A novel C-type lectin secreted by a tissue-dwelling parasitic nematode

Alex Loukas*, Nicholas P. Mullin[†], Kevin K.A. Tetteh*, Luc Moens[‡] and Rick M. Maizels*

Many parasitic nematodes live for surprisingly long periods in the tissues of their hosts, implying sophisticated mechanisms for evading the host immune system. The nematode *Toxocara canis* survives for years in mammalian tissues, and when cultivated *in vitro*, secretes antigens such as TES-32. From the peptide sequence, we cloned TES-32 cDNA, which encodes a 219 amino-acid protein that has a domain characteristic of host calcium-dependent (C-type) lectins, a family of proteins associated with immune defence. Homology modelling predicted that TES-32 bears remarkable structural similarity to mammalian immune-system lectins. Native TES-32 acted as a functional lectin in affinity chromatography. Unusually, it bound both mannose- and galactose-type monosaccharides, a pattern precluded in mammalian lectins by a constraining loop adjacent to the carbohydrate-binding site. In TES-32, this loop appeared to be less obtrusive, permitting a broader range of ligand binding. The similarity of TES-32 to host immune cell receptors suggests a hitherto unsuspected strategy for parasite immune evasion.

Addresses: *Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, UK. †Department of Chemistry, University of Edinburgh, Edinburgh EH9 3JJ, UK. ‡Department of Biochemistry, University of Antwerp (UIA), Wilrijk B-2610, Belgium.

Correspondence: Rick M. Maizels
E-mail: Rick.Maizels@ed.ac.uk

Results and discussion
Identification of TES-32 as a C-type lectin

TES-32 is a 32 kDa N-glycosylated protein secreted by tissue-stage *T. canis* parasites [1]. Secreted *Toxocara* excretory–secretory (TES) products collected from larval *T. canis* in culture [2] were separated by SDS–PAGE, electrotransferred to PVDF membrane and the TES-32 band excised. Peptides from trypsinised TES-32 were then separated by reverse-phase HPLC and sequenced. Peptide sequences were matched against a database of 250 expressed sequence tags (ESTs) from a larval *T. canis* cDNA library

[3]. Clone *Tc*-EST-070 (*Tc-ctl-1*) encoded one of the TES-32 peptides. On complete sequencing, *Tc-ctl-1* proved to contain a full-length 762 bp cDNA insert, and to encode a second peptide derived from gel-purified TES-32.

The deduced CTL-1 protein sequence comprises 219 amino acids, including a putative 18 amino-acid signal peptide and three potential N-glycosylation sites (Figure 1a). The amino-terminal portion (amino acids 19 to 80) of the mature protein includes eight cysteine residues, four of which are in (V)CVAP-like motifs (using the single-letter amino-acid code), and a short threonine-rich tract (Figure 1a). This amino-terminal portion shows no significant homology to other known proteins. The carboxy-terminal domain (amino acids 81 to 219), however, was found to have strong similarities to mammalian and invertebrate C-type lectins (CTLs), including the macrophage mannose receptor (MMR), CD23 and E-selectin (Figure 1b). Maximal amino-acid identity was 31% (43/139) between amino acids 81 to 219 of TES-32 and the carbohydrate-recognition domain (CRD) 8 of human MMR (amino acids 1230 to 1359). The gene encoding TES-32 was therefore designated *Tc-ctl-1*.

Antibodies to recombinant *Tc*-CTL-1 (rCTL-1) reacted with native TES-32 in Western blots (Figure 2), and with the 27 kDa recombinant protein (a dimeric form is also apparent). Furthermore, rCTL-1 was also recognised by monoclonal antibody Tcn-3, which specifically binds TES-32 [2] (Figure 2). The difference in size between the predicted mature *Tc*-CTL-1 protein (22 kDa) and the actual protein found in TES products can be attributed to N-glycosylation [4,5]. Monoclonal antibody Tcn-3 showed TES-32 to be localised to the parasite epicuticle [1]; an anti-rCTL-1 antiserum showed a similar surface-binding pattern (data not shown). Additional bands of 45 kDa and > 200 kDa were recognised in TES products by the anti-rCTL-1 antiserum; peptide sequence data indicated that the native 45 kDa component is a second lectin derived from a distinct gene (unpublished observations).

TES-32 is a bispecific lectin

The lectin-like properties of native TES products were then tested on affinity columns of different monosaccharides in a Ca^{2+}-containing buffer. A 32 kDa protein was selectively retarded on a GalNAc column (Figure 3). Dissociation from the ligand occurred slowly, with protein evident throughout ten column-volume washes, and was completed in the presence of EDTA. Binding was greatly

Figure 1
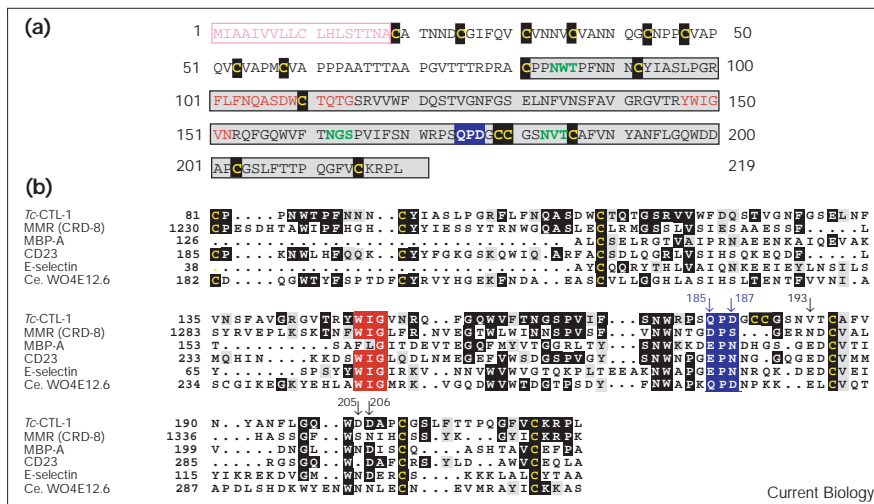
Sequence of *Tc*-CTL-1 and its similarity to other CTLs. **(a)** Amino-acid sequence of *Tc*-CTL-1, showing the signal peptide (pink), the CTL CRD (grey shading) and *N*-glycosylation sites (green lettering). Cysteine residues are yellow, the carbohydrate-binding motif is in white lettering on a blue background, and tryptic peptide sequences are red. **(b)** Alignment of CTLs. Cysteine residues are yellow; the five conserved Ca^{2+}-ligating residues (site 2) in MBP-A are indicated by arrows (MBP-A numbering). The sugar-binding motifs QPD and EPN are in white lettering on a blue background. Residues that are identical between *Tc*-CTL-1 and the other CTLs are in white lettering on a black background; similar amino acids are shaded grey. The WIG motif (red shading) in *Tc*-CTL-1 is conserved with minor substitutions [10]. GenBank accession numbers: human MMR, M93221; rat MBP-A, M14105; mouse CD23, M99371; human E-selectin, M24736. Ce WO4E12.6 is a predicted protein from *Caenorhabditis elegans*.

reduced when free GalNAc was added to the TES products before loading (data not shown). Surprisingly, the same 32 kDa molecule bound almost as well to D-mannose–agarose but not to unconjugated agarose alone (Figure 3). Although different affinity matrices may not bear comparable ligand densities, distinctly weaker

Figure 2
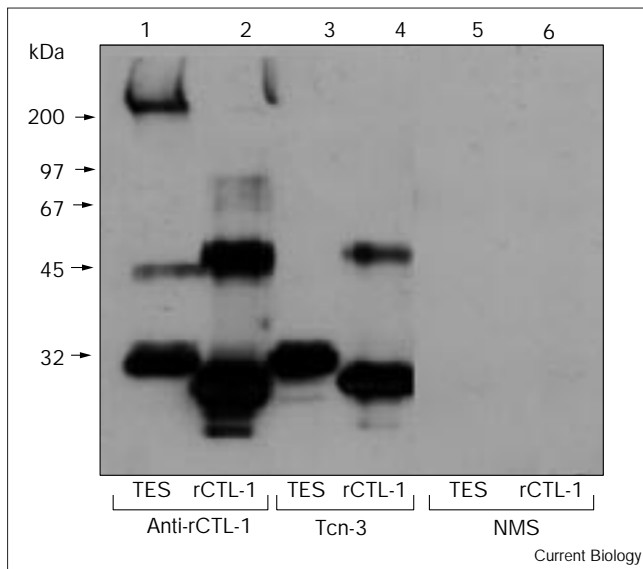
Identity of *Tc*-CTL-1 and TES-32. A western blot of native TES products (lanes 1,3 and 5) and of rCTL-1 (lanes 2,4 and 6) showing that both proteins are recognised by anti-rCTL-1 serum (lanes 1,2) and by monoclonal antibody Tcn-3 (lanes 3,4). Normal mouse serum (NMS; lanes 5,6) was used as a control.

Figure 3
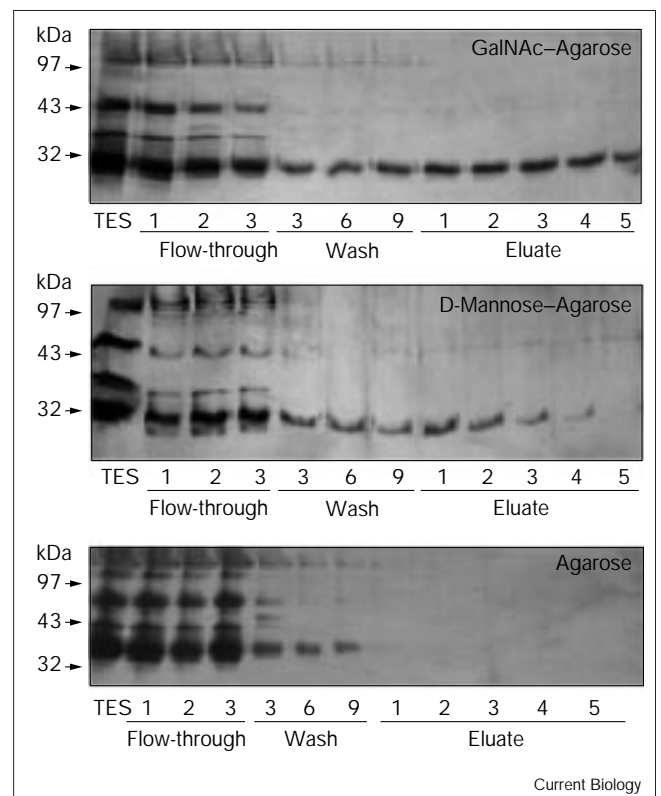
Carbohydrate binding by *Tc*-CTL-1. SDS–PAGE profiles of TES products after affinity chromatography on GalNAc–agarose, D-mannose–agarose or unconjugated agarose. Fractions of 1 ml were collected from a 1 ml column loaded successively with 4 ml loading buffer (flow-through), 10 ml wash buffer and 5 ml elution buffer.

binding of the 32 kDa protein was observed to GlcNAc and D-galactose compared with GalNAc or D-mannose (data not shown). Both the anti-rCTL-1 serum and mono-clonal antibody Tcn-3 bound the 32 kDa protein eluted from both GalNAc and D-mannose columns, confirming that the retarded native protein in each case was *Tc*-CTL-1. In addition, a second 45 kDa protein, similar to the one detected in western blotting, exhibited weak binding to D-mannose (Figure 3), consistent with data indicating that it is encoded by a distinct lectin gene (unpublished observations).

Structural conservation between mammalian and nematode lectins

Co-crystallisation of rat mannose-binding protein-A (MBP-A) and an oligomannose ligand has identified the residues at which CTLs ligate Ca^{2+} and saccharides [6]. We aligned the CRDs of *Tc*-CTL-1 and CTLs with known structure, and modelled the CRD of *Tc*-CTL-1 (Figure 4). In this prediction, *Tc*-CTL-1 shows a remarkable structural similarity to MBP-A, with the major beta sheets and alpha helices being in corresponding positions.

All CTLs possess a conserved Ca^{2+}-binding site (designated Ca^{2+} number 2) [7], and *Tc*-CTL-1 contains four of the five residues necessary to ligate Ca^{2+} in MBP-A (Figure 4); the fifth Ca^{2+}-binding residue cannot be determined from the alignment. But the similarity of this site to that of MBP-A indicates that sugar ligation by *Tc*-CTL-1 is likely to resemble that observed in the crystal structure of MBP-A [6]. Both MBP-A and tetranectin, but not E-selectin, contain a second (auxiliary) Ca^{2+} binding site (Ca^{2+} number 1), which does not bind directly to sugar. This site appears to be absent from *Tc*-CTL-1 which in this regard is more similar to E-selectin than to MBP-A or tetranectin.

CTLs can be loosely grouped into those that bind D-galactose and its derivatives (Gal-type), and those that bind D-mannose, D-glucose and their derivatives (Man-type). The specificity of MBP-A, a Man-type CTL, is determined by calcium-binding residues at positions 185 and 187 (MBP-A numbering) [6] and by residue 189 and the adjacent loop [8]. Thus, Man-type CRDs typically have Glu185 and Asn187, whereas Gal-type CRDs have Gln185 and Asp187, with Trp- or Phe189 [8,9]. The mannose-specific MBP-A can be converted to bind galactose by replacement of residues at positions 185 and 187 with Gal-type residues, and rendered galactose-specific by gaining Trp189 and the adjacent five-residue loop from a natural galactose-binding lectin [8]. In *Tc*-CTL-1, this loop appears significantly modified. Thus, in place of a fold which brings His/Trp189 close to the ligand, the protein backbone in *Tc*-CTL-1 is more extended and the bulky aromatic residue replaced by an unusual pair of cysteines at positions 179 and 180 (equivalent residues are 189 and 190 in MBP-A). The predicted position of these cysteine residues is too distant for them to have a direct role in forming the binding site for Ca^{2+} and saccharides. As a consequence of these changes, the ligation of *Tc*-CTL-1 to either Gal-type or Man-type saccharides may be unimpeded. Site-directed mutagenesis to convert this region of *Tc*-CTL-1 would test this hypothesis.

Biological significance of parasite lectins

This is the first report of a gene encoding a CTL from a pathogenic organism; in contrast, most other CTLs are involved in host defence against infection [10]. The latter include the selectins, the low affinity IgE receptor CD23, the NK cell receptors Ly-49 and CD94, and antigen-presenting cell lectins such as the MMR and the dendritic cell receptor DEC-205. Genes encoding CTLs have now

Figure 4

Similarity of mammalian and nematode lectins. Structural model of the CRD of *Tc*-CTL-1 based on the crystal structure of MBP-A. The His189 residue of MBP-A is close to the sugar-binding site, but no equivalent large residue exists in *Tc*-CTL-1. The conserved and auxiliary Ca^{2+}-binding sites in MBP-A are depicted as black spheres. It is not yet known whether the two newly identified cysteine residues (Cys179 and Cys180) in *Tc*-CTL-1 retain free SH groups, as shown. CRD sequences of *Tc*-CTL-1 (amino acids 81–219), rat MBP-A (M14105), human E-selectin (M24736) and human tetranectin (X64559) were aligned for the MODELLER program [19], and figures were produced using MOLSCRIPT [20]. Five separate models of *Tc*-CTL-1 with near-identical structures were produced, differing only in the carboxy-terminal loop (residues 201 onwards).
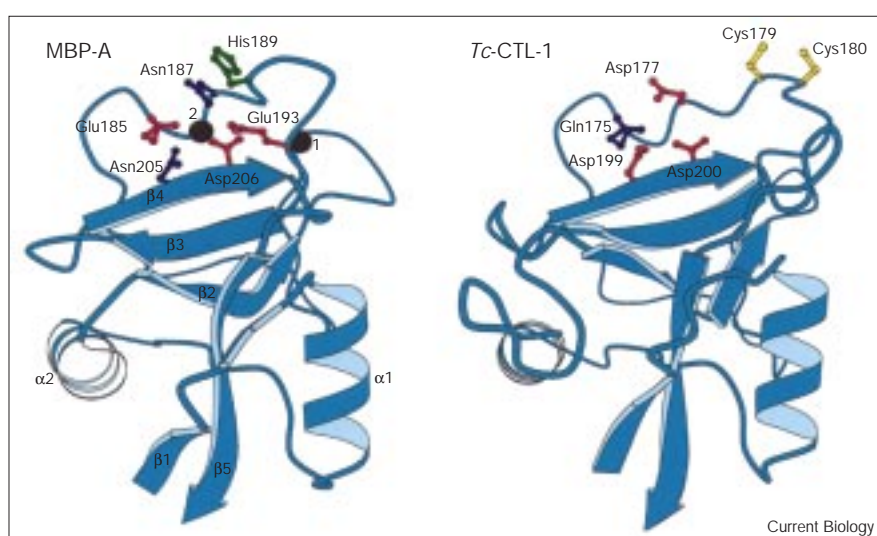
been identified in invertebrates, including arthropods, molluscs, echinoderms and tunicates, and some of the proteins encoded by these genes are thought to be involved in agglutinating pathogens [11]. The free-living nematode *C. elegans* contains 102 genes encoding CRD-like domains of CTLs [12]. Interestingly, *Tc*-CTL-1 has only weak similarity with these putative gene products, sharing greater identity with mammalian CTLs than with the closest *C. elegans* homologue (Figure 2b).

T. canis secretions contain a substantial glycan content — ~40% by weight [4] — and the predominant *O*-linked glycan contains a terminal fucose which is *O*-methylated at position 2 [13]. These carbohydrates are abundant on other secreted proteins such as mucins [2,14]. As adjacent hydroxyls at carbons 2 and 3 of fucose (or 3 and 4 of other hexoses) are required for binding to mammalian CTLs[6], *O*-methylation may prevent the parasite lectin from binding to carbohydrate structures being secreted at the same time.

It is intriguing that helminth parasites secrete proteins with significant homology to host molecules that are instrumental in the immune response against the very same pathogens. This finding suggests an exciting new parasite immune-evasion strategy [15]. Although there is much information on the role of host lectins in the innate and acquired immune responses to infectious organisms, little is known about the role of pathogen-derived CTLs. We are now investigating the potential host ligands of *Tc*-CTL-1, at both a molecular and cellular level, in an attempt to elucidate the biological function of this class of CTL proteins in host–parasite interactions.

Materials and methods

Parasites, protein sequencing and expression
Larval *T. canis* were maintained *in vitro* in serum-free medium, from which secreted TES products were collected [16]. TES proteins were separated by SDS–PAGE and transferred to PVDF membrane as published [17]. TES-32 was excised and digested with trypsin [17], and HPLC-purified peptides sequenced on an ABI-471-B sequencer (Applied Biosystems). The larval cDNA library made in Uni-ZAP XR (Stratagene) was a generous gift from C. Tripp and R. Grieve (Heska). The mature region of *Tc*-CTL-1 (Cys19–Leu219) was expressed by cloning into the pET-29T expression vector (Novagen), and purified on metal chelation columns. BALB/c mice were immunised with three doses of 10 µg rCTL-1.

Detection of lectin activity using immobilised monosaccharides
Ligand-binding assays [18] employed 100 µg TES product diluted in 500 µl of low-calcium loading buffer (1.25 M NaCl, 25 mM Tris pH 7.4, 2.5 mM CaCl$_2$) and applied to 1 ml columns of 4% beaded agarose cross-linked to GalNAc (Sigma A2787, 6 hydroxy-linked), GlcNAc, β-D-glucose or D-mannose (Sigma A2278, G2019 and M6400, respectively; all hydroxy-linked). Columns were rinsed with 4 ml 2.5 mM CaCl$_2$ buffer and washed with 10 ml 25 mM CaCl$_2$ buffer. Bound proteins were eluted with 2.5 mM EDTA in calcium-free buffer.

References
1. Page AP, Hamilton AJ, Maizels RM: *Toxocara canis*: monoclonal antibodies to carbohydrate epitopes of secreted (TES) antigens localize to different secretion-related structures in infective larvae. *Exp Parasitol* 1992, **75**:56-71.
2. Maizels RM, Kennedy MW, Meghji M, Robertson BD, Smith HV: **Shared carbohydrate epitopes on distinct surface and secreted antigens of the parasitic nematode *Toxocara canis*.** *J Immunol* 1987, **139**:207-214.
3. Tetteh KKA, Loukas A, Tripp C, Maizels RM: **Identification of abundantly-expressed novel and conserved genes from the infective larval stage of *Toxocara canis* by an expressed-sequence tag strategy.** *Infect Immun* 1999, in press.
4. Meghji M, Maizels RM: **Biochemical properties of larval excretory–secretory (ES) glycoproteins of the parasitic nematode *Toxocara canis*.** *Mol Biochem Parasitol* 1986, **18**:155-170.
5. Page AP, Maizels RM: **Biosynthesis and glycosylation of serine/threonine-rich secreted proteins from *Toxocara canis* larvae.** *Parasitology* 1992, **105**:297-308.
6. Weis WI, Drickamer K, Hendrickson WA: **Structure of a C-type mannose-binding protein complexed with an oligosaccharide.** *Nature* 1992, **360**:127-134.
7. Drickamer K: **Engineering galactose-binding activity into a C-type mannose-binding protein.** *Nature* 1992, **360**:183-186.
8. Iobst ST, Drickamer K: **Binding of sugar ligands to Ca(2+)-dependent animal lectins. II. Generation of high-affinity galactose binding by site-directed mutagenesis.** *J Biol Chem* 1994, **269**:15512-15519.
9. Kolatkar AR, Weis WI: **Structural basis of galactose recognition by C-type animal lectins.** *J Biol Chem* 1996, **271**:6679-6685.
10. Weis WI, Taylor ME, Drickamer K: **The C-type lectin superfamily in the immune system.** *Immunol Rev* 1998, **163**:19-34.
11. Wilson R, Chen C, Ratcliffe NA: **Innate immunity in insects: the role of multiple, endogenous serum lectins in the recognition of foreign invaders in the cockroach, *Blaberus discoidalis*.** *J Immunol* 1999, **162**:1590-1596.
12. The *C.elegans* Genome Sequence Consortium: **Genome sequence of *Caenorhabditis elegans*: a platform for investigating biology.** *Science* 1998, **282**:2012-2018.
13. Khoo K-H, Maizels RM, Page AP, Taylor GW, Rendell N, Dell A: **Characterisation of nematode glycoproteins: the major *O*-glycans of *Toxocara* excretory secretory antigens are methylated trisaccharides.** *Glycobiology* 1991, **1**:163-171.
14. Gems DH, Maizels RM: **An abundantly expressed mucin-like protein from *Toxocara canis* infective larvae: the precursor of the larval surface coat glycoproteins.** *Proc Natl Acad Sci USA* 1996, **93**:1665-1670.
15. Maizels RM, Bundy DAP, Selkirk ME, Smith DF, Anderson RM: **Immunological modulation and evasion by helminth parasites in human populations.** *Nature* 1993, **365**:797-805.
16. Page AP, Richards DT, Lewis JW, Omar HM, Maizels RM: **Comparison of isolates and species of *Toxocara* and *Toxascaris* by biosynthetic labelling of somatic and ES proteins from infective larvae.** *Parasitology* 1991, **103**:451-464.
17. Coligan JE, Kruisbeek AM, Margulies DH, Shevach EM, Strober W: *Current Protocols in Immunology.* New York: Greene Publishing Associates and Wiley-Interscience; 1997.
18. Burrows L, Iobst ST, Drickamer K: **Selective binding of N-acetylglucosamine to the chicken hepatic lectin.** *Biochem J* 1997, **324**:673-680.
19. Sali A, Blundell TL: **Comparative protein modelling by satisfaction of spatial restraints.** *J Mol Biol* 1993, **234**:779-815.
20. Kraulis PJ: **MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures.** *J Appl Crystallogr* 1991, **24**:946-950.