FA5-MS44-P13

Fragment-based Electron Density Interpretation at Low Resolution. Shao-Yang Ku, Thomas R. Schneider
EMBL Hamburg, Germany.
E-mail: s.ku@embl-hamburg.de

The electron density of large macromolecular complexes determined by X-ray crystallography and electron microscopy is often only available to low resolution and is difficult to interpret by conventional methods. Fortunately, a complex often contains known structural components, which can be further divided into rigid fragments [1,2]. To position a fragment in an experimentally phased density map at low resolution requires density fitting in real space. Unlike its reciprocal space counterpart, real space molecular replacement has the advantages of ignoring the "missing part" of the structure and allowing the calculation of correlation between the calculated electron density of a search fragment and the experimental electron density within the mask. Here we use the fast spherical averaged density matching as implemented in Molrep [3] to automatically replace fragments into electron density maps between 4 and 10Å resolution, followed by rigid body refinement to maximize the correlation coefficient between the calculated and experimental maps. The method is implemented by using the Clipper libraries [4]. We use the Rab5 GDP/GTP exchange factor (Rabex-5) in complex with human ubiquitin [5] as a test case. The position of the placed ubiquitin fragment is validated by comparing the rmsd between the placed fragment and the refined Rabex-5 structure. We investigate how resolution, phase error, model error and B factors modeling would influence the quality of density fitting.

[1] Schneider T.R. *Acta Cryst. D.* 2004, 60, 2269. [2] Mosca R., Schneider T.R. *Nucleic Acids Res.* 2008, 36, W42. [3] Vagin A.A., Isupov M.N. *Acta Cryst. D.* 2001, 57, 1451. [4] Cowtan K. (2003) IUCr Computing Commission Newsletter, 2, The Clipper C++ libraries for X-ray crystallography, 4-9 [5] Penengo L., Mapelli M., Murachelli A.G., Confalioneri S., Magri L., Musacchio A., Di Fiore P.P., Polo S., Schneider T.R. *Cell* 2006, 124, 1183.

Keywords: Low resolution crystallography, electron density fitting, algorithms

FA5-MS44-P14

Validation of B-factor Distributions in Protein Crystal Structures. Jacopo Negroni[a], Garib Murshudov[b], Thomas R. Schneider[a]. [a]*EMBL-Hamburg, Germany.* [b]*YSBL, Chemistry Department, University of York, Heslington, York, England.*
E-mail: jacopo.negroni@embl-hamburg.de

Many tools for the analysis of protein models from X-ray crystallography are available nowadays. They check the distribution of geometrical and stereo-chemical properties [1], the agreement of the model with the data [2], or both [3]. These analyses can be either at local or global level. Despite that, a systematic procedure for the analysis and validation of B-factor distributions is still missing. This is surprising since temperature factors play an important role in model interpretation. Moreover, anomalies in the distribution of B-factors can be symptoms of errors introduced during model building and/or refinement. A tool for the detection of these cases would be useful for the interpretation of a protein model already deposited into the Protein Data Bank (PDB) or at the end of the refinement stage.
Here we propose a new approach for the identification of suspicious B-factor distributions in protein models. The main assumption underlying the method is derived from Bayesian statistics and states that isotropic B-Factors in a protein crystal structure follow an Inverse-Gamma Distribution (IGD). A Maximum Likelihood Estimation (MLE) approach is used to estimate the parameters of the IGD that best fit the distribution of B-factors of a protein structure. A Kolmogorov-Smirnov test (K-S test) is then used to evaluate the goodness of fit and compute a p-value.
We developed and tested the new approach on a set of 14229 protein crystal structures selected from the PDB with a resolution of 2Å or higher. We found that for 82% of the PDB structures the p-value was equal or higher than 0.01, indicating a reasonable agreement between the observed distribution and the expected IGD. For some of the structures with a p-value lower than 0.01, their B-factors still satisfied the IGD assumption if their chains were individually analysed. Thus we analysed only single chains from the original set of PDB structures and we found that around 90% of the chains had a p-value equal or higher than 0.01. Furthermore, a re-refinement protocol performed with the experimental version 5.6 of REFMAC [4] was able to rescue some of the outlier structures found with the single chain analysis.
Our work shows that the IGD distribution is a reasonable assumption for the validation of B-factor distributions and the new approach can be used for the detection of suspicious B-factor distributions in protein models.

[1] Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB 3rd, Snoeyink J, Richardson JS, Richardson DC., *Nucleic Acids Res.* 2007 Jul; 35(Web Server issue): W375-83. [2] Vaguine AA, Richelle J, Wodak SJ., *Acta Crystallogr D Biol Crystallogr.* 1999 Jan; 55(Pt 1):191-205. [3] Urzhumtseva L, Afonine PV, Adams PD, Urzhumtsev A., *Acta Crystallogr D Biol Crystallogr.* 2009 Mar; 65(Pt 3): 297-300. [4] Murshudov GN, Vagin AA, Dodson EJ., *Acta Crystallogr D Biol Crystallogr.* 1997 May 1;53(Pt 3):240-55.

Keywords: structure validation, crystal structure analysis, crystal structure properties

FA5-MS44-P15

A stepwise approach for the X-ray diffraction data in Rietveld refinement. O.A. Smirnova. *Institute for Chemical Research, Kyoto University, Uji, Kyoto-fu 611-0011, Japan.*

Consideration of the diffraction data in a way they are collected, i.e., with a step applied by a diffractometer, seems a reasonable way to enhance the structure refinement. With this regard, a number of R-factors to evaluate Rietveld fit are suggested. They account for number of points, automatically referring to an equipment resolution. The new R-factors reflect both goodness of background and peaks fitting while conventional Rietveld R-factors neglect background contribution, sometimes making the R-factors artificially high. The true expressions to the R-factors are:

$R_1 = \Sigma(|I_{obs}-I_{calc}|\ /I_{obs})/N$ $\quad R_{1w} = \Sigma w_i(|I_{obs}-I_{calc}|/I_{obs})/N$

$R_2 = (\Sigma(|I_{obs}-I_{calc}|\ )/N)/(\Sigma I_{bragg}/k_h)$ $\quad R_{2w} = \Sigma(w_i|I_{obs}-I_{calc}|\ )/N)/(\Sigma w_k I_k/k$

$R_3 = (\Sigma(|I_{obs}-I_{calc}|\ )/N)/\ I_{bragg}'$ $\quad R_{3w} = \Sigma(w_i|I_{obs}-I_{calc}|)/N)/w_k I_k')$

where n – number of points;
w_i – weight, $w = 1/\sigma_i$

σ — standard deviance;
I_k – peaks intensuities with $I_k \geq 10 \cdot \sigma_k$
k — a number of peaks with intensities $I_k \geq 10 \cdot \sigma_k$
I_k' – a highest peak intensity;
w_k – weight, $w_k = 1/\sigma_k+1$

The second set of R-factors and goodness of fit are calcuated only for $|I_{obs}-I_{calc}|- \sigma_i$., i.e.,

$R'_1= \Sigma((| I_{obs}-I_{calc}|-\sigma)/I_{obs})/N$ $R'_{1w}= \Sigma w((| I_{obs}-I_{calc}|-\sigma)/I_{obs})/N$
$R'_2= (\Sigma((| I_{obs}-I_{calc}|-\sigma))/N)/(\Sigma I_k/k_h)$ $R'_{2w}= \Sigma(w(| I_{obs}-I_{calc}|-\sigma)/N)/(\Sigma w I_k/k_h)$
$R'_3= (\Sigma(| I_{obs}-I_{calc}|-\sigma))/N)/ I_k'$ $R'_{3w}= \Sigma(w(|I_{obs}-I_{calc}|-\sigma))/N)/(I_k'))$

Keywords: powder diffraction, Rietvld method, R-factors

FA5-MS44-P16

Multialiquot cell approach for the SDPD of high-symmetry compounds. O.A. Smirnova. *Institute for Chemical Research, Kyoto University, Uji, Kyoto-fu 611-0011, Japan.*

Recently mentioned [1] inconsistency of figures of merit [2,3] when indexing high symmetry lattices turned to a conclusion the smaller cells of lower symmetry can be applied as building units when solving a structure by direct space methods. The approach is expected to decrease time necessary for simulated annealing of one structure solution and may appear particularly useful for large organic structures. The poster illustrates the approach based of example compounds with a small pyrochlore structure. The indexing program suggest several possible solutions and the correct solution of highest symmetry among them. The repetition of the same lattice described by different cells should be considered as an indication of the correct indexing solution. From the other hand, that might be random and unfruitful indexing solution if the lattice is non-primitive but is not observed among proposed cells with its primitive representation. The extension to indexing algorithms, eliminating lower symmetry cells for the same lattice described by high-symmetry cell, and the corrected figures of merit taking into account the number of equal proposal cells might be drawn as follows:

$$M'(20) = M(20)_h \cdot N_{ep}$$
$$F'(20) = F(20)_h \cdot N_{ep}$$

where $M(20)_h$ and $F(20)_h$ are M(20) and F(20) for the ***h***ighest symmetry cell

N_{ep} is the number of ***e***qual ***p***roposal cells

Then, one may start to search for the structural model applying a cell of lower/volume symmetry providing it may represent a building unit for the larger cell of higher symmetry or may asssist to find a sublattice..

[1] O.A. Smirnova, in Abstracts, Denver X-ray conference 2009, Colorado Springs 2009, USA. [2] De Wolff, P.M., J. APPL. CRYST. 5, 108-113 (1968). [3] Smith, G. S. & Snyder, R. L., J. APPL. CRYST. 12, 60-65 (1979).

Keywords: powder diffraction, indexing, figures of merit

FA5-MS44-P17

***DDMSuite* – a powder diffraction full-profile analysis system.** Yaroslav I. Yakimov[a], Leonid A. Solovyov[b], Alexander N. Zaloga[a], Igor S. Yakimov[a]. [a]*Siberian Federal University, Krasnoyarsk, Russia.* [b]*Institute of Chemistry and Chemical Technology, Krasnoyarsk, Russia.* E-mail: yar_yakimov@mail.ru

In this presentation we describe a new freeware program package *DDMSuite* designed for crystal structure analysis from powder diffraction data. The program is based on the recently proposed derivative difference minimization (DDM) method of whole-profile fitting [1] as a background-independent alternative to the conventional Rietveld refinement procedure. In this method the refinement is aimed not at minimizing the absolute difference between the observed and calculated patterns, but at minimizing the oscillations (or curvature) of the difference curve. The difference curve is considered as an estimation of background which, in the absence of crystalline admixtures, usually varies much less rapidly along the powder profile than does the diffraction pattern. The main advantage of this method is that it does not involve background line modelling thus avoiding the background-related systematic errors and allowing structure refinement with higher stability and accuracy [2]. Newly developed software implementation *DDMSuite* is intended to provide a free comprehensive use of the DDM method by means of an easy-to-use graphical user interface (GUI). The program includes both DDM and Rietveld refinement procedures for X-ray and neutron powder diffraction data, profile decomposition routines, the quantitative phase analysis and the size-strain calculations. Great effort has been made to design a versatile tree-type phases/atoms hierarchy navigation. It gives convenient means for macro-editing parameters of multiple selected phases and atoms. A dedicated dialog allows controlling the refinement process facilitated by a number of graphical tools: structure 3D-view; powder pattern plot; Fourier and Patterson mapper. The GUI has interfaces to a number of widely used crystallographic software packages (Diamond, CCDC Mercury etc.) and CIF import-export routines. Applications to various types of powder diffraction data including semi-crystalline substances such as mesostructured materials and complicated multi-phase samples will be demonstrated. The program (to date, a Microsoft Windows version) can be freely downloaded from:
http://www.icct.ru/eng/content/persons/Sol_LA/ddm.html.

[1] Solovyov L.A., *J. Appl. Crystallogr.*, 2004, 37, 743. [2] Solovyov L.A., in *Powder Diffraction Theory and Practice,* ed. Dinnebier R.E., Billinge S.J.L., 2008, 282.

Keywords: powder diffraction software, full-profile refinement, derivative difference minimization

FA5-MS44-P18

On execution of Pawley method without requiring intensity constraints on overlapping reflections. R. Oishi-Tomiyasu[a], M. Yonemura[a], A. Hoshikawa[b], S. Torii[a], T. Ishigaki[b], T. Kamiyama[a]. [a]*High Energy Accelerator Research Organization, Tsukuba, Ibaraki, Japan.* [b]*Ibaraki University, Hitachi, Ibaraki, Japan.* E-mail: ryoko.tomiyasu@kek.jp

The Pawley method and the Le Bail method are two major methods for extraction of integrated intensities from powder diffraction patterns. The advantage of the Pawley method to the Le Bail method is that it can directly obtain the intensity covariance matrix. In both methods, it is considered to be necessary to group overlapping reflections and impose linear