Structure determination of *Streptomyces plicatus* β-*N*-acetylhexosaminidase by MAD phasing

Brian L. Mark[1], Barbara L. Triggs-Raine[2], and Michael N.G. James[1]
[1] *MRC Group in Protein Structure and Function, University of Alberta, Edmonton, Alberta, Canada T6G 2H6*
[2] *Department of Biochemistry and Medical Genetics, University of Manitoba, Winnipeg, Manitoba, Canada, R3E 0W3*

Introduction

Carbohydrates are involved in a wide spectrum of biological processes including structure, food storage, viral invasion, and cellular signaling. Glycosyl hydrolases are the enzymes responsible for degrading carbohydrates and have been classified into families [1]. Family 20 includes β-*N*-acetylhexosaminidases (β-hex) and chitobiases. β-hex enzymes remove *N*-acetylhexosamine residues from the nonreducing ends of various carbohydrates and their conjugates by hydrolyzing the β-1,4 linkage that attaches them to the rest of the carbohydrate chain. Chitobiases hydrolyze the β-1,4 linkage within the disaccharide di-*N*-acetylglucosamine (chitobiose). The only three-dimensional structure of a family 20 protein is that of a chitobiase from *Serratia marcescens* [2].

In humans, there are two β-hex isoforms: Hex A and B. Hex A is a heterodimer of subunits α and β (encoded by *HEXA* and *HEXB*, respectively) and is essential for degrading G_{M2} ganglioside. Hex B is a homodimer of β subunits and degrades soluble carbohydrates. Mutations in *HEXA* or *HEXB* cause the severe neurodegenerative disorders known as Tay-Sachs and Sandhoff disease, respectively. Unfortunately, the three-dimensional structures of these human β-hex isoforms have not been determined primarily due to an insufficient amount of protein. However, it appears that the catalytic mechanism and active site architecture of this family of proteins may be conserved [2, 3]. Thus, investigation into the structure of family 20 enzymes, such as *S. plicatus* β-hex and *Serratia marcescens* chitobiase, should provide structural evidence into the molecular basis of Tay-Sachs and Sandhoff disease.

We have now successfully determined the three dimensional structure of a family 20 β-hex from *Streptomyces plicatus* to 2.2 Å resolution using a multiwavelength anomalous diffraction (MAD) phasing experiment [4] performed on selenomethionine-substituted protein crystals at the BioCARS sector, beamline 14-BM-D.

Methods and Materials

S. plicatus β-hex (56 kDa) was expressed in *Eschericia coli* strain BL21(DE3) pLysS. Bacteria were grown in minimal medium until they reached an OD_{600} of 0.5, whereupon the culture was supplemented with lysine, threonine, phenylalanine, leucine, isoleucine, and valine to inhibit methionine biosynthesis [5]. Seleno-L-methionine was then added followed by IPTG to induce the expression of the Se-met recombinant protein. Crystals of purified *S. plicatus* β-hex were grown by vapor diffusion and were found to belong to space group $P6_122$ with unit cell dimensions of a = b 133.1 Å and c = 176.7 Å.

At beamline 14-BM-D, a crystal (0.3 x 0.3 x 0.2 mm) was flash cooled to ~100 K using a rayon loop (Hampton Research) in N_2(g) cryostat (Oxford Instruments) with 25% glycerol as cryoprotectant. A fluorescence spectrum was measured across the Se K-absorption edge. Reflections were measured using the rotation method to a minimal Bragg spacing of 2.2 Å at 1) the absorption peak (λ = 0.9795 Å, Δφ = 0.5°, total φ = 75°), 2) the inflection point (λ = 0.9796 Å, Δφ = 0.5°, total φ = 75°), and 3) a high-energy remote point (λ = 0.9496 Å, Δφ = 0.5°, total φ = 75°). An ADSC Quantum-4 detector was used to measure the diffraction and the collected data were processed using DENZO and SCALEPACK (R_{sym} = 0.04, completeness = 99.9%).

The program SOLVE was used for local scaling of data and to determine the anomalous and dispersive differences used to find Se sites and calculate phase probability distributions. Density modification was performed using DM. Electron density maps were calculated using CCP4. Skeletonization of the map was carried out using MAPMAN and a molecular model was built using O v7.0.0. Refinement of the molecular model is being carried out with CNS using amplitudes from a data set of native *S. plicatus* β-hex collected to 2.2 Å at the monochromatic beamline 14-BM-C. This native data set was collected by the rotation method using an ADSC Quantum-4 detector (λ = 1.0 Å, Δφ = 0.5°, total φ = 75) and processed using DENZO and SCALEPACK (R_{sym} = 0.03, completeness = 99.8%).

Results

Combination of the MAD-phasing experiment at beamline 14-BM-D and data collection from native *S. plicatus* β-hex crystals at beamline 14-BM-C allowed for the determination of the three-dimensional structure of *S. plicatus* β-hex to 2.2 Å resolution. The quality of the data allowed us to clearly find the five Se atoms present in a 512 amino acid protein using Pattersons calculated from the anomolous and dispersive differences.

An excellent electron density map was generated using the experimental phases obtained from the MAD-phasing experiment (initial figure of merit was 0.8). The initial map was improved slightly by solvent flattening and a molecular model was then readily built into it. The structure of *S. plicatus* β-hex reveals a compact fold with two domains that appear homologous with domains found in the *Serratia marcescens* chitobiase crystal structure. Further analysis of the model will be published shortly.

Discussion

The quality of data collected during the MAD-phasing experiment at 14-BM-D allowed us to easily find five Se atoms in a 512 amino acid protein. Although diffraction was seen out to 1.8 Å, data collection was restricted to 2.2 Å for fear of losing data to overlaps due to the unit cell dimensions. Nonetheless, the MAD-phasing experiment resulted in experimental phases of exception quality that were used to calculate easily interpretable maps. A data collection strategy will be planned to take advantage of the high-resolution reflections in the near future.

The model has currently been refined to R_{cryst} of 0.25, and, once refinement is complete, the model will be used to further understand the catalytic mechanism of this family of glycosyl hydrolases.

Acknowledgments

We thank Ken Ng, Marie Fraser, William Wolodko, and the BioCARS staff for their assistance with data collection at beamline 14-BM-D and 14-BM-C. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract No. W-31-109-Eng-38. Use of BioCARS sector 14 was supported by the National Institutes of Health, National Center for Research Resources, under grant number RR07707.

References

[1] B. Henrissat and A. Bairoch, *Biochem. J.* **293**, 781–788 (1993).
[2] I. Tews, A. Perrakis, A. Oppenheim, Z. Dauter, K.S. Wilson, and C. Vorgias, *Nature Struc. Bio*. **3**, 638–648 (1996).
[3] B.L. Mark, G.A. Wasney, T. Salo, A.R. Khan, Z. Cao, P.W. Robbins, M.N. James, and B.L. Triggs-Raine, *J. Biol. Chem.* **273**, 19618–19624 (1998).
[4] W.A. Hendrickson, *Science* **254**, 51–58 (1991).
[5] G.D. Van Duyne, R.F. Standaert, P.A. Karplus, S.L. Schreiber, and J. Clardy, *J. Mol. Biol* **229**, 105–124 (1993).