Journal of Structural Biology 166 (2009) 79–87

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Crystal structure of a fungal protease inhibitor from *Antheraea mylitta*

Sobhan Roy [a,1], Penmatsa Aravind [b,1], Chaithanya Madhurantakam [a,2], Ananta Kumar Ghosh [a], Rajan Sankaranarayanan [b,*], Amit Kumar Das [a,*]

[a] Department of Biotechnology, Indian Institute of Technology, Kharagpur 721302, India
[b] Centre for Cellular and Molecular Biology, Council of Scientific and Industrial Research, Hyderabad, India

ARTICLE INFO

Article history:
Received 3 October 2008
Received in revised form 19 December 2008
Accepted 20 December 2008
Available online 19 January 2009

Keywords:
Antheraea mylitta
Canonical protease inhibitor
Reactive site
Disulfide linkage
Fungal protease inhibitor

ABSTRACT

Indian tasar silk is produced by a wild insect called *Antheraea mylitta*. Insects do not have any antigen–antibody mediated immune system like vertebrates but they produce a wide variety of effector proteins and peptides possessing potent antifungal and antibacterial activity to combat microbial attack. *Antheraea mylitta* expresses a fungal protease inhibitor AmFPI-1, in the hemolymph that inhibits alkaline protease of *Aspergillus oryzae* for protection against fungal infection. AmFPI-1 is purified from the hemolymph, crystallized and the structure is solved using the single isomorphous replacement with anomalous scattering (SIRAS) method to a resolution of 2.1 Å. AmFPI-1 is a single domain protein possessing a unique fold that consists of three helices and five β strands stabilized by a network of six disulfide bonds. The reactive site of AmFPI-1 is located in the loop formed by residues 46–66, wherein Lys54 is the P_1 residue. Superimposition of the loop with reactive sites of other canonical protease inhibitors shows that reactive site conformation of AmFPI-1 is similar to them. The structure of AmFPI-1 provides a framework for the docking of a 1:1 complex between AmFPI-1 and alkaline protease. This study addresses the structural basis of AmFPI-1's specificity towards a fungal serine protease but not to mammalian trypsin and may help in designing specific inhibitors against fungal proteases.

1. Introduction

Proteases are ubiquitous in living systems and show an intimate relationship with proteins from synthesis to their breakdown. Proteases take part in various physiological processes such as food digestion, blood clotting, clot dissolution, embryogenesis, tissue regeneration, defense mechanisms and immune responses (Laskowski et al., 2000). In general, protease activity is modulated by means of regulated expression/secretion, proteolytic activation of their inactive forms, auto-inactivation or autolysis, transportation and inhibition by specific protease inhibitors (Bode and Huber, 2000; Fodor et al., 2005). Hence protease inhibitors play an important role in all living organisms through regulation of proteolysis. During the last two decades, a number of protease inhibitors have been studied to understand their physiological role and their potential use as therapeutic agents.

Multiple families (~48) of structurally distinct protease inhibitors have been identified with the capability to inhibit serine, cysteine, metallo and aspartyl proteases (Rawlings et al., 2004). These inhibitors interact with proteases through various structural elements, such as N- or C-terminal exposed loops, either separately or in association with other similar structural elements (Otlewski et al., 2005). Serine protease inhibitors are classified into three categories: canonical, non-canonical and serpins. Canonical inhibitors form the largest family and act through a standard mechanism of inhibition (Laskowski and Kato 1980; Krowarsch et al., 2003). These inhibitors are rigid, stable and mostly composed of purely β-sheet or mixed α/β topologies. However, some canonical protease inhibitors are only α-helical or form irregular structures rich in disulfide bridges (Otlewski et al., 2005). About 18 families of canonical inhibitors have been identified which show a common three-dimensional architecture forming the exposed loop surrounding the reactive site. These residues, in the reactive loop have a β-strand architecture and the P_1 residue (Schechter and Berger nomenclature, Schechter and Berger, 1967) has a 3_{10}-helical conformation. Inhibition is achieved through several intermolecular interactions between the protease active site and the inhibitor reactive loop: formation of two hydrogen bonds between the carbonyl oxygen of P_1 and the amides of the oxyanion hole, and a short contact between the P_1 carbonyl carbon and the catalytic serine (Otlewski et al., 2005). Non-canonical protease inhibitors, like hirudin and haemedin interact with the active site of serine proteases like thrombin, through their N-terminal tails in such a manner that the N-terminal tail is inserted into the enzyme active site,

* Corresponding authors. Fax: +91 40 27160591 (R. Sankaranarayanan), +91 3222 278707 (A.K. Das).
E-mail addresses: sankar@ccmb.res.in (R. Sankaranarayanan), amitk@hijli.iitkgp.ernet.in (A.K. Das).
[1] These authors contributed equally.
[2] Present address: Center for Infectious Medicine, Department of Medicine, F59, Karolinska University Hospital, Hudding, Stockholm, 141 86, Sweden.

1047-8477/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2008.12.010

forming a parallel β-sheet with the protease (Grutter et al., 1990; Richardson et al., 2000). Finally, serpins are 45–55 kDa proteins mostly found in the blood plasma, where they account for coagulation, complement activation and fibrinolysis. Serpins inhibit their targets using a reactive center loop (RCL) present at their C-terminus (Potempa et al., 1994; Gettins, 2002). The initial recognition of the exposed RCL is similar to canonical inhibitors and the protease attacks the P_1—P'_1 bond as a potential site of the substrate. The subsequent attack by the catalytic serine residue on the serpin P_1—P'_1 peptide bond leads to an acyl-enzyme intermediate in which the newly formed amino group dissociates from the active site, followed by insertion of the RCL into β-sheet and flipping the covalently attached protease to the opposite side of the serpin (Huntington et al., 2000).

Specificity of inhibition depends on the structures of both the protease and the inhibitor. Species specificity for trypsin inhibition has been well documented for insect protease inhibitors like LMPI-1 and SGPI-1 from *Locusta migratoria* and *Schistocerca gregaria*, respectively (Patthy et al., 2002; Kellenberger et al., 2003). These two inhibitors are able to discriminate between the bovine and porcine trypsin on one side, and fungal and arthropod trypsin on the other side through the unusual presence of a Lys at the P'_1 position, internal rigidity, and the presence of a P_6–P_{10} loop which provides a secondary site of interaction with the protease (Kellenberger and Roussel, 2005). A taxon-specific inhibitor (SGTI) from *S. gregaria* inhibits arthropod trypsins five times more effectively than mammalian trypsins, accomplished by extended interactions in addition to the primary binding residues P_3—P'_3 of SGTI (Fodor et al., 2005).

Though insects lack the immune system like vertebrates involving antigen–antibody reactions, they can protect themselves from entomopathogenic infections efficiently. The humoral part of the insect immune system is characterized by a rapid and transient synthesis of proteins or peptides with potent antibacterial and antifungal activity (Gillespie et al., 1997). Entomopathogenic fungi infect susceptible hosts directly through the integument and utilize different proteases to carry out the digestion of cuticle proteins in order to colonize and to inactivate the host immune systems (Clarkson and Charnley, 1996; Frobius et al., 2000). Insects protect themselves from this type of fungal infections by expressing a wide spectrum of protease inhibitors to inhibit the fungal proteases (Eguchi et al., 1994; Nirmala et al., 2001a,b; Zheng et al., 2007).

Antheraea mylitta is a wild silkworm species which is endemic to India. It produces tasar silk and feeds on non-mulberry plants. A protease inhibitor from hemolymph of *A. mylitta* was isolated, biochemically characterized and named as *A. mylitta* fungal protease inhibitor-1 or AmFPI-1 (Shrivastava and Ghosh, 2003). AmFPI-1 shows sequence homology only with the partially sequenced inducible serine protease inhibitor ISPI-1 of *Galleria mellonella*. But the sequence homology failed to indicate the reactive site residues as well as its structure–function relationships. Hence we determined the crystal structure of AmFPI-1 to understand the mechanism of inhibition. We have previously reported the crystallization of this protease inhibitor, AmFPI-1, purified from the hemolymph of the *A. mylitta* (Roy et al., 2006). Here, we present the crystal structure of AmFPI-1 at a resolution of 2.1 Å and docking studies to understand the structural basis for the specificity of AmFPI-1 towards fungal serine proteases over mammalian trypsin.

2. Materials and methods

2.1. Native protease inhibitor purification, crystallization and heavy atom derivative preparation

Purification and crystallization of the AmFPI-1 from the hemolymph of *A. mylitta* was carried out as described earlier (Roy et al., 2006). In brief, proteins were precipitated from 30 ml of hemolymph using 60% ammonium sulfate and pelleted by centrifugation. The pellet was resuspended in citrate buffer (10 mM sodium citrate, pH 6.0, 50 mM NaCl), dialyzed against the same buffer, heat-treated at 80 °C for 2 min and centrifuged to remove denatured proteins. Supernatant was loaded onto a Q-Sepharose column (Amersham) equilibrated with citrate buffer. Flowthrough fractions were collected, dialyzed against buffer containing 10 mM Tris pH 7.5 and 50 mM NaCl and loaded onto a SP-Sepharose column (Amersham). Bound proteins were eluted from the column with a linear gradient of 0.1–1 M NaCl. Fractions containing AmFPI-1 were pooled and further purified by gel-exclusion chromatography using a Sephadex G-75 column (Amersham). Purified AmFPI-1 was concentrated to 10 mg/ml in 10 mM Tris pH 7.5 and 50 mM NaCl and crystals were obtained from 0.2 M ammonium sulfate, 0.1 M bis–tris pH 6.5 and 30% w/v polyethylene glycol 3,350 at 25 °C. Heavy atom derivative crystal was prepared by overnight (14 h) soaking in a solution of 4.5 mM Potassium tetra chloro platinate (K_2PtCl_4) in mother liquor solution followed by an additional soaking with 6 mM K_2PtCl_4 in mother liquor for 30 min before X-ray diffraction data collection.

2.2. Data collection and processing

X-ray diffraction data collection and processing of the native AmFPI-1 have been reported previously (Roy et al., 2006). A complete dataset of the Pt derivative of an AmFPI-1 crystal was collected on a MAR Research MAR-345dtb image-plate detector attached to a Rigaku RU-H3R rotating-anode generator producing CuKα radiation equipped with an Osmic mirror system and operated at 50 kV and 100 mA. The crystal was flash cooled in a liquid-nitrogen stream at 100 K using an Oxford cryostream controller. Fifteen percentage of glycerol in mother liquor solution was used as a cryoprotectant. The intensities of reflections were integrated and scaled using DENZO and SCALEPACK (Otwinowski and Minor, 1997), respectively. Data collection statistics are shown in Table 1.

2.3. Phasing

The structure of AmFPI-1 was determined by phasing with the single isomorphous replacement anomalous scattering (SIRAS) method using the Pt derivative. The native and Pt derivative datasets were merged and scaled together using the programs CAD and Scaleit respectively (Collaborative Computational Project Number 4, 1994). Single heavy atom position with a clear Patterson peak was located using SOLVE (Terwilliger, 2004) and the position was checked in isomorphous difference and anomalous difference Patterson maps made using the program Generate Patterson map (Collaborative Computational Project Number 4, 1994). Refinement of heavy atom parameters and phase calculation was carried out using SOLVE (Terwilliger, 2004). Identification of a single Pt site helped in phasing with a mean FOM of 0.45 to a resolution of 2.4 Å. The electron density map was generated through Fourier transformation. Iterative solvent flattening and histogram matching were performed using DM (Collaborative Computational Project Number 4, 1994). RESOLVE (Terwilliger, 2004) was used for maximum likelihood density modification and automated model building. The resulting electron density map with a partial model revealed clear main chain density with substantial side chain details. The model was manually skeletonized using O (Jones et al., 1991). One residue at the C-terminus (Arg86) was not modeled due to lack of electron density.