ON INTEGRATING THE TECHNIQUES OF DIRECT METHODS AND SIRAS. II. THE INITIAL APPLICATIONS

H. Xu[1, 2] H. Hauptman[1, 2]
[1]Hauptman-woodward Medical Research Institute, 73 High Street, Buffalo, NY 14203-1196, USA [2]Department of Structural Biology, SUNY at Buffalo, USA

The joint probability distribution of three structure factors in the SIRAS case has been recently derived [1]. This distribution leads directly to formulae for estimating a single phase of a derivative structure, when a heavy-atom substructure is known. The formulae were tested with the protein S-adenosylhomocystein (AdoHcy) hydrolase, having molecular weight about 48000, space group *C*222, and its selenium derivative obtained by replacing 30 sulfur atoms in the native protein by selenium atoms [2]. The Se-atom substructure, presumed to be known, had been previously obtained by the Shake-and-Bake direct method using single wavelength anomalous diffraction data for the derivative. All error free SIRAS data, normal diffraction for the native protein and single wavelength anomalous diffraction for derivative, used in our test were calculated from known atomic parameters to a resolution of 2.8 Å. The resultant phase errors, arranged in decreasing order of expected reliability (cumulative mean-phase error in degrees for the derivative) show that these probabilistic formulae, combined with the use of heavy-atom substructure information, are capable of estimating large numbers of individual phases with very high accuracy (Table).

Group No.	1	2	3	4	5	6	7	8
No. phases in group	100	500	1000	2000	5000	10000	20000	26713
Mean-phase error	13.5	23.4	26.9	30.6	32.7	34.7	43.7	51.6

This Research was supported by grants GM-46733 (NIH) and ACI-9721373 (NSF). References

1. Hauptman, H.A. & Xu, H. 19th Congress and General Assembly, August 2002, Geneva, Switzerland.
2. Turner, M.A. et al. (1998). Nature Struct. Biol. 5, 369-376.

Keywords: **DIRECT METHODS ANOMALOUS SCATTERING**

CRYSTAL STRUCTURE OF A CORONAVIRUS MAIN PROTEINASE: TGEV Mpro

K. Anand G. Palm J. Mesters J. Ziebuhr R. Hilgenfeld
Institute of Molecular Biotechnology Dept. of Structure Biology Beutenbergstrasse 11 JENA THURINGIA 07745 GERMANY

Porcine transmissible gastroenteritis virus (TGEV), a positive-strand RNA virus of the coronavirus family, employs extensive proteolytic processing of large polyprotein precursors to control the activities of the viral replication-transcription complex.The key enzyme in this process is a 33.1 kDa cysteine protease called coronavirus main proteinase (Mpro) or 3CLpro. The structure of TGEV proteinase was determined using multiple-wavelength anomalous dispersion (MAD) [1] data. The bottleneck of the phasing process was the location of 60 selenium positions (6 monomers/a.u. with 10 SeMet residues each). Solutions for the selenium substructure were obtained with the program SnB v2.0 [2]. Five data sets were merged together to get high redundancy. The atomic positions derived from the best SnB solutions were examined for NCS. Six fold NCS symmetry predicted a further 11 positions in addition to the 37 sites found by SnB. All 48 positions were used in MLPHARE for phasing, followed by solvent flattening in DM. Subsequent refinement using the program CNS [3] converged to Rfree = 25.6% and R = 21.0%. With the typical chymotrypsin-like β-barrel fold of domains I and II and a Cys-His catalytic dyad at the active site, TGEV proteinase shares structural features of both serine and cysteine proteases. Domain III has a unique, largely α-helical fold.

References

[1] Hendrickson, W.A., et al., (1990) EMBO J., 9 1665-1672
[2] Miller, R., et al., (1994) J. Appl. Crystallogr.27, 613-621
[3] Brunger, A.T et al., (1998) Acta Crystallogr. D54, 905-921

Keywords: **CORONAVIRUS PROTEASE, MAD PHASING, 60 Se SITES**

CRYSTALLOGRAPHIC STUDIES OF THERMUS THERMOPHILUS UDP-N-ACETYLGLUCOSAMINE 2-EPIMERASE AT 2.28 Å RESOLUTION

B. Bagautdinov N. Kunishima K. Nishio A. Paehler M. Miyano
RIKEN Harima Institute, High Throughput Factory, 1-1-1, Kouto, Mikazuki-cho, Sayo-gun, Hyogo, 679-5148 Japan

Recent development of the software tools available for the determination of heavy atom positions have dramatically improved the efficiency and speed of the calculation needed in the case of a very large number of incorporated heavy atom sites in the asymmetric unit. We applied the programs SnB, SOLVE and SHELXD to find the positions of the 30 expected Se atoms in the structure of SeMet UDP-N-acetylglucosamine 2-epimerase. Selenomethionine was introduced into the structure to allow MAD phasing. Native crystals of *Thermus thermophilus* UDP-N-acetylglucosamine 2-epimerase crystallize in the hexagonal space group $P3_12_1$ with unit cell parameters: a = b = 70.773 Å, and c = 139.002 Å. The SeMet analog also crystallizes in the same hexagonal space group $P3_12_1$, but with unit cell parameters a = b = 70.097 Å, and c = 282.96 Å. Data to 2.28 Å resolution were collected with an ADSC Quantum detector at the SPring-8 beamline BL38B1 using cryo-cooled crystals. The diffraction intensities of the SeMet analog at four different wavelengths near the Se K-edge were used to solve the phase problem. The best Se positions were used for the initial phasing. Efforts to improve the initial phases and extending the phases to higher resolution by model building, phase combination and density modification techniques are under way. In our poster details of the structure analysis and a comparison of the molecular packing and conformational differences between *Thermus thermophilus* UDP-N-acetylglucosamine 2-epimerase and the SeMet analog and the *E. coli* monoclinic form determined at 2.5 Å by R.E.Campbell et al. (Biochemistry, 14993-15001, 39, 2000) will be presented.

Keywords: PHASE PROBLEM MAD ENZYME

SAD/MAD PHASING AT HIGH RESOLUTION USING TANTALUM BROMIDE CLUSTER

S. Banumathi[1] M. Dauter[2] Z. Dauter[3]
[1]NCI, Brookhaven National Laboratory [2]SAIC, Brookhaven National Laboratory [3]NCI, Brookhaven National Laboratory

The tantalum bromide cluster has been used for phasing large macromolecular structures at relatively low resolution. Tantalum exhibits a significant anomalous scattering signal, with the strong white line at its LIII absorption edge at 1.25 Å wavelength, which can be used for MAD or SAD phasing. The cluster consists of six tantalums in an octahedral arrangement surrounded by twelve bromides. At low resolution the cluster can be approximated by a spherical superatom of about 5 Å diameter. Owing to the compact shape and symmetry of the cluster, the anomalous effect is well pronounced at low and high resolution ranges, but diminishes at the intermediate range of 3 – 4 Å. Bromides may also contribute a potentially useful anomalous signal. After soaking into the protein crystals, the cluster binds at the protein surface in an ordered or rotationally disordered fashion. At low resolution it can be easily identified as a superatom, but high resolution is required to resolve individual tantalum atoms, which then can be used to estimate phases. Results obtained with tantalum bromide cluster using different phasing approaches will be presented.

Keywords: **SAD/MAD PHASING, SOAKING, TANTALUM**