MS.92.5

Acta Cryst. (2008). A**64**, C155

Quality checks for carbohydrate structures in PDB entries

Thomas Lutteke[1], Claus W von der Lieth[2]
[1]Radboud University Nijmegen, CMBI, Geert Grooteplein 26-28, Nijmegen, Gelderland, 6525 GA, The Netherlands, [2]German Cancer Research Centre, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany, E-mail：t.luetteke@uu.nl

The quality of the carbohydrates in PDB entries is rather poor compared with the protein parts [1], with about 30% of the PDB entries with carbohydrates being erroneous [2]. The main reasons for this are the complexity of carbohydrates and the lack of check programs for carbohydrate 3D structures. Recently, such check tools were established. Pdb-care (PDB CArbohydrate REsidue check) [www.glycosciences.de/tools/pdb-care/] checks if the carbohydrate residue names used in a PDB file match the monosaccharide units present in the structure [3]. Furthermore, the connectivities given in a PDB file are checked. These bond checks can be applied not only to carbohydrates but to any residue. To evaluate the conformation of a carbohydrate chain, plots of the phi / psi angles of the glycosidic linkages can be used, similar to the Ramachandran Plot for proteins [4]. However, the preferred glycan torsions depend on the involved residues and the linkage position. Thus, separate plots have to be generated for each type of disaccharide fragment in the structure. This can be done with carp [www.glycosciences.de/tools/carp/]. Detected torsions can be compared with either all torsions of the respective type that are present in the PDB or with computed energy maps taken from GlycoMapsDB [www.glycosciences.de/modeling/glycomapsdb/] [5]. Use of these tools will help to increase the reliability of carbohydrate 3D structures.

[1] Crispin M, Stuart DI, Jones EY (2007) Nat Struct Mol Biol. 14, 354
[2] Lutteke T, Frank M, von der Lieth CW (2004) Carbohydr Res. 339, 1015-1020
[3] Lutteke T, von der Lieth CW. (2004) BMC Bioinformatics 5, 69
[4] Hooft RW, Sander C, Vriend G (1997) Comput Appl Biosci. 1997, 13, :425-430
[5] Frank M, Lutteke T, von der Lieth CW (2007) Nucleic Acids Res. 35, :287-290

Keywords: structure validation, carbohydrates, glycoproteins

MS.92.6

Acta Cryst. (2008). A**64**, C155

Data quality in the PDB archive

John Westbrook[1], Kim Henrick[2], Zukang Feng[1], Dimitris Dimitropoulos[2], Shuchismita Dutta[1], Catherine L Lawson[1], John L Markley[3], Haruki Nakamura[4], Jawahar Swaminathan[2], Eldon L Ulrich[3], Jasmine Young[1], Helen M Berman[1]
[1]Rutgers, The State University of New Jersey, Chemistry and Chemical Biology, 610 Taylor Road, Piscataway, nj, NJ, 08854, USA, [2]European Bioinformatics Institute, EMBL Outstation-Hinxton, Cambridge CB10 1SD, United Kingdom, [3]University of Wisconsin-Madison, Department of Biochemistry, 433 Babcock Drive, Madison, WI 53706, USA, [4]Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan, E-mail：jwest@rcsb.rutgers.edu

As part of the deposition and annotation process, a variety of geometrical and experimental validation diagnostics are performed by the wwPDB (www.wwpdb.org). These diagnostics are provided to assist PDB users in the preparation and deposition of new entries, and internally by the wwPDB as entries are processed and annotated. As part of an ongoing effort to improve the consistency and usability of entries within the PDB archive, the wwPDB released a set of remediated data files in 2007. Remediation efforts have continued, and a second set of remediated will also be released. The lessons learned in this remediation effort, the scope of these corrections, and additional content in these entries will be described. The wwPDB sponsored the Workshop on Next Generation Validation Tools for the wwPDB at EBI-Hinxton, April 14-16, 2008. The workshop was designed to collect recommendations and develop consensus on additional validation that should be performed on PDB entries, and to identify software applications to perform validation tasks. The recommendations of this workshop will form the basis for the next generation validation procedures.

Keywords: protein structure database, structure validation, data validation

MS.93.1

Acta Cryst. (2008). A**64**, C155

The annotations of non-Watson-Crick base pairs and comparisons between RNA structures and sequences

Eric Westhof
Université Louis Pasteur, Architecture et réactivité de l'ARN, IBMC, 15 rue René Descartes, Strasbourg, France, 67084, France, E-mail：e.westhof@ibmc.u-strasbg.fr

RNA architecture is now visualized as the hierarchical assembly of preformed double-stranded helices defined by Watson-Crick base pairs and RNA modules maintained by non-Watson-Crick base pairs. All base-base pairwise interactions present in nucleic acids have been classified in twelve families where each family is a 4x4 matrix of the bases A, G, C, U (1). This classification allows to deduce the isotericity matrices which yield all the possible and geometrically equivalent base pairs in a given family. These isostericity matrices have been verified for several RNA motifs using structural alignments anchored by crystallographic structures. Various recurrent motifs have been analyzed along such lines. Three-way junctions observed in crystal structures with two helices approximately co-axially stacked can be divided into three main families depending on the relative lengths of the segments linking the three Watson-Crick helices (2). Each family has topological characteristics with some conservation in the non-Watson-Crick pairs within the linking segments as well as in the types of contacts between the segments and the helices. Computer programs have been developed for automatic annotations and manipulations of RNA structures and sequence alignments (3).

(1) Leontis, NB. Stombaugh, J. and Westhof, E. 2002. The non-Watson-Crick base pairs and their associated isostericity matrices. Nucleic Acids Res 30:3497-3531.
(2) Lescoute, A. and Westhof, E. (2006) Topology of three-way junctions in folded RNAs. RNA 12, 83-93.
(3) Jossinet, F. and Westhof, E. (2005). Sequence to Structure (S2S): display, manipulate and interconnect RNA data from sequence to structure. Bioinformatics, 1, 3320-3321.

Keywords: nucleic acid topology, nucleic acid crystallography, RNA