Journal of Structural Biology 175 (2011) 97–103

Contents lists available at ScienceDirect

Journal of Structural Biology

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

Structure Report

Crystal structure of the 30 K protein from the silkworm *Bombyx mori* reveals a new member of the β-trefoil superfamily

Jie-Pin Yang, Xiao-Xiao Ma, Yong-Xing He, Wei-Fang Li, Yan Kang, Rui Bao, Yuxing Chen *, Cong-Zhao Zhou *

Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei Anhui 230027, People's Republic of China

ARTICLE INFO

Article history:
Received 28 December 2010
Received in revised form 7 April 2011
Accepted 9 April 2011
Available online 14 April 2011

Keywords:
Silkworm
Crystal structure
30 K proteins
β-Trefoil superfamily
Sugar-binding site

ABSTRACT

The hemolymph of the fifth instar larvae of the silkworm *Bombyx mori* contains a group of homologous proteins with a molecular weight of approximately 30 kDa, termed *B. mori* low molecular weight lipoproteins (Bmlps), which account for about 5% of the total plasma proteins. These so-called "30 K proteins" have been reported to be involved in the innate immune response and transportation of lipid and/or sugar. To elucidate their molecular functions, we determined the crystal structure of a 30 K protein, Bmlp7, at 1.91 Å. It has two distinct domains: an all-α N-terminal domain (NTD) and an all-β C-terminal domain (CTD) of the β-trefoil fold. Comparative structural analysis indicates that Bmlp7 represents a new family, adding to the 14 families currently identified, of the β-trefoil superfamily. Structural comparison and simulation suggest that the NTD has a putative lipid-binding cavity, whereas the CTD has a potential sugar-binding site. However, we were unable to detect the binding of either lipid or sugar. Therefore, further investigations are needed to characterize the molecular function of this protein.

1. Introduction

The silkworm *Bombyx mori* has been domesticated in China as a major economic insect for silk production for 5000 years. Similar to other *Lepidoptera*, the silkworm has an open circulatory hemolymph system (Vierstraete et al., 2003). The silkworm "30 K" proteins are a group of plasma proteins with a molecular weight of approximately 30 kDa, which are synthesized in fat bodies and secreted into the hemolymph in the later stage of the fourth through to the fifth instar larvae (Mori et al., 1991). 30 K proteins are transported from the hemolymph back to the fat body cells during the larval-pupal transformation (Mine et al., 1983). In sexually mature female moths, these proteins are transported from the fat bodies to yolk granules, where they become the second major yolk protein of the eggs, after vitellin (Maki and Yamashita, 1997). During embryonic development, the 30 K proteins are detectable and eventually vanished after the larval hatching (Zhong et al., 2005).

30 K proteins were previously identified as lipoproteins (Gamo, 1978). However, the lipid content of these proteins and their functional relationships to human lipoproteins were unknown (Chapman, 1980). 30 K proteins had also been shown to bind to glucose, maltose and glucans, suggesting their involvement in insect self-defense systems (Ujita et al., 2005). In addition, 30 K proteins had been found to exhibit anti-apoptotic activity (Kim et al., 2001) and to be involved in translocating chymotrypsin inhibitor-8 to the membrane of the midgut (Ueno et al., 2006). Compared with the classic lipoproteins in the insect hemolymph or human serum, the 30 K proteins exhibit a lower molecular weight, and might, at most, represent protein–fatty acid complexes (Chapman, 1980); therefore, they were classified into the lepidopteran low molecular weight lipoproteins family (Lipoprotein_11 family, PF03260, http://pfam.sanger.ac.uk/).

Based on the analysis of the primary sequences deduced from the genomic sequences (Xia et al., 2004) and expressed sequence tags (ESTs), ten 30 K proteins (Bmlp1–10) of the silkworm were identified and grouped into three subfamilies (Sun et al., 2007). We overexpressed five of them (Bmlp1–3, Bmlp7 and Bmlp8), and determined the crystal structure of Bmlp7, which is the first structure in the lipoprotein_11 family. It reveals a novel β-trefoil family, in addition to the current 14 families of the β-trefoil superfamily (http://pfam.sanger.ac.uk/), which are proposed to arise from a common ancestor, despite sharing no characteristic binding sites or subcellular localization motif (Hazes, 1996). Bmlp7 has an all-α N-terminal domain (NTD) fused to a all-β trefoil C-terminal domain (CTD). Structural analysis in combination with simulation indicated a putative lipid-binding cavity at the NTD and one potential sugar-binding site at the CTD; however, the *in vitro* binding for neither lipid nor sugar was detected.

2. Expression, purification, crystallization and data collection

Considering an N-terminal signal peptide (residues from 1–45) was predicted by SignalP 3.0 server, we cloned the truncated *Bmlp7*

* Corresponding authors. Fax: +86 551 3600406.
E-mail addresses: cyxing@ustc.edu.cn (Y. Chen), zcz@ustc.edu.cn (C.-Z. Zhou).

1047-8477/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2011.04.003

(Silkworm Knowledgebase, http://silkdb.genomics.org.cn/silkworm/index.jsp, Accession code: Bmb021422, residues from 46 to 284) into the *Nco*I/*Not*I digested pET28b expression vector with a hexahistidine tag (AAALEHHHHHH) at the C-terminus. The non-labeled protein was expressed in *Escherichia coli* BL-21 RIL (DE3) with LB medium. The seleno-methionine (Se-Met) labeled Bmlp7 was expressed in *E. coli* B834 (DE3) with Se-Met medium (25 mg/L L-Se-Met and 50 mg/L essential amino acids with M9 medium). After reaching an OD_{600nm} of 0.6–0.8, isopropyl-β-d-1-thiogalactopyranoside was added to a final concentration of 0.2 mM, and the cells continued growing for another 4 h at 37 °C. The cells were harvested by centrifugation and resuspended in 5 ml of 20 mM Tris–Cl, pH 8.0, 200 mM NaCl (Buffer A).

After sonication and centrifugation, the supernatant was loaded on a 0.5 ml Ni–NTA column and eluted with Buffer A containing 20 mM imidazole to remove the contaminants. The target protein was eluted with Buffer A containing 200 mM imidazole and further purified by gel filtration with a HiLoad 16/60 Superdex 75 column, which equilibrated with 20 mM Tris–Cl, pH 8.0, 50 mM NaCl, 14 mM β-mercaptoethanol. Crystallization trials were carried out at 16 °C by the hanging drop vapor diffusion method, with the initial condition of mixing 0.8 μl protein at 7 mg/ml with an equal volume of the reservoir solution. Se-Met Bmlp7 crystals were obtained in 0.05 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, pH 7.5, 0.2 M KCl, 28% pentaerythritol propoxylate (17/8, PO/OH). The non-labeled Bmlp7 crystals were obtained by mixing the protein at 12.9 mg/ml with 1.8 M ammonium sulfate, 0.1 M Tris–Cl, pH 8.5, 20 mM DL-dithiothreitol, 1.72 mM palmitic acid, and DMSO 1.72%.

The crystals were flash-cooled in liquid nitrogen and subjected to data collection. The multi-wavelength anomalous diffraction (MAD) dataset of the Se-Met Bmlp7 crystal at 2.60 Å was collected with radiation wavelengths of 0.9788 and 0.9793 Å at the Beijing Synchrotron Radiation Facility (BSRF). The dataset of the non-labeled Bmlp7 at 1.91 Å was collected with radiation wavelength of 0.9795 Å at the Shanghai Synchrotron Radiation Facility (SSRF). All diffraction data were processed with HKL2000 (Otwinowski and Minor, 1997).

3. Structure determination

Before overexpressing Bmlp7 in *E. coli*, we have successfully purified the natural Bmlp7 from silkworm hemolymph for crystallization and collected the diffraction data at 1.60 Å at BSRF. However, the dataset of natural Bmlp7 crystal was rather poor (results not shown). Therefore, we cloned eight of the ten 30 K genes, purified five of them (those encoding Bmlp1–3, Bmlp7 and Bmlp8) and obtained the crystal of Bmlp7. After optimization, the crystal of non-labeled Bmlp7 was diffracted to 1.91 Å at SSRF and the Se-Met derivative crystal was diffracted to 2.60 Å at BSRF.

The Bmlp7 diffraction was phased against the MAD data of the Se-Met-substituted protein to a maximum resolution of 2.60 Å (*C*2 space group). Two out of six methionine positions were identified with the SHELX program suite (Sheldrick, 2008). The phase was calculated and further improved with the program SOLVE/RESOLVE (Terwilliger and Berendzen, 1999). Automatic model building with the ARP/wARP program (Perrakis et al., 1997) was carried out against the MAD data from 25.00 Å to 2.60 Å. Then the initial model was used for the molecular replacement process of the non-labeled dataset at 1.91 Å (*P*1 space group) with MOLREP (Vagin and Teplyakov, 2010) in CCP4 (Collaborative Computational Project No. 4, 1994). Refinement was carried out using the maximum likelihood method implemented in REFMAC (Murshudov et al., 1997) and the interactive rebuilding process in COOT (Emsley and Cowtan, 2004). The overall assessment of model quality was checked with MOLPROBITY (Most favored: 98.9%) (Lovell

Table 1
Data collection, phasing and refinement statistics.

Data collection	MAD Peak	MAD Edge	Non-labeled
Wavelength (Å)	0.9788	0.9793	0.9795
Space group	C2	C2	P1
Unit cell (Å, °)	a = 59.96, b = 56.86 c = 70.73, $\alpha = \gamma$ = 90.00 β = 114.23	a = 59.91, b = 56.83 c = 70.68, $\alpha = \gamma$ = 90.00 β = 114.23	a = 41.86, b = 49.32, c = 54.79, α = 93.26 β = 94.04, γ = 103.47
Resolution limit (Å)	25.00–2.60 (2.69–2.60)	25.00–2.60 (2.69–2.60)	50.00–1.91 (1.94–1.91)
Unique Reflections	6756 (643)	6742 (646)	31,866 (1572)
[a]Completeness (%)	99.9 (98.6)	99.8 (98.3)	96.8 (95.9)
[b]R_{merge}	6.1 (12.3)	5.4 (12.4)	6.9 (19.4)
$I/\sigma(I)$	45.5 (18.0)	38.2 (15.2)	8.3 (4.8)
Redundancy	6.1	6.5	2.7
Heavy atom sites	2	2	
Refinement			
Resolution limit (Å)			50.00–1.91 (1.96–1.91)
[c]R_{factor} (%)			19.3
[d]R_{free} (%)			23.5
[e]RMSD bond length (Å)			0.010
RMSD bond angles (°)			1.186
Average B factor (Å^2)			25.83
[f]Ramachandran plot			
Most favored (%)			98.9
Additional allowed (%)			1.1
PDB entry			3PUB

[a] The values in parentheses refer to statistics in the highest bin.
[b] $R_{merge} = \sum_{hkl}\sum_i|I_i(hkl) - < I(hkl)/\sum_{hkl}\sum_i I(m)$, where $I_i(hkl)$ is the intensity of an observation and <I(hkl)> is the mean value for its unique reflection; Summations are over all reflections.
[c] $R_{factor} = \sum_h|Fo(h) - Fc(h)|/\sum_h Fo(h)$ where Fo and Fc are the observed and calculated structure-factor amplitudes, respectively.
[d] R_{free} was calculated with 5% of the data excluded from the refinement.
[e] Root-mean square-deviation from ideal values.
[f] Categories were defined by Molprobity.

et al., 2003). The final atomic coordinates and structure factors were deposited in the Protein Data Bank under the accession code of 3PUB (R_{factor} = 19.3%, R_{free} = 23.5%). The crystallographic parameters of the three datasets are listed in Table 1. All structure figures were prepared with PyMOL (DeLano, 2008). In the later sections, "Bmlp7" would indicate the non-labeled Bmlp7 (*P*1 space group), if not specially annotated.

4. Overall structure of Bmlp7

Each asymmetric unit of the Bmlp7 crystal contains two identical molecules with a root mean square deviation (RMSD) of 0.31 Å over 229 Cα atoms. We could trace the electron density from Val50 to Phe284 for chain A, and Val50 to Ala285 for chain B. The dimeric interface in the asymmetric unit predominantly consists of electrostatic interactions via an interface adding up to 1600 Å^2. This dimeric form could be observed in both crystals of space group *P*1 and *C*2. Results from PISA server (Krissinel and Henrick, 2007) showed this dimeric form is most likely an aggregation pattern. However, we did not observe the peak corresponding to the dimer during gel filtration, thus the CTD-CTD dimeric interface might be the result of crystal packing rather than specific dimerization.

Bmlp7 folds into two distinct domains: an all-α NTD (Ala46–Asn133) and an all-β CTD (Ala134–Phe284) (Fig. 1). However, residues 46–49 and those after Phe284 are disordered that they could not be traced from the electron density map. The two domains are mainly interacted via hydrophobic interactions with an interface adding up to 2000 Å^2. The NTD contains six helices, arranged as a right-handed superhelix, forming a well-defined hydrophobic cavity at the center. The overall structure of the CTD is a β-trefoil barrel composed of 12 anti-parallel β-strands that are arranged in six two-stranded hairpin-shaped segments, three of which form the barrel sidewall, whereas the remaining three form a triangular cap. These segments fold into three structural repeats (CTD_α, β and γ), each of which consists of four β-strands, with the second and third strands separated by a β-hairpin and the other two connected by loops of variable length.

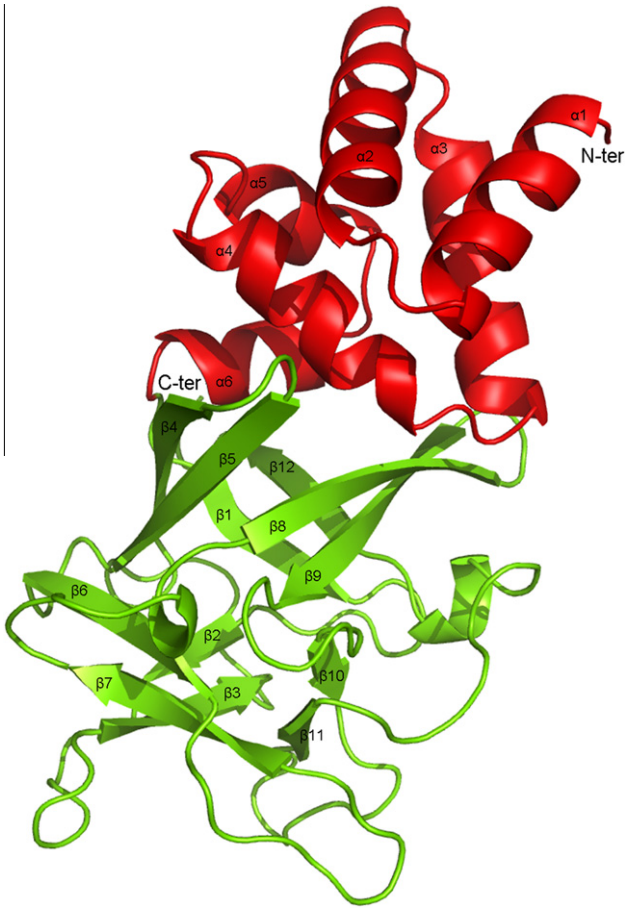

Fig. 1. Overall structure of Bmlp7. The NTD (red) contains six helices, arranged as a right-handed superhelix. The CTD (green) is a typical β-trefoil barrel. The two domains are mainly linked via hydrophobic interactions.

5. The all-α NTD and the putative lipid binding cavity

Structural comparison using the DALI server (Holm and Sander, 1998) showed that, although the arrangement of the NTD of Bmlp7 is similar to that of a monomer of the cleavage and polyadenylation factor Pcf11 (Meinhart and Cramer, 2004) from *Saccharomyces cerevisiae* (PDB: 1SZA, RMSD = 2.20 Å, over 69 Cα atoms), residues involved in hydrophobic-core packing or binding sites are totally different. Another hit of the DALI output is the barley lipid transfer protein (Lerche and Poulsen, 1998) (LTP, PDB: 1BE2, RMSD = 2.50 Å, over 60 Cα atoms) which has a palmitic acid in its U-shaped hydrophobic cavity. A deep cavity at the center of the NTD of Bmlp7 was also revealed by the CAVER program (Petrek et al., 2006) (Fig. 2A). The surface of this cavity is made up of residues Leu59, Ile63, Ala71, Leu78, Val86, Ile87, Val90, Val91, Leu94, Tyr105, Leu109, Ile117 and Phe122. Outlets of this hydrophobic cavity are sealed by residues Tyr60, Val64, Lys99, Met100 and Asn101 on one side, and by Leu55, Gln58, Lys74, His77 and Lys83 on the other side (Fig. 2A). Despite the structural similarity between Bmlp7 and LTP, the loops connecting helices α1, α2 and α3 of Bmlp7 are shorter than those in LTP. Moreover, two short helices, α5 and α6, cover the entrance of the hydrophobic cavity, whereas the same place is held by a flexible loop in LTP (Fig. 2B). We have tried to obtain the Bmlp7–palmitic acid complex, or to detect their interaction by isothermal titration calorimetry (ITC), but we failed. Calculation showed that the volume of palmitic acid is 293.75 Å^3 (Generate 3D server), while the cavity volume of Bmlp7 NTD is only about 234 Å^3 (CASTp server) which is too small to accommodate a palmitic acid molecule. Although 30 K family proteins were reported as lipoproteins, the natural lipid associated with these proteins is unknown. The failure to detect palmitic acid binding may imply Bmlp7 has a different lipid ligand.

6. The all-β CTD and the putative sugar-binding sites

In addition to the all-α NTD, Bmlp7 has an all-β CTD which belongs to the β-trefoil fold. Several hydrophobic residues (Leu137, Leu146, Trp174, Ile189, Leu198, Trp225, Ile240, Leu249 and Trp279) form the hydrophobic core of the β-trefoil barrel (Fig. 3A). According to the Pfam database, the β-trefoil fold has been found in 14 families (CL0066, http://pfam.sanger.ac.uk/), including fibroblast growth factors, Kunitz soybean trypsin inhibitors, and ricin-like toxins. The relationship between these 14 families can be defined by the similarity of their sequences or structures or by profile-hidden Markov models. The previous structures of β-trefoil family show that the lectin domain tends to fuse to a wide variety of functionally and/or structurally unrelated domains, thus forming a new family. Bmlp7 is the first example of the 15th β-trefoil fold family.

To find some clues to the function of this domain, we performed a structural homology search with DALI server and found that all of the top 40 hits were β-trefoil proteins. The ricin B-type domain of mosquitocidal holotoxin (Hazes and Read, 1995; Treiber et al., 2008) (PDB: 2VSE, Z-score = 17.6) is the most similar to the CTD

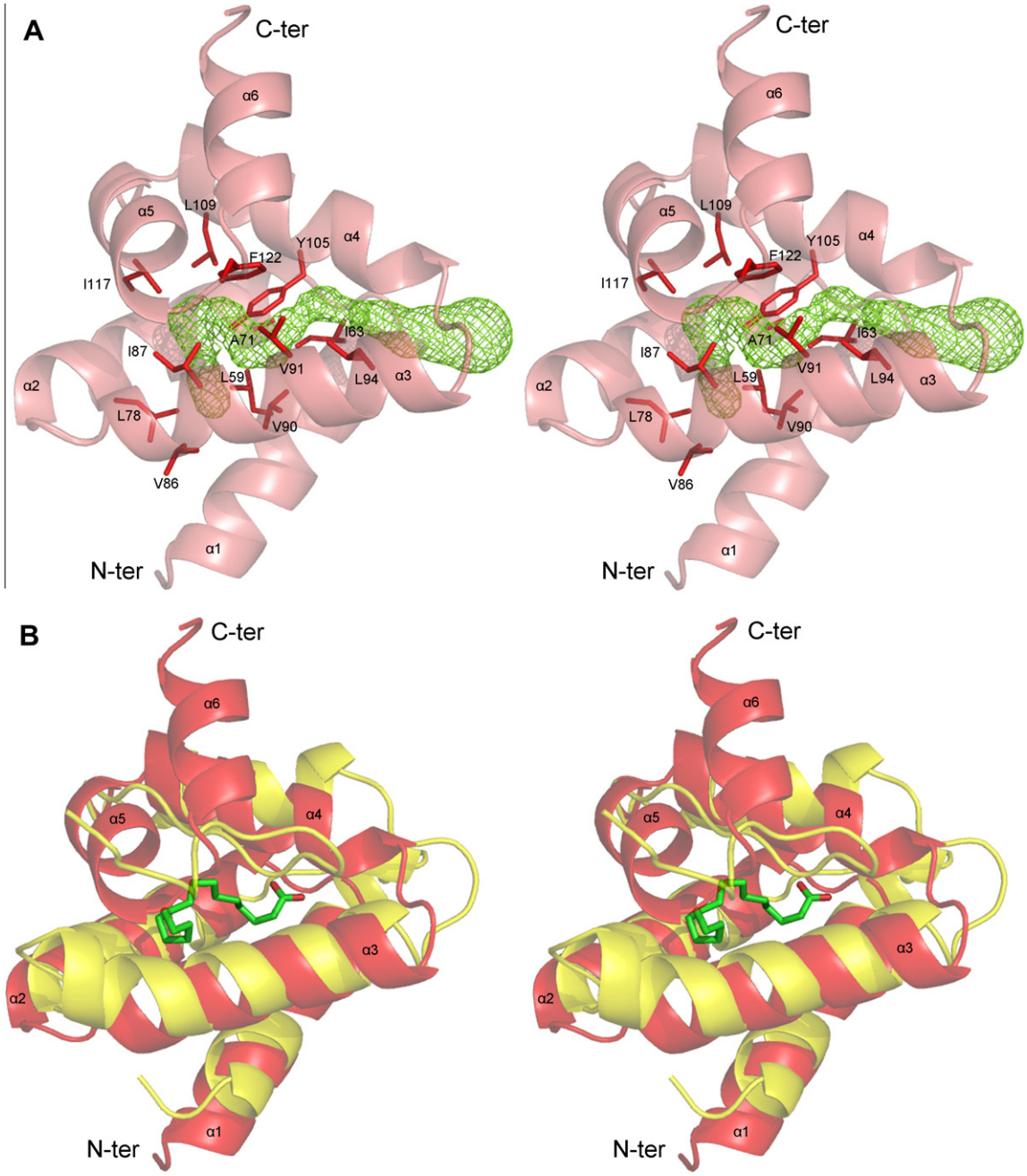

Fig. 2. Structural analysis of the NTD of Bmlp7. (A) The six helices from NTD form a well-defined hydrophobic cavity at the center. (B) Superposition of the Bmlp7 NTD (red) and the barley LTP (yellow) shows the cavity in the two structures are similar. However, the C-terminus of LTP is a flexible loop, while that of NTD is composed of two small helices. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

of Bmlp7. Despite a relatively low sequence identity (17%), the two structures could be readily superimposed, giving an RMSD of 2.60 Å over 137 Cα atoms. Moreover, the key residues at the hydrophobic core are highly conserved (Fig. 3A).

In these 40 β-trefoil structures, we found 9 proteins with a sugar at the their sugar-binding sites (Fujimoto et al., 2002; Grahn et al., 2007; Mancheno et al., 2005; Maveyraud et al., 2009; Nakamura et al., 2008; Notenboom et al., 2002; Rutenber et al., 1991; Suzuki et al., 2009). The sugar molecule could bind to the protein via π-π stacking interactions with the aromatic residue located at β3 in each β-repeat. The sugar molecule is also stabilized by hydrogen bonds between the hydroxyl moiety of the sugar and several polar residues around. The alignment of the three β-repeats of Bmlp7 and the sugar-binding sites of the above nine proteins revealed the relatively conserved sugar-binding residues are Trp/Tyr/His, all of which could offer π-π stacking interactions (Fig. 3B). For example, the lactose in the structure of plant cytotoxin ricin (Rutenber et al., 1991) (PDB: 2AAI, Z-score = 13.2, RMSD = 2.60 Å over 122 Cα atoms, compared to Bmlp7) was stabilized by π-π interaction to Tyr248, together with several hydrogen bonds. Superposition revealed that Tyr266 from the CTD_γ repeat of Bmlp7 is the counterpart of Tyr248 of the plant cytotoxin ricin (Fig. 3C). However, corresponding residues in the CTD_α and CTD_β repeats of Bmlp7 are Gly162 and Val215, which could not offer the π-π interactions (Fig. 3B). Thus we presumed that only the CTD_γ repeat of Bmlp7 has the potential for sugar binding.

In order to check the sugar-binding capacity of Bmlp7, we tested two major sugar molecules in silkworm hemolymph,

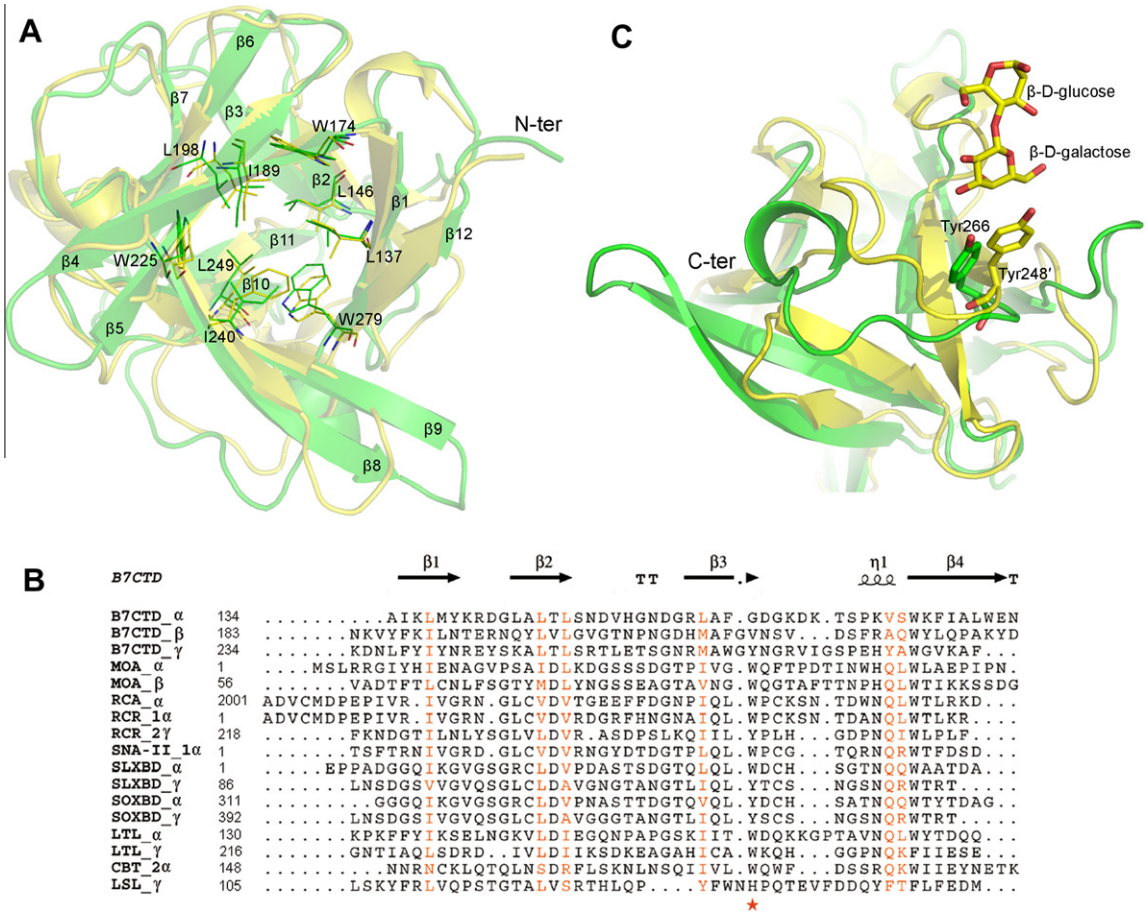

Fig. 3. Structural comparison of the CTD of Bmlp7. (A) Superposition of the CTD (green) against the mosquitocidal holotoxin (yellow). The key residues at the hydrophobic core are highly conserved. (B) Multiple-sequence alignment of the three β-repeats of Bmlp7 and sugar-binding motifs of the β-trefoil proteins. The three β-repeats of Bmlp7 are termed B7CTD_α, β and γ, respectively. The other β-trefoil proteins include: MOA (*Marasmius oreades* agglutinin); RCA (*Ricinus communis* agglutinin); RCR (*R. communis* cytotoxin ricin); SNA-II (galactose/*N*-acetylgalactosamine-specific lectin of the bark of elderberry *Sambucus nigra*); SLXBD (xylan-binding domain of *Streptomyces lividans*); SOXBD (xylan-binding domain of *Streptomyces olivaceoviridis* E-86); LTL (galactose-binding lectin EW29 from the earthworm *Lumbricus terrestris*); CBT (*Clostridium botulinum* type C 16S progenitor toxin); and LSL (lectin from the parasitic mushroom *Laetiporus sulphureus*). The sugar-binding residues are marked. The residue number at the N-terminus of each β-repeat is indicated. (C) Superposition of the third β-repeat (green) of Bmlp7 CTD against the cytotoxin ricin of *R. communis* (yellow). Tyr266 of Bmlp7 and Tyr248′ of *R. communis* ricin are putative residues for sugar binding. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

trehalose and glucose. However, we were unable to detect the binding of d-trehalose or d-glucose (both at a maximum concentration of 2 mM) to Bmlp7, either using ITC or fluorescence spectrometry. We supposed that a sugar of different type may be the natural ligand for Bmlp7.

7. All 30K proteins have a similar organization

Based on the complete genome sequence draft of the silkworm (Xia et al., 2004) and ESTs, Sun et al. systematically analyzed and renamed ten isoforms of 30 K proteins which shared a sequence similarity above 85% (Sun et al., 2007). In order to further analyze the 30 K proteins, we submitted the sequence of Bmlp7 to BLAST (Altschul et al., 1990) against the non-redundant protein sequences database. Among the 32 top hits, 29 are silkworm proteins, whereas three proteins are from two other species (*Pseudaletia separate* and *Manduca sexta*). All 32 proteins belong to the lipoprotein_11 family. After removing 19 sequences, which have a sequence identity of more than 85%, we selected 13 proteins with an identity from 83% to 29% to perform the alignment. The alignment showed that these 13 sequences are relatively conserved, especially the residues forming the NTD hydrophobic cavity (Fig. 4A). Moreover, these proteins have an organization similar to that of Bmlp7: a signal peptide and an all-α NTD followed by an all-β CTD (Fig. 4B).

8. Putative structure–function relationship of 30K proteins

In all innate immune response systems, the most important step is that pattern-recognition proteins bind sugar molecules (e.g. β-1,3-glucan) on the surfaces of invading pathogens (Yu et al., 2002). A recombinant *B. mori* 30 K protein 6G1 (Bmlp1) was reported to bind glucans and participate in the activation of prophenoloxidase cascade. Also, Bmlp1 could interfere with hyphal growth of the entomopathogenic fungus *Paecilomyces tenuipes* (Ujita et al., 2005). Recently, several anti-apoptotic fractions from *B. mori* hemolymph were obtained. The one exhibiting the highest activity was a 30 K protein (Choi et al., 2006). These results suggest that 30 K proteins probably play a crucial role in the protection of *B. mori* against invading pathogens. Considering the similar primary sequence and organization of 30 K proteins, we proposed that the putative sugar-binding site may be the basis for its immune related function. The CTD might bind some special sugar molecules on the surface of invading pathogens. Then the NTD probably would bind some special lipids or downstream protein factors, which could activate the immune system. Various isoforms

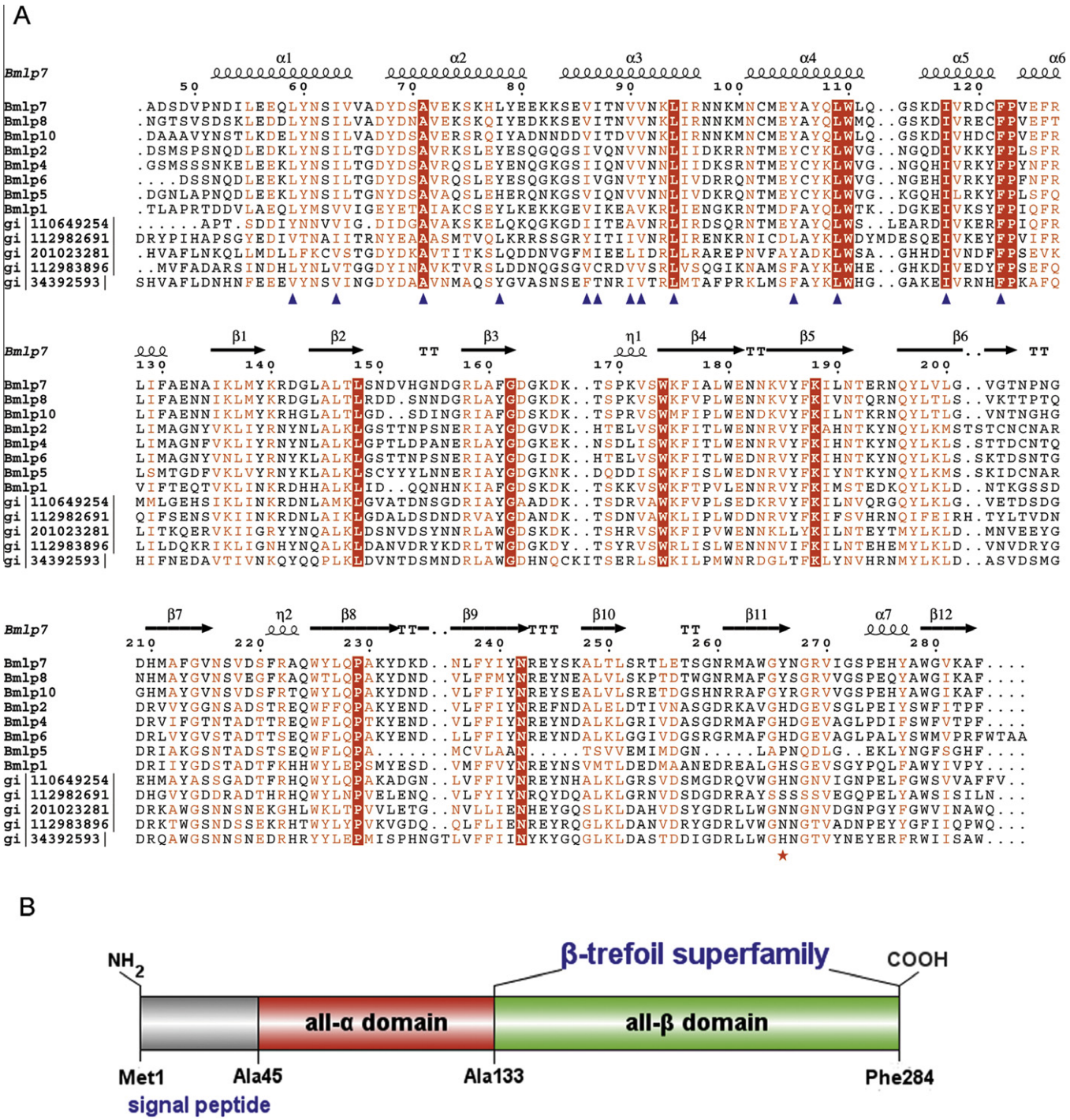

Fig. 4. Multiple-sequence alignment and organization of 30 K proteins. (A) Multiple-sequence alignment of 30 K proteins. We selected 13 proteins with an identity from 83% to 29% to perform the alignment (gi|110649254| from *M. Sexta*; gi|34392593| from *P. separate*; others all from *B. mori*). These 13 sequences are relatively conserved, especially the hydrophobic cavity-forming residues (blue triangles) and the putative sugar-binding residue (red star). (B) Organization of 30 K proteins. They all consist of a signal peptide, an all-α NTD, and an all-β CTD. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

of 30 K proteins might be responsible for diverse pathogens, thanks to different sugar molecules binding to their surfaces. However, more investigations are needed to define the interacting partners of NTD and CTD of Bmlp7 or other 30 K proteins.

Acknowledgements

We would like to thank Dr. Sheng-Peng Wang at the Sericultural Research Institute, Chinese Academy of Agricultural Sciences for offering the cDNA of *B. mori* P50 strain. We also would like to thank the staff at BSRF and SSRF for their help. This work was supported by the National Natural Science Foundation of China (Program 90608027) and the Ministry of Science and Technology of China (Projects 2005CB121002 and 2006AA10A119).

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Chapman, M.J., 1980. Animal lipoproteins: chemistry, structure, and comparative aspects. J. Lipid Res. 21, 789–853.
Choi, S.S., Rhee, W.J., Kim, E.J., Park, T.H., 2006. Enhancement of recombinant protein production in chinese hamster ovary cells through anti-apoptosis engineering using *30Kc6* gene. Biotechnol. Bioeng. 95, 459–467.
Collaborative Computational Project, No. 4, 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. 50, 760-3.
DeLano, W.L., 2008. The PyMOL Molecular Graphics System. DeLano Scientific LLC, Palo Alto, CA, USA. http://www.pymol.org.
Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132.
Fujimoto, Z., Kuno, A., Kaneko, S., Kobayashi, H., Kusakabe, I., et al., 2002. Crystal structures of the sugar complexes of *Streptomyces olivaceoviridis* E-86 xylanase: sugar binding structure of the family 13 carbohydrate binding module. J. Mol. Biol. 316, 65–78.
Gamo, T., 1978. Low molecular weight lipoproteins in the haemolymph of the silkworm, *Bombyx mori*: Inheritance, isolation and some properties. Insect Biochem. 8, 457–470.
Grahn, E., Askarieh, G., Holmner, A., Tateno, H., Winter, H.C., et al., 2007. Crystal structure of the *Marasmius oreades* mushroom lectin in complex with a xenotransplantation epitope. J. Mol. Biol. 369, 710–721.
Hazes, B., 1996. The $(QxW)_3$ domain: a flexible lectin scaffold. Protein Sci. 5, 1490–1501.
Hazes, B., Read, R.J., 1995. A mosquitocidal toxin with a ricin-like cell-binding domain. Nat. Struct. Biol. 2, 358–359.
Holm, L., Sander, C., 1998. Touring protein fold space with Dali/FSSP. Nucleic Acids Res. 26, 316–319.
Kim, E.J., Rhee, W.J., Park, T.H., 2001. Isolation and characterization of an apoptosis-inhibiting component from the hemolymph of *Bombyx mori*. Biochem. Biophys. Res. Commun. 285, 224–228.
Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797.
Lerche, M.H., Poulsen, F.M., 1998. Solution structure of barley lipid transfer protein complexed with palmitate. Two different binding modes of palmitate in the homologous maize and barley nonspecific lipid transfer proteins. Protein Sci. 7, 2490–2498.
Lovell, S.C., Davis, I.W., Arendall 3rd, W.B., de Bakker, P.I., Word, J.M., et al., 2003. Structure validation by Calpha geometry: phi, psi and Cbeta deviation. Proteins 50, 437–450.
Maki, N., Yamashita, O., 1997. Purification and characterization of a protease degrading 30 kDa yolk proteins of the silkworm, *Bombyx mori*. Insect Biochem. Mol. Biol. 27, 721–728.
Mancheno, J.M., Tateno, H., Goldstein, I.J., Martinez-Ripoll, M., Hermoso, J.A., 2005. Structural analysis of the *Laetiporus sulphureus* hemolytic pore-forming lectin in complex with sugars. J. Biol. Chem. 280, 17251–17259.
Maveyraud, L., Niwa, H., Guillet, V., Svergun, D.I., Konarev, P.V., et al., 2009. Structural basis for sugar recognition, including the Tn carcinoma antigen, by the lectin SNA-II from *Sambucus nigra*. Proteins 75, 89–103.
Meinhart, A., Cramer, P., 2004. Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. Nature 430, 223–226.
Mine, E., Izumi, S., Katsuki, M., Tomino, S., 1983. Developmental and sex-dependent regulation of storage protein synthesis in the silkworm, *Bombyx mori*. Dev. Biol. 97, 329–337.
Mori, S., Izumi, S., Tomino, S., 1991. Structures and organization of major plasma protein genes of the silkworm *Bombyx mori*. J. Mol. Biol. 218, 7–12.
Murshudov, G.N., Vagin, A.A., Dodson, E.J., 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. 53, 240–255.
Nakamura, T., Tonozuka, T., Ide, A., Yuzawa, T., Oguma, K., et al., 2008. Sugar-binding sites of the HA1 subcomponent of *Clostridium botulinum* type C progenitor toxin. J. Mol. Biol. 376, 854–867.
Notenboom, V., Boraston, A.B., Williams, S.J., Kilburn, D.G., Rose, D.R., 2002. High-resolution crystal structures of the lectin-like xylan binding domain from *Streptomyces lividans* xylanase 10A with bound substrates reveal a novel mode of xylan binding. Biochemistry 41, 4246–4254.
Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326.
Perrakis, A., Sixma, T.K., Wilson, K.S., Lamzin, V.S., 1997. WARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. Acta Crystallogr. D Biol. Crystallogr. 53, 448–455.
Petrek, M., Otyepka, M., Banas, P., Kosinova, P., Koca, J., et al., 2006. CAVER: a new tool to explore routes from protein clefts, pockets and cavities. BMC Bioinf. 7, 316.
Rutenber, E., Katzin, B.J., Ernst, S., Collins, E.J., Mlsna, D., et al., 1991. Crystallographic refinement of ricin to 2.5 A. Proteins 10, 240–250.
Sheldrick, G.M., 2008. A short history of SHELX. Acta Crystallogr. A 64, 112–122.
Sun, Q., Zhao, P., Lin, Y., Hou, Y., Xia, Q.Y., et al., 2007. Analysis of the structure and expression of the 30K protein genes in silkworm, *Bombyx mori*. Insect Sci. 14, 5–14.
Suzuki, R., Kuno, A., Hasegawa, T., Hirabayashi, J., Kasai, K.I., et al., 2009. Sugar-complex structures of the C-half domain of the galactose-binding lectin EW29 from the earthworm *Lumbricus terrestris*. Acta Crystallogr. D Biol. Crystallogr. 65, 49–57.
Terwilliger, T.C., Berendzen, J., 1999. Automated MAD and MIR structure solution. Acta Crystallogr. D Biol. Crystallogr. 55, 849–861.
Treiber, N., Reinert, D.J., Carpusca, I., Aktories, K., Schulz, G.E., 2008. Structure and mode of action of a mosquitocidal holotoxin. J. Mol. Biol. 381, 150–159.
Ueno, Y., He, N., Ujita, M., Yamamoto, K., Banno, Y., et al., 2006. Silkworm midgut proteins interacting with a hemolymph protease inhibitor, CI-8. Biosci. Biotechnol. Biochem. 70, 1557–1563.
Ujita, M., Katsuno, Y., Kawachi, I., Ueno, Y., Banno, Y., et al., 2005. Glucan-binding activity of silkworm 30-kDa apolipoprotein and its involvement in defense against fungal infection. Biosci. Biotechnol. Biochem. 69, 1178–1185.
Vagin, A., Teplyakov, A., 2010. Molecular replacement with MOLREP. Acta Crystallogr. D Biol. Crystallogr. 66, 22–25.
Vierstraete, E., Cerstiaens, A., Baggerman, G., Van den Bergh, G., De Loof, A., et al., 2003. Proteomics in *Drosophila melanogaster*: first 2D database of larval hemolymph proteins. Biochem. Biophys. Res. Commun. 304, 831–838.
Xia, Q., Zhou, Z., Lu, C., Cheng, D., Dai, F., et al., 2004. A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*). Science 306, 1937–1940.
Yu, X.Q., Zhu, Y.F., Ma, C., Fabrick, J.A., Kanost, M.R., 2002. Pattern recognition proteins in *Manduca sexta* plasma. Insect Biochem. Mol. Biol. 32, 1287–1293.
Zhong, B.X., Li, J.K., Lin, J.R., Liang, J.S., Su, S.K., et al., 2005. Possible effect of 30K proteins in embryonic development of silkworm *Bombyx mori*. Acta Biochim. Biophys. Sin. (Shanghai) 37, 355–361.