Structure determination of *Ascaris* pepsin inhibitor-3 by MAD phasing

Kenneth K.S. Ng[†], Jens F.W. Petersen[†], Maia M. Cherney[†], Craig Garen[†], Michael N.G. James[†], Jeffrey L. Zalatoris[¶], Chetana Rao-Naik[¶], Ben M. Dunn[¶], Mark R. Martzen[‡], and Robert J. Peanasky[‡]

[†]*MRC Group in Protein Structure and Function, University of Alberta, Edmonton, Alberta, Canada T6G 2H6*
Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, FL 32610-0245 USA
[‡]*Department of Biochemistry and Molecular Biology, University of South Dakota, Vermillion, SD USA*

Introduction

Aspartic proteinases play important roles in a wide range of normal and pathological processes. For example, pepsin is the major gastric proteinase of mammals, renin plays a key role in haemostasis by cleaving angiotensinogen to angiotensin I, and the HIV protease is essential for the maturation and replication of the virus that causes AIDS [1]. Because of the important roles played by aspartic proteinases in many human diseases, they have been attractive targets for the design of enzyme inhibitors acting as drugs. Most drug design efforts have focused on low-molecular-weight inhibitors, some of which are already in widespread use in the treatment of AIDS. In contrast, natural proteinaceous inhibitors of aspartic proteinases have been poorly studied, even though many proteinaceous inhibitors are known to be highly potent and specific inhibitors of proteinases.

The best studied proteinaceous inhibitor of aspartic proteinases is pepsin inhibitor-3 (PI-3) from the intestinal nematode parasite *Ascaris* [2]. PI-3 is believed to protect the worm from gastric aspartic proteinases in the stomach of the worm's host [3]. PI-3 inhibits pepsin, gastricsin, and cathepsin E with K_is of 1–10 nM, but it does not inhibit closely related enzymes like renin and chymosin [4]. Understanding the molecular basis for the tight binding and specificity of PI-3 requires a detailed analysis of the structure of PI-3 and its complexes with target enzymes using high-resolution structural methods like x-ray crystallography.

PI-3 does not share sequence homology with proteins of known three-dimensional structure. Previous attempts to solve the structure of PI-3 (heavy-atom soaking and isomorphous replacement phasing and molecular replacement using the known structure of porcine pepsin and crystals of the pepsin:PI-3 complex) failed to provide sufficient phasing information for determining the structure of PI-3. As a result, a MAD (multiwavelength anomalous dispersion) [5] phasing experiment was performed on selenomethionine-substituted protein crystals at the BioCARS sector, beamline 14-BM-D.

Methods and Materials

PI-3 was expressed in *Eschericia coli* strain B834(DE3) pLysS (a methionine auxotroph) using a T7-phage expression system (pET-3a). Bacteria were grown in minimal medium supplemented with seleno-L-methionine. Inclusion bodies of PI-3 were solubilized in 8 M urea and the denatured protein was refolded by stepwise dialysis into Tris-Cl and Na-MOPS buffers at pH 11 and 7, respectively. The protein was further purified by ammonium sulfate precipitation and gel filtration chromatography. The K_i for porcine pepsin was found to be approximately 5 nM.

Crystals of PI-3 were grown by vapor diffusion using PEG 8000 as a precipitant. The crystal belongs to space group $P2_1$ with unit cell dimensions a = 30.5 Å, b = 72.7 Å, c = 31.1 Å, $\beta = 102°$. There is one protein molecule in the asymmetric unit (V_m = 2.05 Å^3/Da).

At beamline 14-BM-D, a crystal (0.2 mm x 0.2 mm x 0.1 mm) was transferred for 2–3 seconds to a cryosolvent solution containing 25% glycerol and the crystal was flash cooled to ~100 K using a rayon loop (Hampton Research) and N_2(g) cryostat (Oxford Instruments). A fluorescence spectrum was measured around the Se K-absorption edge. Data were measured by the rotation method at the absorption peak (λ = 0.9793 Å; $\Delta\phi = 0.5°$, total $\phi = 360°$), inflection point (λ = 0.9797 Å; $\Delta\phi = 1°$, total $\phi = 180°$), and high-energy remote point (λ = 0.9496 Å; $\Delta\phi = 1°$, total ϕ = 180°) using an ADSC Quantum-4 detector. Data were processed using DENZO and SCALEPACK (d_{min} = 2.1 Å, R_{sym} = 0.07, completeness = 0.98).

Anomalous and dispersive differences, sites for Se atoms, and phase probability distributions were calculated using SOLVE. Density modification and phase combination were carried out using DM and SIGMAA. Electron density maps were calculated using CCP4 and XtalView. A molecular model for 125 residues was constructed using wARPnTRACE and XFIT. Higher-resolution data were measured from a native crystal using CuK_α radiation from a rotating anode generator and a DIP 2030B image plate detector (d_{min} = 1.75 Å, R_{sym} = 0.06, completeness = 0.99). The molecular model of PI-3 was refined against the native data set (R_{cryst}= 0.22, R_{free} = 0.26) using CNS.

Results

The MAD-phasing experiment conducted at beamline 14-BM-D allowed for the determination of the structure of PI-3 to 2.1 Å resolution. Five Se atoms in the asymmetric unit were clearly identified from Patterson syntheses using anomalous and dispersive differences (Figure 1). Phase probability distributions were calculated and improved by density modification techniques (initial figure of merit 0.69, final 0.87). The resulting electron-density maps were of

exceptional quality and allowed the rapid construction of a molecular model (Figure 2). The model has subsequently been refined against native data to 1.75 Å resolution.

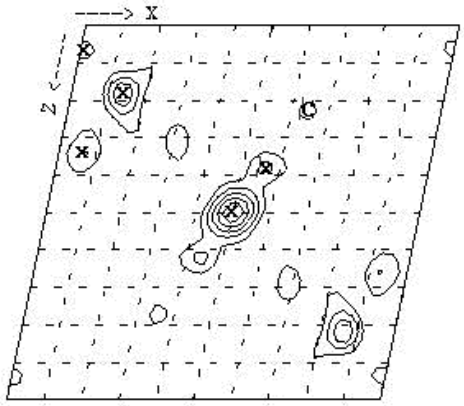

Figure 1: Patterson map (Harker section, y = 1/2) calculated from $|F_A|$ coefficients using SOLVE. Contoured at one σ steps. x denotes Harker peak. c denotes cross peak.

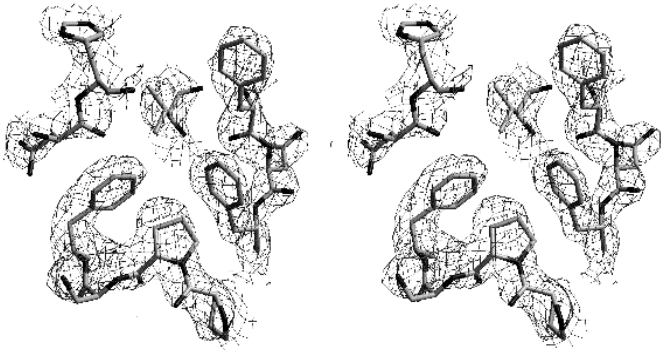

Figure 2: Electron density map ($|F_o|\exp(\alpha_{MAD/DM})$ coefficients) contoured at one σ. Final refined model drawn as sticks. Stereoscopic view.

The structure of PI-3 reveals a novel fold and explains patterns of sequence conservation seen in homologous proteins, as well as site-directed mutagenesis and chemical modification experiments. The structure of PI-3 has also been used to solve the structure of the porcine pepsin:PI-3 complex using molecular replacement techniques. The structure of the pepsin:PI-3 complex has been refined to 2.45 Å resolution, allowing a detailed explanation of the mechanism of inhibition and indicating the determinants of binding specificity.

Discussion

The use of MAD-phasing methods was essential for determining the structure of PI-3. Previous attempts to obtain isomorphous heavy-atom derivatives were not successful, and previous attempts to solve the structure of the pepsin:PI-3 complex using phases calculated from pepsin alone did not provide sufficient information to build a model of PI-3 [6].

The favorable ratio of anomalous scattering atoms (5 ordered Se) to total protein atoms (997 C, O, N, and S atoms in 125 residues in the asymmetric unit) provided a clear fluorescence signal for the optimization of monochromator settings around the K-absorption edge of Se. Accurate diffraction measurements at the BioCARS 14-BM-D beamline allowed the calculation of high quality difference Patterson (Figure 1) and experimentally phased electron density maps (Figure 2).

The high quality of the experimentally phased electron density maps allowed the structure of PI-3 to be built and refined in a few weeks. This has subsequently allowed the rapid solution of the structure of the pepsin:PI-3 complex. Both of these structures have now been refined at 1.75 Å and 2.45 Å resolution, respectively. These structures reveal, for the first time, the novel fold for a small family of nematode proteins related to PI-3 and indicate how PI-3 acts as a potent and specific inhibitor of aspartic proteinases. These structures provide a basis for the design of a novel class of inhibitors targeted against a wide range of aspartic proteinases involved in human disease.

A preliminary report on the structures of PI-3 and the pepsin:PI-3 complex has been presented at the International Conference on Proteinase Inhibitors (Gainesville, Florida, Dec. 3–6, 1999). A paper describing the structure determination and providing a detailed analysis of the structures of PI-3 and the pepsin:PI-3 complex has been prepared and will be submitted for publication.

Acknowledgments

We thank Marie Fraser, Brian Mark, and the staff of BioCARS for assistance with the MAD x-ray data collection. 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] M.N.G. James, *Aspartic Proteinases: Retroviral and Cellular Enzymes*, (Plenum Press, New York, 1998).

[2] M.R. Martzen, B.A. McMullen, N.E. Smith, K. Fujikawa, and R.J. Peanasky, *Biochem.* **29**, 7366–7372 (1990).

[3] R.J. Peanasky, M.R. Martzen, G.A. Homandberg, J.M. Cash, D.R. Babin, and B. Litweiler, *Molecular paradigms for eradicating helminthic parasites* (ed.) (Alan R. Liss, Inc., New York, 1987) 349–366.

[4] M.J. Valler, J. Kay, T. Aoyagi, and B.M. Dunn, *J. Enzyme Inhibition* **1**, 77–82 (1985).
[5] W.A. Hendrickson, *Science* **254**, 51–58 (1991).
[6] J.F.W. Petersen, M.M. Chernaia, C. Rao-Naik, J.L. Zalatoris, B.M. Dunn, and M.N.G. James, *Aspartic Proteinases: Retroviral and Cellular Enzymes* (ed. M.N.G. James) (Plenum Press, New York, 1998) 391–395.