unstructured regions have been classified in the respect of the secondary structure elements situated before and after them. We did not use C-terminal and N-terminal regions of proteins represented by random coil and regions with amino acid residues which were not located in crystallographic experiment.

The percent of amino acid residues in alpha helices is equal to 45.01% in our data set (total number of amino acid residues is equal to 44136), while the percent of amino acid residues situated in alpha helices among those binding Mn^{+2} ions (their total number is equal to 588) is significantly lower (27.04%). In contrast, the percent of amino acid residues situated in beta strands is significantly higher among those involved in Mn^{+2} coordination (31.46%) than in the complete data set (18.91%).

Amino acids involved in Mn^{+2} binding are overrepresented in regions of coil situated between beta strand and alpha helix (17.01% vs. 7.47%, $P<0.001$) and underrepresented in regions of coil between alpha helix and beta strand (0.51% vs. 6.28%, $P<0.001$). Mn^{+2} binding sites are overrepresented in regions of coil between beta strand and 3/10 helix (4.08% vs. 1.96%, $P<0.01$) and underrepresented in regions of coil between beta strand and 3/10 helix (0.34% vs. 1.55%, $P<0.001$). Mn^{+2} binding residues are also overrepresented in regions of coil between two beta strands (12.59% vs. 7.39%, $P<0.001$) and underrepresented in regions of coil between two alpha helices (1.36% vs. 3.90%, $P<0.001$). The clear preference to be situated after the beta strand characteristic for Mn^{+2} binding residues may be explained, at least partially, by the fact that Asp has the highest frequency of usage (6.74%) in the last position of beta strands (average usage of Asp in beta strands is equal to 3.47%). Indeed, amino acid residue coordinating Mn^{+2} ion more frequently can be found in the C-terminal border of beta strand than inside it. Moreover, the usage of lysine (Lys) which should disturb interactions with positively charged ions is significantly lower in regions of coil between beta strand and alpha helix than in regions between alpha helix and beta strand (4.30% vs. 6.50%, $P<0.001$). The percent of binding sites situated between beta strand and alpha helix is significantly higher than the percent of those sites between two beta strands, probably, also because of the significantly lower Lys usage (4.30% vs. 5.85%, $P<0.01$).

1. M. Brylinski, J. Skolnick (2011) FINDSITE-metal: Integrating evolutionary information and machine learning for structure-based metal binding site prediction at the proteome level, *Proteins,* **79:** 735–751.

2. R. Aurora, G.D. Rose (1998) Helix capping, *Protein Science,* **7:**21–38.