MS100.O06

Check your metal - not every density blob is a water molecule

H. Zheng[1,5], M. Chordia[1,5], D. Cooper[1,5], I. Shabalin[1,5], M. Chruszcz[1,2,5], P. Müller[3], G. Sheldrick[4], W. Minor[1,5]

[1]University of Virginia, Department of Molecular Physiology and Biological Physics, Charlottesville, Virginia, USA, [2]University of South Carolina, Department of Chemistry and Biochemistry, Columbia, South Carolina, USA, [3]Massachusetts Institute of Technology, Department of Chemistry, Cambridge, Massachusetts, USA, [4]Universität Göttingen, Lehrstuhl für Strukturchemie, Göttingen, Germany, [5]Center for Structural Genomics of Infectious Diseases (CSGID) Consortium, USA

Metals play vital roles in both the mechanism and architecture of biological macromolecules, and are the most frequently encountered ligands (i.e. non-solvent heterogeneous chemical atoms) in the determination of macromolecular crystal structures. However, metal coordinating environments in protein structures are not always easy to check in routine validation procedures, resulting in an abundance of misidentified and/or suboptimally modeled metal ions in the Protein Data Bank (PDB). We present a solution to identify these problems in three distinct yet related aspects: (1) coordination chemistry; (2) agreement of experimental B-factors and occupancy; and (3) the composition and motif of the metal binding environment. Due to additional strain introduced by macromolecular backbones, the patterns of coordination of metal binding sites in metal-containing macromolecules are more complex and diverse than those found in inorganic or organometallic chemistry. These complications make a comprehensive library of "permitted" coordination chemistry in protein structures less feasible, and the usage of global parameters such as the bond valence method more practical, in the determination and validation of metal binding environments. Although they are relatively infrequent, there are also cases where the experimental B-factor or occupancy of a metal ion suggests careful examination. We have developed a web-based tool called CheckMyMetal [1](http://csgid.org/csgid/metal_sites/) for the quick validation of metal binding sites. Moreover, the acquired knowledge of the composition and spatial arrangement (motif) of the coordinating atoms around the metal ion may also help in the modeling of metal binding sites in macromolecular structures. All of the studies described herein were performed using the NEIGHBORHOOD SQL database [2], which connects information about all modeled non-solvent heterogeneous chemical motifs in PDB structure by vectors describing all contacts to neighboring residues and atoms. NEIGHBORHOOD has broad applications for the validation and data mining of ligand binding environments in the PDB.

[1] *H. Zheng, M. Chordia, D. Cooper et al, Nat Protoc., 2014, 9(1), 156-70,* ***[2]*** *H. Zheng, M. Chruszcz, P. Lasota et al, J Inorg Biochem., 2008, 102(9), 1765-76*

Keywords: CheckMyMetal (CMM), metal-containing proteins, structure validation