MS.98.1

Acta Cryst. (2008). A**64**, C163

Quality of protein crystal structures in the protein data bank

Ramaswamy Subramanian, Eric N Brown
University of Iowa, Biochemistry, 4-612, Bowen Science Building, 51 Newton Road,, Iowa City, Iowa, 52242, USA, E-mail：s-ramaswamy@uiowa.edu

One of the challenges of the genomic era is the validation and quality assessment of data deposited in databases. Unfortunately, users of databases are not fully aware of the limitations of the experimental techniques used to generate these data sets or the subjective interpretation of the data by the depositor. In the case of models of protein structures deposited in the PDB (RCSB), there is significant validation efforts carried out. However, all current validation tools compare the deposited model to an ideal scenario (both in terms of chemical knowledge or agreement to experimental data). An ideal structural model hence is one that will have ideal geometries and perfect agreement with other known chemical knowledge and to experimental data. Since, most models are a result of least squares minimization (in some form or another), and the information of this 'ideal' is used in minimization; one can argue that validation is simply checking how wellthe minimization process worked. We decided to take a different approach. Using a combination of standard statistical techniques and some assumptions we argue that the average structure in the PDB is average. We then proceeded to determine possible attributes that contribute to deviations from the average. This analysis resulted in several interesting findings. In the presentation, I will elaborate on the advantages of this methodology in its ability to point us towards experimental methods to improve the average (statistical oxymoron?), further work that needs to be done to fully exploit the method (including redundancy and R-sym in the equations for quality), and what additional information should be deposited to make the data more useful to the scientific community.

Keywords: structural accuracy, protein database, erroneous structures

MS.98.2

Acta Cryst. (2008). A**64**, C163

Discovering the world's best organic non-linear optical materials

Jacqueline M Cole, Daniel D. Hickstein, Ze F. Weng
University of Cambridge, Physics, Cavendish Laboratory, J. J. Thomson Avenue, Cambridge, Cambridgeshire, CB3 0HE, UK, E-mail：jmc61@cam.ac.uk

This paper presents two generic approaches that predict the world's best organic non-linear optical (NLO) compounds. The results arise from a data-mining project that drew its source data from the Cambridge Structural Database, the world's resource of all published organic crystal structures. We undertook two complementary systematic data-mining strategies to realise these predictions: (a) construction of a new data-mining tool, coded with mathematical algorithms derived from the area of decision mathematics. This enabled one to systematically search the entire database for certain fragment types or other aspects of a molecule that are known to be of key importance in existing industry-tested organic NLO materials (certain electron donor-acceptor pairs, conjugation lengths, optimum conjugation pathways, etc); (b) semi-empirical MOPAC calculations that use the molecular geometry of each crystal structure in this database, duly optimised, to predict the value of μ and β, a measure of the NLO effect. There is a very high correlation between these two very separate approaches, which provides us already with good confidence that these predictions are valid. Furthermore, the fact that we have found that several of these predicted top 100 molecules have already been patented as NLO materials demonstrates that several of these molecules are certainly very useful for the NLO industry. Complementary to the above studies, we have also conducted crystallography experiments on several NLO materials: both conventional crystal structure analysis and high resolution charge-density studies. In the latter respect, we calculate the optical properties (μ, α, β, γ) using X-ray constrained wavefunction fitting and compare these to values obtained using traditional multipolar refinement.

Keywords: non-linear optics, charge-density, data-mining

MS.98.3

Acta Cryst. (2008). A**64**, C163

The crystal structures of para-acetanilides analysed systematically

Susanne Huth[1], Mike B. Hursthouse[1], Chris S. Frampton[2]
[1]University of Southampton, Chemistry, Highfield Campus, Southampton, Hampshire, SO17 1BJ, UK, [2]SAFC-Pharmorphix Limited, 250 Cambridge Science Park, Milton Road, Cambridge, CB4 0WE, UK., E-mail：S.Huth@soton.ac.uk

Over the last two decades small molecule X-ray crystallography has become an important tool for the detailed investigation of the solid state with the ultimate aim to understand the (supra)molecular assemblies in crystal structures. The systematic study of crystal packing patterns together with the application of solid state energy calculations can provide an improved insight into the solid state assembly providing feedback for design and prediction procedures. Libraries of closely related compounds comprise suitable systems for such an analysis since the core structure and its interactions remain the same so that variations in the crystal packing can be directly associated with substituent effects. Para-Substituted acetanilides comprise a pharmaceutically important group of molecules; in particular the compound p-hydroxy-acetanilide (aka paracetamol) has attracted exceptional interest. A thorough understanding of substituent effects on crystal packing is hence highly desirable with respect to the utilisability of these compounds. The program XPac [1] has proved to be an excellent tool for the search of common structural patterns in solid state assemblies and was thus used for the systematic cross-comparison of the crystal structures of a series of para-substituted acetanilides. Supplementing this information a variety of lattice energy calculations were performed comparing ab initio (CRYSTAL06) with semi-classical density sum (OPiX) methods. The results of this systematic study are presented and discussed.

[1] T. Gelbrich, M. B. Hursthouse, Cryst. Eng. Comm., 2005, 7(53), 324-336.

Keywords: structural systematics, lattice energy calculations, para-substituted acetanilides

MS.98.4

Acta Cryst. (2008). A**64**, C163-164

An investigation into deuteration effects: Implications for protein crystallography

Stuart J Fisher[1,2], John R Helliwell[2]
[1]Institut Laue-Langevin, BP 156, 6, rue Jules Horowitz, Grenoble, Cedex

9, 38042, France, [2]Department of Chemistry, The University of Manchester, Brunswick Street, Manchester, M13 9PL, UK, E-mail：fisher@ill.fr

Per-deuteration of proteins is becoming more common place and the assumption is in general that deuteration does not affect protein structure. It should be noted that functional changes upon deuteration are known, e.g. D_2O is toxic to living systems, reaction kinetics change [1], proteins stiffen in D_2O [2] and ferroelectrics alter their properties [3]. There are two exceptions known to us for protein structures. Firstly in Kuhn et al [4] they observe a difference in position of a critical H versus D atom for subtilisin versus trypsin respectively. Secondly in haloalkane dehydrogenase Liu et al 2007 [5] observe a rotation of an Asp towards a His for the deuterated enzyme. In the absence of a statistical database of protein neutron AND ultra-high resolution X-ray crystal structures we have instead examined the effect of deuteration on structure by data-mining of the Cambridge Structural Database [6] for deuterated and hydrogenated pairs of small molecule structures which have been analysed by neutron or X-ray crystallography. There are, mainly, examples of isomorphous crystal pairs but also some non-isomorphous pairs. Differences between these structures for both types have been calculated and their statistical significance assessed. There are precisely enough measured structural differences but we find that they are in each case small enough that they do not upset the general assumption that deuteration does not affect protein structure.

1. P. Atkins & J. de Paula (2002). Atkins' Physical Chemistry, 7th ed.
2. P. Cioni and G. B. Strambini (2002) Biophysical Journal 82, 3246-3253.
3. A. Katrusiak (1995). Physical Review B 51, 589.
4. P. Kuhn et al Biochemistry 1998, 37, 13446-13452
5. X. Liu et al Acta Cryst. (2007). D63, 1000-1008
6. CCDC (2007). http://www.ccdc.cam.ac.uk

Keywords: structure mining, deuterated structures, neutron protein crystallography

MS.98.5

Acta Cryst. (2008). A**64**, C164

Structural database using semantic Web concepts to support structure-Based drug design for AIDS

Talapady N Bhat[1], Talapady N Bhat[1], A Nguyen[1], G Noble[2], L Cooney[2], M Nasr[3], A Wlodawer[4], K Das[5], E Arnold[5]
[1]NIST, Biochemical Science Division, Bldg 227, 100 Bureau Drive, Gaithersburg, MD, 20899, USA, [2]Cygnus Corporation, Inc, Rockville, MD 20852, USA, [3]NIAID, Bethesda, MD 20892, USA, [4]NCI/FCRDC, Frederick, MD 21702, USA, [5]CABM/Rutgers University, 679 Hoes Lan Picatway, NJ 08854, USA, E-mail：bhat@nist.gov

The HIV structural databases (HIVSDB, http://bioinfo.nist.gov/SemanticWeb_pr2d/chemblast.do , http://chemdb2.niaid.nih.gov) distribute one of the largest comprehensive collections of structural, biological and pre-clinical data on inhibitors, drug leads and clinical drugs for AIDS. These databases contain info on several thousand biologically active compounds from all classes (HIV PR, RT, CCR5, Integrase) of FDA approved drugs. Efficient and yet user friendly data management systems that support state-of-the-art annotation, visualization and query capabilities are crucial for the effective use of data for fragment based structural pharmacology and rational drug design. Semantic Web is the vision of the World Wide Web Consortium for enabling seamless integration of electronic data for data mining and knowledge generation across the Web. Robust and functionally relevant ontology plays a critical role in developing the data elements for a Semantic Web. Presentation will illustrate how Semantic Web concepts are used for novel annotation, data integration, storage, and query to manage and display structural (fragments, 2-D images and text-based) biological, and pre-clinical data. One of these techniques (Chem-BLAST(Prasanna, Vondrasek et al. 2006)) developed allows rapid comparison of structural fragments using the Semantics commonly used in drug discovery process. At present majority of the data in HIVSDB are obtained by us by weaning through publications. Our intension is to seek greater participation by the community by depositing data to HIVSDB at the time of publication.

Prasanna, M. D., J. Vondrasek, et al. (2006). "Chemical compound navigator: a web-based chem-BLAST, chemical taxonomy-based search engine for browsing compounds." Proteins 63(4): 907-17.

Keywords: bioinformatics: the future, knowledge-based applications in structural chemist, structure-based drug design, structural informatics