

Biotechnology Advances

Hookworm SCP/TAPS protein structure – a key to understanding host-parasite interactions and developing new interventions

Asiah Osman^a, Conan K. Wang^a, Anja Winter^{a,b}, Alex Loukas^{c,d}, Leon Tribolet^{c,d},
Robin B. Gasser^e, Andreas Hofmann^{a,d,e,*}

^a*Structural Chemistry Program, Eskitis Institute for Cell & Molecular Therapies, Griffith University, Brisbane, Queensland, Australia*

^b*Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom*

^c*James Cook University, Cairns, Queensland, Australia*

^d*Queensland Tropical Health Alliance*

^e*Department of Veterinary Science, The University of Melbourne, Victoria, Australia*

*Corresponding author: Telephone/Facsimile: +61-7-3735-4425; Email: a.hofmann@griffith.edu.au

ARTICLE INFO

Keywords:

SCP/TAPS proteins

Activation-associated secreted proteins

Ancylostoma- secreted proteins

Protein structure

Host-parasite interactions

Pathogenesis-related proteins

Note:

"Hinge" denotes the C-terminal moiety found in many SCP/TAPS proteins; "hinge" is used to describe a flexible region linking two domains or moieties.

ABSTRACT

SCP/TAPS proteins are a diverse family of molecules in eukaryotes, including parasites. Despite their abundant occurrence in parasite secretomes, very little is known about their functions in parasitic nematodes, including blood-feeding hookworms. Current information indicates that SCP/TAPS proteins (called *Ancylostoma*-secreted proteins, ASPs) of the canine hookworm, *Ancylostoma caninum*, represent at least three distinct groups of proteins. This information, combined with comparative modelling, indicates that all known ASPs have an equatorial groove that binds extended structures, such as peptides or glycans. To elucidate structure-function relationships, we explored the three-dimensional crystal structure of an ASP (called *Ac*-ASP-7), which is highly upregulated in expression in the transition of *A. caninum* larvae from a free-living to a parasitic stage. The topology of the N-terminal domain is consistent with pathogenesis-related proteins, and the C-terminal extension that resembles the fold of the Hinge domain. By anomalous diffraction, we identified a new metal binding site in the C-terminal extension of the protein. *Ac*-ASP-7 is in a monomer-dimer equilibrium, and crystal-packing analysis identified a dimeric structure which might resemble the homo-dimer in solution. The dimer interaction interface includes a novel binding site for divalent metal ions, and is proposed to serve as a binding site for proteins involved in the parasite-host interplay at the molecular level. Understanding this interplay and the integration of structural and functional data could lead to the design of new approaches for the control of parasitic diseases, with biotechnological outcomes.

1. Introduction

Soil-transmitted helminths (geohelminths) are causative agents of neglected tropical diseases (NTDs), mainly in developing countries (Hotez et al., 2007). In particular, blood-feeding hookworms, including species of *Necator* and *Ancylostoma* (Nematoda), infect ~740 million people in rural areas of the tropics and subtropics (de Silva et al., 2003), causing an estimated disease burden of 22 million disability-adjusted life years (Hotez et al., 2006). Hookworms infect humans by direct transmission of infective, third-stage (L3) larvae from the environment. These larvae usually penetrate human skin and, following migration via the circulatory system and lungs, develop through to adults, which reside in the duodenum. The dioecious adults attach to the intestinal mucosa, rupture capillaries and feed on blood. Consequently, hookworm disease is characterized by anaemia, and often leads to physical and mental retardation and sometimes death in children and adverse maternal-fetal outcomes (Bethony et al., 2006; Loukas et al., 2006). Although anthelmintic drugs are used to reduce the burden of disease in a range of countries, mass treatment programs carry a significant risk of inducing drug resistance in hookworm populations (Cantacessi et al., 2010). Therefore, in recent years, there has been a major focus on developing new approaches to control hookworm disease, based on improved knowledge of the pathogen, the pathogen-host interactions and the disease at the molecular and immunological levels. Often, the canine hookworm, *Ancylostoma caninum*, has been used as a model for this purpose.

Various studies have implicated activation-associated proteins (ASPs) in an immunomodulatory function during a hookworms' invasion of the host, migration through tissues, attachment to the intestinal wall and blood-feeding (reviewed by (Cantacessi et al., 2009)). These are amongst the ten most abundant groups of proteins in hookworms (Cantacessi et al., 2009). Moreover, during the transition of *A. caninum* from a free-living to parasitic L3, 17 of the 30 most highly up-regulated mRNAs encoded members of the ASP family (Datu et al., 2008). ASPs belong to a large group of proteins, the 'sperm-coating protein (SCP)-like extracellular proteins', also called SCP/Tpx-1/Ag5/PR-1/Sc7 (SCP/TAPS; Pfam accession number no. PF00188), characterized by the presence of a single or double 'SCP-like extracellular domain' (InterPro: IPR014044). In *A. caninum*, double and single SCP-domain ASPs, designated *Ac*-ASP-1 and *Ac*-ASP-2, respectively, are secreted in response to host-specific signals during the infection process (Hawdon et al., 1996; Bethony et al., 2005). ASPs also dominate the excretory/secretory (ES) proteins released by the blood-feeding adult stage hookworms (Mulvenna et al., 2009). Homologues of *Ac*-ASP-1 and *Ac*-ASP-2 have been identified also in the L3 stage of *N. americanus* (Zhan et al., 1999; Goud et al., 2004, 2005). Results from crystallography (Asojo et al., 2005), combined with the observation that one of these ASPs (*Na*-ASP-2) induces neutrophil and monocyte migration (Bower et al., 2008), suggest that this molecule has a role as an antagonistic ligand of complement receptor 3 (CR3) and alters the immune cascade by preventing the binding of chemotaxin (Asojo et al., 2005). Because of its immunogenic and protective properties, *Na*-ASP-2 is being explored as a vaccine candidate against human hookworm disease (Bethony et al., 2005; Loukas et al., 2006; Mendez et al., 2008; Xiao et al., 2008). In adult *A. caninum*, at least four other ASPs (named *Ac*-ASP-3, *Ac*-ASP-4, *Ac*-ASP-5 and *Ac*-ASP-6) have been identified to date (Zhan et al., 2003). Another ASP-like molecule, designated neutrophil inhibitory factor (NIF), has been isolated and shown to play an immunomodulatory role by blocking the adhesion of activated neutrophils to vascular endothelial cells and the subsequent release of H₂O₂ from activated neutrophils (Moyle et al., 1994). NIF further interferes with the function of integrin receptors located on the cell surface, resulting in the inhibition of the aggregation and adhesion of platelets (Del Valle et al., 2003).

Clearly, investigating the structure of ASPs has major implications for understanding the molecular interactions between hookworms and their mammalian hosts. The topology of proteins of this group comprises an N-terminal PR (or CAP) domain, a hinge region and a variable C-terminal extension domain, such as the ion channel regulator domain in CRISPs, C-type lectin domain or the

LCCL (= Limulus clotting factor C, Coch-5b2, and Lg11) domain (Gibbs et al., 2008). This latter C-terminal domain is found in ASPs of mammals but is absent from invertebrates (Gibbs et al., 2008). To date, six SCP/TAPS proteins (see Table S1) have been characterized at the structural level, revealing the common fold of the core domain, first described for the PR-1 protein P14a from plants (Fernández et al., 1997). The SCP/TAPS core consists of an α - β - α sandwich in which the three-stranded anti-parallel β -sheet is flanked by three helices on one side, and a fourth helix on the other. These three, stacked layers are held together at the interfaces by a hydrophobic core. With the exception of GAPR-1 (Gibbs et al., 2008), the fold of the core domain includes three conserved disulphide bonds. In spite of this information, the structural classification of ASPs is unknown. Here, we establish, for the first time, that there are at least three distinct groups based on structural characteristics. Using this classification system, we determine the three-dimensional crystal structure of an entirely novel ASP (designated *Ac*-ASP-7; belonging to “group 2”) that is highly upregulated upon transition of the *A. caninum* L3 stage to parasitism, and examine its structure-function relationships by comparison with the known structure of *Na*-ASP-2 (Asojo et al., 2005) representing “group 1”. In addition, we identify a metal binding site in the C-terminal extension of *Ac*-ASP-7 which might constitute a conserved feature of ASPs representing group 2.

2. Structure determination

Recombinant *Ac*-ASP-7 was produced as a secreted protein in the yeast *Pichia pastoris* as described for *Na*-ASP-2 (cf. Goud et al., 2005). Briefly, the open reading frame (ORF) of the mature protein (excluding the endogenous signal peptide, beginning at Leu-19) was cloned into pPIC-Z α , in frame with the cleavable N-terminal alpha-mating factor signal peptide encoded by the plasmid and C-terminal hexa-His tag. The secreted protein was purified from culture medium using nickel-NTA affinity chromatography and an AKTA UPC protein purification workstation (GE Health Care). Initial crystallization screens were carried out at 16 °C in sitting drops in 96-well MRC plates (Molecular Dimensions Ltd, Suffolk, UK) and our large in-house factorial collection (> 1000 pre-formulated conditions). An ideal concentration of 4.7 mg/ml was determined. First crystals were obtained using either of the following conditions: 10% PEG2000, 20 mM KH₂PO₄, pH 8.0 or 20% MPD, 0.1 M TRIS, pH 6.8. The conditions were then refined to 15-30% MPD, and 0.1 M bis-TRIS, pH 6.5 or 0.1 M TRIS pH 7.0 or 7.5 in hanging drops using 24-well plates. Crystals appeared after 10 days. The X-ray crystal structure of *Ac*-ASP-7 was determined using multiple isomorphous replacement with anomalous diffraction (MIRAS) using data from two derivatives (ethyl mercury phosphate, LaCl₃) and one native crystal. Data sets were indexed with Mosflm (Leslie, 1992), and scaling, truncation and analysis were performed using programs from the CCP4 suite (Collaborative Computational Project Number 4, 1994). The program autoSHARP was used to find heavy atom sites and conduct heavy atom refinement, density modification and solvent flattening. Data collection and phasing statistics are given in Table S2. Multiple rounds of manual model building with O (Jones et al., 1991) and Coot (Emsley and Cowton, 2004) were interspersed with computational refinement (rigid body refinement, positional and B-factor refinement) with moderate NCS restraints using Phenix (Adams et al., 2010). Structure refinement statistics are given in Table 1. The atomic coordinates and structure factors have been deposited with the Protein Data Bank (accession numbers 3s6s, 3s6u and 3s6v). Interface analysis was conducted using the PISA web server (Krissinel and Henrick, 2007). Figures were prepared with PyMOL (DeLano, 2002).

3. Classification of *A. caninum* ASPs

The ASP sequences inferred from transcriptomic datasets (Datu et al., 2008) were subjected to secondary structure predictions and used to assemble a structure-based sequence alignment,

including *Na*-ASP-2 from *N. americanus* (Fig. S1). Using the structure-based alignment, three different groups (1-3) of proteins could be distinguished. In general, dithioether bonds **1-3** are located in the SCP/TAPS protein core, whereas dithioether bonds **4** and **5** (if present) occur in the C-terminal moiety. In groups 1 and 3, all proteins possess five dithioether bonds, with the exception of *Ac_SSH_C_00143* (group 1) and *Ac_SSH_C_0065*, *Ac_SSH_C_0118* and *Ac_SSH_C_0158* (Group 3). Proteins in group 2 are distinguished from those of groups 1 and 3 by the absence of dithioether bond **2** from the core.

Group 1 includes *Na*-ASP-2, and the tandem-PR domain protein *Na*-ASP-1 (Asojo, 2011), and is characterised by the presence of two conserved histidine residues (His-69 and His-129; see Fig. S1), which may confer a metal binding site to the protein. Interestingly, the proteins VAP-1 and SCL-22 of *Caenorhabditis elegans* (GenBank accession nos. NP_001024553 and NP_509802, respectively) also belong to this group. While groups 1 and 3 appear to be closely related based on phylogenetic analysis (Cantacessi et al., 2009), the distinguishing criterion between both groups is the lack of the conserved histidine residues (Fig. S1) in ASPs of group 3. Based on current structural data for ASPs, these residues might be involved in metal-dependent protein-ligand interactions. The absence of these residues from ASPs of groups 2 and 3 indicates that these proteins are highly unlikely to engage this type of interaction. Apart from the conserved histidine residues, group 2 ASPs also lack a considerable number of amino acids (18 compared with *Na*-ASP-2) in the loop preceding the CAP1 motif. At the structural level, this deletion is expected to result in a different shape for ASPs of group 2 when compared with those of group 1.

The C-terminal moiety of these proteins (when present) is ~ 30 amino acids in length and typically starts with a G-X-P/V motif. Three-dimensional structures predicted by comparative modeling of ASPs for all groups (not shown) suggested an equatorial groove is present in all of these proteins. The shape and extent of this groove might allow the binding of extended structures, such as peptides or glycans, but there is presently no experimental evidence to support this hypothesis. Following the classification of hookworm ASPs, we selected *Ac_0042A* (renamed *Ac*-ASP-7; GenBank accession no JN038056), as a representative of group 2, for detailed structural analysis.

4. The structure of *Ac*-ASP-7

Despite the low amino acid sequence identities among members of the ASPs, the fold of the α - β - α sandwich, as first described for P14a (Fernández et al., 1997), is highly conserved. In *Ac*-ASP-7, the N-terminal 150 residues closely resemble those in this fold, including two of the conserved disulphide bonds linking the very N-terminal end to helix $\alpha 2$ and beta-strands $\beta 5$ and $\beta 6$ (see Fig. 1). The third conserved disulphide bond within the α - β - α sandwich is absent from ASPs of group 2.

At the C-terminal end, *Ac*-ASP-7 possesses a moiety of 27 amino acids, which is characterized by the cysteine pattern C-X₅-C-X₇-C-X₄-C of the Hinge-like sequence of CRISPs (Gibbs et al., 2008). The Hinge moiety forms intimate contacts with the α - β - α sandwich domain and all four cysteine residues are engaged in disulphide bonds within the Hinge itself. Structurally, this C-terminal extension moiety consists of a helical turn, a β -hairpin and second helical turn, all connected by loops.

In agreement with the classification criteria for group 2, the area of the putative binding site around the conserved histidine residues in ASPs of group 1 is re-arranged, due to the deletion of residues between helices $\alpha 4$ and $\alpha 5$. The position of His-129 in *Na*-ASP-2 (group 1) is occupied by Asp-102 and Ser-105 in *Ac*-ASP-7. The side chain of Arg-79 in *Ac*-ASP-7 fills approximately the space occupied by His-69 and Glu-80 in *Na*-ASP-2 (see Fig. 1).

4.1. Divalent metal binding site in the C-terminal extension

The structure of *Ac*-ASP-7 shows a conformation in the C-terminal Hinge-like moiety that is reminiscent of binding sites for divalent metal ions. To evaluate this hypothesis, *Ac*-ASP-7 was co-crystallized with the divalent metal ions magnesium, calcium and manganese. Although we did not observe anomalous density for magnesium, a significant peak was detected at the putative binding site at Glu-171 for calcium and manganese (Fig. 2). The metals are coordinated by the backbone carbonyls CO-170 and CO-171 as well as the bidentate side chain of Glu-171 as η^1 -ligands. Notably, the metals are only bound in one of the two molecules within the asymmetric unit, since the side chain O-atom of residue Asn-7 of a symmetry-related molecule also coordinates the metal. This interaction is not possible at the second monomer within the asymmetric unit, because of the different packing environment. The metal binding capability of *Ac*-ASP-7 was assessed in solution using isothermal titration calorimetry. For each of three metals tested (MgSO_4 , CaCl_2 , MnCl_2), although no significant binding could be detected, the affinity for these metals might increase in the presence of a protein binding partner (see section 5).

4.2. Quaternary structure

Results from size exclusion chromatography revealed that *Ac*-ASP-7 exists as a mixture of a monomeric (~75%) and dimeric (~25%) species at a neutral pH (Fig. 3). The molecular masses determined by MALS were $M = 22520 \text{ g mol}^{-1}$ and 41130 g mol^{-1} for the monomer and dimer species, respectively. These figures are in excellent agreement with the theoretically expected masses of 23479 g mol^{-1} and 46958 g mol^{-1} .

The crystal structure shows five possible dimeric assemblies of *Ac*-ASP-7, three of which possess an interface area of at least 350 \AA^2 (see Table 2, Fig. 4). The interface of dimer 3 (P2-symmetric) comprises the hydrophobic residues Leu-66 and Leu-72 (helices α_2 and α_3) in both molecules, and has an interface area of 350 \AA^2 . The non-symmetric dimer 2 possesses a 377 \AA^2 interface, which includes residues from helix α_4 (molecule A) and helices α_1 and α_4 (molecule B). The interface of this dimer features four polar interactions, including a salt bridge between Asp-102 and Arg-16. The largest interface area is found in the P2-symmetric dimer 1 (463 \AA^2), and this assembly is held together mainly by hydrophobic interactions between Trp-172 and Phe-178 (beta-strand β_{10}) from molecule A as well as Phe-178 and Phe-179 (beta-strand β_{10}) from molecule B. Additionally, Glu-171 (β_9) in molecule A engages in polar interactions with Asn-7 (β_1) and Arg-116 (β_5) from molecule B, thus harbouring the divalent metal ion within the groove constituted by the interfacing of the two molecules. In this dimer, the aforementioned hydrophobic residues are arranged with Ile-3, Val-5 and Gly-167 in a ring that is enclosed in the dimer interface and, thus, excluded from "unfavourable" contact with hydrophilic solvent. The ring also features a cluster of negative surface potential on one side (Asn-157, Glu-158, Glu-171 and several backbone carbonyl groups) and has a small cavity in its centre. The bottom of the cavity contains the aromatic ring of the side chain of Tyr-176.

5. Concluding remarks and implications

Using information on the primary and predicted secondary structures of novel hookworm SCP/TAPS proteins, at least three distinct groups of ASPs can be classified. A recent study of 28 SCP/TAPS proteins from the human blood fluke *Schistosoma mansoni* (Trematoda) (Chalmers et al., 2008) revealed only two groups of ASPs that differed mainly by insertions/deletions of short amino acid sequences in two regions, one following the CAP3 sequence and another after CAP4. In both studies, the conservation of cysteine residues involved in disulphide linkages differed among the different groups. An important distinguishing criterion between groups of hookworm ASPs is the existence of a cleft located in the equatorial groove, which harbours two histidine and two glutamate residues for group 1. *Ac*-ASP-7 is a member of group 2, and the crystal structure reported

in this study illustrates and confirms the predicted differences compared with the structures of the ASPs *Na*-ASP-1 and *Na*-ASP-2 (group 1). A novel and unexpected observation was the presence of a metal binding site in the C-terminal extension of *Ac*-ASP-7 which constitutes a Hinge-like moiety, including two disulphide linkages. Crystallographic experiments indicate that a backbone/side chain conformation around Glu-171 can serve as a binding site for calcium and manganese, where the metals are coordinated on three sites by two backbone carbonyls and the side chain carboxyl group of one molecule; a fourth coordination is donated by a side chain residue in the N-terminal region of a symmetry-related molecule in the crystal. Combined with the low binding affinity of the protein for metal ions in solution, as observed by isothermal titration calorimetry, these results suggest that this metal binding site could be important in protein-protein interactions.

Although the binding partner for *Ac*-ASP-7 (group 2) has not yet been discovered, neutrophil inhibitory factor (NIF; group 3) is known to bind to an extracellular domain of integrin (Muchowski et al., 1994). Based on the crystal structure of *Na*-ASP-2 (group 1), a model of the complex has been proposed (Asojo et al., 2005). Intriguingly, this model proposes that the residue involved in the interaction of integrin with NIF was Glu-106 (*Na*-ASP-2 numbering), one of the residues of the putative active site that is conserved among members of group 1. Although this residue was found to point inside the cleft of *Na*-ASP-2, a solvent-exposed conformation projecting towards a binding partner seems possible. If Glu-106 is indeed a crucial residue for interaction of ASPs with integrin, the absence of this residue from group 2 ASPs suggests that this interaction does not occur in group 2.

In solution, *Ac*-ASP-7 exists as a mixture of monomeric and dimeric molecular species. The orthorhombic crystal structure of the protein offers several possible options for a quaternary structure. The most likely dimer is the P2-symmetric species, having the largest interaction interface (dimer 1; see Table S3 and Fig. 3). Two reasons support the proposal that this dimer is the most likely oligomer of physiological relevance. Firstly, it possesses the largest interface area, and secondly, the majority of the interaction interface comprises the C-terminal extension, which is specific to this group 2 ASP. Intriguingly, the metal binding site at Glu-171 is accessible in the assembled dimer and might constitute a trigger for dimer disassembly upon occupation. The disassembly of this dimer will expose the ring-like hydrophobic surface, with its characteristic area of negative surface potential located on one side, and the cleft restricted by Tyr-176 in the centre (Fig. 4). Thus, it is highly likely that this interface provides a unique interaction surface for a target protein. Future studies should identify the host targets of the group 2 ASPs, and we are currently panning phage libraries that display random 12-mer peptides to address this ligand binding issue.

In general terms, the analysis of the structure of *Ac*-ASP-7 and the comparison with other SCP/TAPS proteins highlights the interesting and specific design concepts of this protein family. The constitutive topology of an N-terminal PR (or CAP) domain, folding into an α - β - α sandwich, and a C-terminal extension sourced from a pool of different domains, allows for individual targeting of various host proteins by using only limited structural folds. The PR domain, with its disulphide-stabilized fold, provides stability in the extracellular environment within the host. It might also confer some binding, recognition or enzymatic activity, although these aspects await experimental evaluations. The individual targeting of host proteins might be mediated by the particular C-terminal extension that provides specific interaction characteristics for the host target. The delivery of the extension domain (which may be unstable when exposed to the host in isolation) to the required location is ensured by using the N-terminal PR domain as a 'vehicle'.

Given the key biological roles that SCP/TAPS proteins have been proposed to play in a range of eukaryotic systems (e.g., defence against pathogens, allergens and induction of cell proliferation), a detailed understanding of the function of such molecules in each biological system has become a priority. Advances in -omic technologies now provide unique prospects for 'systems biological' investigations of the structure, function and molecular interactions of a range of ASPs.

Such an approach would be advantageous for elucidating the molecular biology and functional roles of these proteins on a fundamental level, and could lead to important biotechnological outcomes, including the design of new drugs or vaccines, designed to disrupt or interrupt key biological pathways linked to ASPs.

Acknowledgements

We gratefully acknowledge funding from the Australian Research Council (RG, AH and AL) and support from the Australian Synchrotron. We thank Darren Pickering, Najju Ranjit and Ben Datu for helpful discussions. CKW is a Postdoctoral Fellow and AL a Principal Research Fellow of the National Health and Medical Research Council (NHMRC) of Australia. Other support to the Gasser Lab from the NHMRC, Australian Academy of Science, the Australian-American Fulbright Commission, Melbourne Water Corporation, the Victorian Life Sciences Computation Initiative (VLSCI) and the IBM Collaboratory is gratefully acknowledged.

References

- Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis. I. W., Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr* 2010;D66:213-21.
- Asojo O. Structure of a two-CAP-domain protein from the human hookworm parasite *Necator americanus*. *Acta Crystallogr D* 2011;67:455-62.
- Asojo OA, Goud G, Dhar K, Loukas A, Zhan B, Deumic V, et al. X-ray structure of *Na*-ASP-2, a pathogenesis-related-1 protein from the nematode parasite, *Necator americanus*, and a vaccine antigen for human hookworm infection. *J Mol Biol* 2005;346:801-14.
- Bethony J, Brooker S, Albonico M, Geiger SM, Loukas A, Diemert D, et al. Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. *Lancet* 2006;367:1521-32.
- Bethony J, Loukas A, Smout M, Brooker S, Mendez S, Plieskatt J, et al. Antibodies against a secreted protein from hookworm larvae reduce the intensity of hookworm infection in humans and vaccinated laboratory animals. *FASEB J* 2005;19:1743-5.
- Bower MA, Constant SL, Mendez S. *Necator americanus*: the *Na*-ASP-2 protein secreted by the infective larvae induces neutrophil recruitment in vivo and in vitro. *Exp Parasitol* 2008;118:569-75.
- Cantacessi C, Campbell BE, Visser A, Geldhof P, Nolan MJ, Nisbet AJ, et al. A portrait of the "SCP/TAPS" proteins of eukaryotes--developing a framework for fundamental research and biotechnological outcomes. *Biotechnol Adv* 2009;27:376-88.
- Cantacessi C, Mitreva M, Jex AR, Young ND, Campbell BE, Hall RS, et al. Massively parallel sequencing and analysis of the *Necator americanus* transcriptome. *PLoS Negl Trop Dis* 2010;4:e684.
- Chalmers IW, McArdle AJ, Coulson RMR, Wagner MA, Schmid R, Hirai H, et al. Developmentally regulated expression, alternative splicing and distinct sub-groupings in members of the *Schistosoma mansoni* venom allergen-like (SmVAL) gene family. *BMC Genomics* 2008;9:89-108.
- Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr* 1994;D50:760-3.
- Datu BJD, Gasser RB, Nagaraj SH, Ong EK, O'Donoghue P, McInnes R, et al. Transcriptional changes in the hookworm, *Ancylostoma caninum*, during the transition from a free-living to a parasitic larva. *PLoS Negl Trop Dis* 2008;2:e130.
- de Silva A, Atukorala S, Weerasinghe I, Ahluwalia N. Iron supplementation improves iron status and reduces morbidity in children with or without upper respiratory tract infections: a

- randomized controlled study in Colombo, Sri Lanka. *Am J Clin Nutr* 2003;77:234-41.
- Del Valle A, Jones BF, Harrison LM, Chadderdon RC, Cappello M. Isolation and molecular cloning of a secreted hookworm platelet inhibitor from adult *Ancylostoma caninum*. *Mol Biochem Parasitol* 2003;129:167-77.
- DeLano W. The PyMOL Molecular Graphics System. <http://www.pymol.org> 2002.
- Emsley P, Cowton K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr* 2004;D60:2126-32.
- Fernández C, Szyperski T, Bruyère T, Ramage P, Möisinger E, Wüthrich K. NMR solution structure of the pathogenesis-related protein P14a. *J Mol Biol* 1997;266:576-93.
- Gibbs GM, Roelants K, O'Bryan MK. The CAP superfamily: cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins--roles in reproduction, cancer, and immune defense. *Endocr Rev* 2008;29:865-97.
- Goud GN, Bottazzi ME, Zhan B, Mendez S, Deumic V, Plieskatt J, et al. Expression of the *Necator americanus* hookworm larval antigen Na-ASP-2 in *Pichia pastoris* and purification of the recombinant protein for use in human clinical trials. *Vaccine* 2005;23:4754-64.
- Goud GN, Zhan B, Ghosh K, Loukas A, Hawdon J, Dobardzic A, et al. Cloning, yeast expression, isolation, and vaccine testing of recombinant *Ancylostoma*-secreted protein (ASP)-1 and ASP-2 from *Ancylostoma ceylanicum*. *J Infect Dis* 2004;189:919-29.
- Hawdon JM, Jones BF, Hoffman DR, Hotez PJ. Cloning and characterization of *Ancylostoma*-secreted protein. A novel protein associated with the transition to parasitism by infective hookworm larvae. *J Biol Chem* 1996;271:6672-8.
- Hotez P, Molyneux D, Fenwick A, Kumaresan J, Sachs S, Sachs J, et al. Control of neglected tropical diseases. *N Engl J Med* 2007;357:1018-27.
- Hotez PJ, Molyneux DH, Stillwaggon E, Bentwich Z, Kumaresan J. Neglected tropical diseases and HIV/AIDS. *Lancet* 2006;368:1865-6.
- Jones TA, Zou JY, Cowan S, Kjeldgaard M. Improved methods for building protein models in electron density maps and location of errors in these models. *Acta Crystallogr A* 1991;47:110-9.
- Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 2007;372:774-97.
- Leslie A. Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* 1992;26.
- Loukas A, Bethony J, Brooker S, Hotez P. Hookworm vaccines: past, present, and future. *Lancet Infect Dis* 2006;6:733-41.
- Mendez S, D' Samuel A, Antoine AD, Ahn S, Hotez P. Use of the air pouch model to investigate immune responses to a hookworm vaccine containing the Na-ASP-2 protein in rats. *Parasite Immunol* 2008;30:53-6.
- Moyle M, Foster DL, McGrath DE, Brown SM, Laroche Y, De Meutter J, et al. A hookworm glycoprotein that inhibits neutrophil function is a ligand of the integrin CD11b/CD18. *J Biol Chem* 1994;269:10008-15.
- Muchowski PJ, Zhang L, Chang ER, Soule HR, Plow EF, Moyle M. Functional interaction between the integrin antagonist neutrophil inhibitory factor and the I domain of CD11b/CD18. *J Biol Chem* 1994;269:26419-23.
- Mulvenna J, Hamilton B, Nagaraj SH, Smyth D, Loukas A, Gorman JJ. Proteomics analysis of the excretory/secretory component of the blood-feeding stage of the hookworm, *Ancylostoma caninum*. *Mol Cell Proteomics* 2009;8:109-21.
- Xiao S, Zhan B, Xue J, Goud GN, Loukas A, Liu Y, et al. The evaluation of recombinant hookworm antigens as vaccines in hamsters (*Mesocricetus auratus*) challenged with human hookworm, *Necator americanus*. *Exp Parasitol* 2008;118:32-40.
- Zhan B, Hawdon J, Qiang S, Hainan R, Huiqing Q, Wei H, et al. *Ancylostoma* secreted protein 1

(ASP-1) homologues in human hookworms. *Mol Biochem Parasitol* 1999;98:143-9.

Zhan B, Liu Y, Badamchian M, Williamson A, Feng J, Loukas A, et al. Molecular characterisation of the *Ancylostoma*-secreted protein family from the adult stage of *Ancylostoma caninum*. *Int J Parasitol* 2003;33:897-907.

Figure Legends

Fig. 1.

Left panel: The fold of *Ac*-ASP-7 resembles the general fold of SCP/TAPS proteins. The colour scheme highlights the N-terminal PR domain (blue) and the C-terminal Hinge domain (red). **Middle panel:** $2F_o - F_c$ electron density around Cys-12 - Cys-54 disulphide bridge. The density is contoured at 1σ . **Right panel:** Superposition of *Ac*-ASP-7 and *Na*-ASP-2 (1u53) showing the region of the putative metal binding site of group 2 ASPs.

Fig. 2.

Binding site for divalent metal ions at Glu-171 in the C-terminal Hinge region. The presence of Ca^{2+} (**left**) and Mn^{2+} (**middle**) was confirmed by the presence of anomalous density (contoured at 2σ and coloured in magenta) in the proximity of the side chain carboxylate of Glu-171. No density was observed for Mg^{2+} . The coordination shell of the metal ions is completed by residue side chains of a symmetry-related molecule. The **right** panel shows the location of the metal with respect to the putative interaction interface in the same orientation as in Fig. 4 (**right**).

Fig. 3.

Quaternary structure of *Ac*-ASP-7. Size exclusion chromatography (SEC) shows the existence of a dimeric (peak 1) and monomeric (peak 2) species in solution. SEC was monitored with by UV absorbance at 280 nm (blue curve), multi-angle light scattering (red curve) and refractive index detection (orange curve). From the crystal packing, three different dimers with interaction interfaces larger than 350 \AA^2 can be identified. The colour scheme of the three-dimensional structures identifies the N-terminal PR domain (blue) and the C-terminal Hinge-like region (red).

Fig. 4.

Dimer 1 interaction interface. In the assembled dimer (**left**), the interface encloses a flat ring-like hydrophobic surface with a cluster of negative surface charge on one side (**right**). In the centre of the ring is a cleft restricted at the bottom by the aromatic ring of Tyr-176. The ring-like surface might serve as an interaction interface for *Ac*-ASP-7 target proteins.

Tables

Table 1
Data collection and refinement statistics

	Native <i>Ac</i> -ASP-7			<i>Ac</i> -ASP-7:Ca ²⁺			<i>Ac</i> -ASP-7:Mn ²⁺		
	Total	Molecule A	Molecule B	Total	Molecule A	Molecule B	Total	Molecule A	Molecule B
Data collection									
Space group	P2 ₁ 2 ₁ 2 ₁			P2 ₁ 2 ₁ 2 ₁			P2 ₁ 2 ₁ 2 ₁		
Cell dimensions									
<i>a</i> , <i>b</i> , <i>c</i> (Å)	59.4, 79.2, 94.9			58.8, 78.0, 96.3			59.0, 78.7, 95.3		
α , β , γ (°)	90, 90, 90			90, 90, 90			90, 90, 90		
Wavelength (Å)	0.93224			1.3778			1.3778		
Resolution (Å)	2.4			2.7			2.6		
<i>R</i> _{sym}	0.074 (0.260)			0.063 (0.169)			0.081 (0.376)		
<i>I</i> / σ <i>I</i>	6.8 (2.8)			9.0 (4.2)			7.3 (2.0)		
Completeness	1.0 (1.0)			0.992 (0.986)			0.999 (0.994)		
Redundancy	7.0 (7.2)			14.1 (14.5)			14.1 (13.7)		
Refinement									
Resolution (Å)	2.4			2.7			2.6		
No. reflections	18086			12548			14201		
<i>R</i> _{work} / <i>R</i> _{free}	0.196 (0.231) / 0.241 (0.319)			0.180 (0.218) / 0.243 (0.335)			0.189 (0.282) / 0.254 (0.354)		
No. atoms									
Protein	2693	1341	1352	2718	1331	1387	2733	1351	1382
Ligand/ion				1 (Ca ²⁺)			1 (Mn ²⁺)		
Water	100			100			78		
Average B-factors									
Protein (Å ²)	36.6	36.3	36.9	30.8	29.9	31.7	40.6	39.9	41.4
Ligand/ion (Å ²)				42.9			70.6		
Water (Å ²)	36.6			31.3			39.5		
R.m.s deviations									
bond lengths (Å)	0.008			0.008			0.008		
bond angles (°)	1.216			1.157			1.209		
B-factor for bonded atoms	3.63	3.64	3.63	3.95	3.89	3.99	3.93	3.92	3.95

*Datasets were obtained from one crystal each. Values in parentheses are for highest-resolution shell.