

Extensive diversity in repeat unit sequences of the cDNA encoding the polyprotein antigen/allergen from the bovine lungworm *Dictyocaulus viviparus*

Collette Britton^{a,*}, Joyce Moore^a, John S. Gilleard^{b,1}, Malcolm W. Kennedy^a

^a Wellcome Laboratories for Experimental Parasitology, University of Glasgow, Glasgow, UK

^b Department of Veterinary Parasitology, University of Glasgow, Glasgow, UK

Received 3 January 1995; accepted 21 April 1995

Abstract

The complete sequence of the cDNA encoding the nematode polyprotein allergen/antigen (NPA) of the bovine lungworm *Dictyocaulus viviparus* was obtained by immunoscreening of cDNA expression libraries and by 5' RACE (rapid amplification of cDNA ends). The encoded polypeptide is similar in sequence to the ABA-1 allergen of *Ascaris*, the gp15/400 'ladder' protein of *Brugia malayi*, *Brugia pahangi* and *Wuchereria bancrofti*, and a 15-kDa antigen of *Dirofilaria immitis*. As with these, the predicted amino-acid sequence comprises a head-to-tail array of similar polypeptides with regularly spaced consensus proteinase cleavage sites. The *D. viviparus* protein was designated DvA-1 (*D. viviparus* antigen-1) and the gene *dva-1*. The deduced amino-acid sequence of DvA-1 showed features not observed before in other NPAs: (i) a hydrophobic leader peptide is present, (ii) none of the 12 units in the array are identical and the sequences diverge to a degree hitherto unseen in the NPAs of other nematode parasites, (iii) the predicted proteinase cleavage sites are also diverse in sequence and, in two instances, no consensus cleavage site was identifiable at the expected position, (iv) a short repeat unit is present, which is the only one containing a consensus N-glycosylation site and (v) a C-terminal extension peptide is encoded which shows no similarity to that from *A. suum* ABA-1. Comparison of independent cDNAs revealed slight variations in the sequence of the gene within the parasite population. Antisera to recombinant DvA-1 polypeptide identified 14–15-kDa antigens in both parasite somatic and excretory-secretory material. DvA-1 is the only NPA for which the complete coding sequence is available and the new principles which it illustrates may lie unsuspected in the NPA-encoding genes of all nematode parasites.

Keywords: *Dictyocaulus viviparus*; Nematode; Excretory-secretory antigen; Nematode polyprotein allergen/antigen

Abbreviations: DvA-1, *Dictyocaulus viviparus* antigen-1; ES, excretory-secretory; NPA, nematode polyprotein allergen/antigen; RACE, rapid amplification of cDNA ends; SL, spliced leader; UTR, untranslated region.

Note: Nucleotide sequence reported in this paper have been submitted to the GenBankTM data base under the accession No. U02568.

* Corresponding author. **Present address:** Department of Pathology, Laboratory of Molecular Pathology, University of California at San Francisco, 4150 Clement Street 113B, San Francisco, CA 94121, USA. Fax: (1-415) 750-6947; e-mail: britton@cgl.ucsf.edu

¹ **Present address:** Institute of Genetics, University of Glasgow, Glasgow, UK.

1. Introduction

Several species of parasitic nematode produce an approx. 15-kDa antigen which is abundant in both somatic and excreted/secreted (ES) material [1–5]. These proteins are produced as large precursors which are subsequently proteolytically cleaved to the 15-kDa polypeptide units. Sequencing of cDNAs has revealed not only similarities in sequence between species, but also that the structure of the genes is unusual in comprising a head to tail array encoding multiple units of the 15-kDa polypeptides, with consensus proteinase cleavage sites at the junctions. This manner of synthesis and processing led to the term ‘nematode polyprotein allergens/antigens’ (NPA) for this family of proteins [6]. In most cases there is little variation between the amino-acid sequence of the repeated units within a given array, or in the structure of the putative proteinase cleavage sites [1–5]. Only a small proportion of NPA units have been examined for any species, however, so these generalisations may not withstand complete analysis of NPA-encoding genes.

The NPA of *Ascaris suum*, ABA-1, is known to be a major allergen, eliciting an IgE response in experimentally infected rodents [7]. It has also been reported that a significant number of individuals infected with brugian filariasis show IgE reactivity to the *Brugia* NPA, gp15/400, and there is a positive correlation between those showing IgE reactivity and those exhibiting pathology [8]. In the present study, which aimed to characterise genes encoding ES antigens of the cattle lungworm *Dictyocaulus viviparus*, we isolated the gene encoding an antigen with significant similarity to the NPAs of other nematodes. The presence of this antigen, named DvA-1, in ES material suggests it is exposed to the host and may play a role in the pathology of lungworm infection.

We present here the complete cDNA sequence of the *D. viviparus* *dva-1* gene. The encoded protein was found to have features which are unique amongst, yet may lie unsuspected in, the NPAs of other species. These features include an unprecedented degree of heterogeneity between the sequence of individual repeat units, the presence of one short unit, diverse putative proteinase cleavage sites and the presence of a hydrophobic leader peptide.

2. Materials and methods

2.1. Parasites

Infective third-stage larvae (L3) were kindly provided by Intervet UK (Cambridge, UK). Adult worms were collected from the lungs of calves killed 28–35 days after infection with 5000 infective larvae or from field cases of patent lungworm infection. Adult in vitro excretory-secretory (ES) products and Tris-soluble extracts of L3 and adult worms were prepared as previously described [9]. The protein concentration of samples was estimated using a Coomassie-blue-based assay (Pierce, Rockford, IL, USA).

2.2. Construction of cDNA libraries in λ ZAPII

Total RNA was isolated from infective larvae and from adult worms by the hot phenol extraction method [10]. mRNA was purified using oligo(dT)-Sephrose (Poly(A) Quik mRNA isolation kit, Stratagene, Cambridge, UK) and its integrity tested by in vitro translation (Promega L4960, Southampton, UK). L3 and adult-stage cDNA libraries were constructed from 5 μ g mRNA using Uni-ZAP-XR-cDNATM synthesis kit according to the manufacturer's instructions (Stratagene).

2.3. Immunoscreening of cDNA expression libraries

The amplified L3 cDNA library (approx. 1.4×10^6 plaques) was screened with a 1:500 dilution of antiserum to *D. viviparus* adult ES products prepared as described previously [9]. The antiserum was first absorbed against *E. coli* lysate prior to immunoscreening [11] and all wash and incubation steps were carried out by standard procedures using an alkaline phosphatase conjugated secondary antibody. Positive plaques were isolated following three rounds of immunoscreening.

Approx. 1.5×10^6 plaques from the amplified *D. viviparus* adult stage cDNA library were screened as described above with anti-ES serum. Fifteen agar plugs from the first screen, which each contained approx. 100 individual plaques, were then screened with antibody specific for phage clone CB10 which

was prepared by the antibody elution method [12]. Phage clone CB10 had been previously isolated from the L3 cDNA library by screening with anti-ES

serum. Positive plaques from the adult cDNA library were purified through three rounds of immuno-screening.

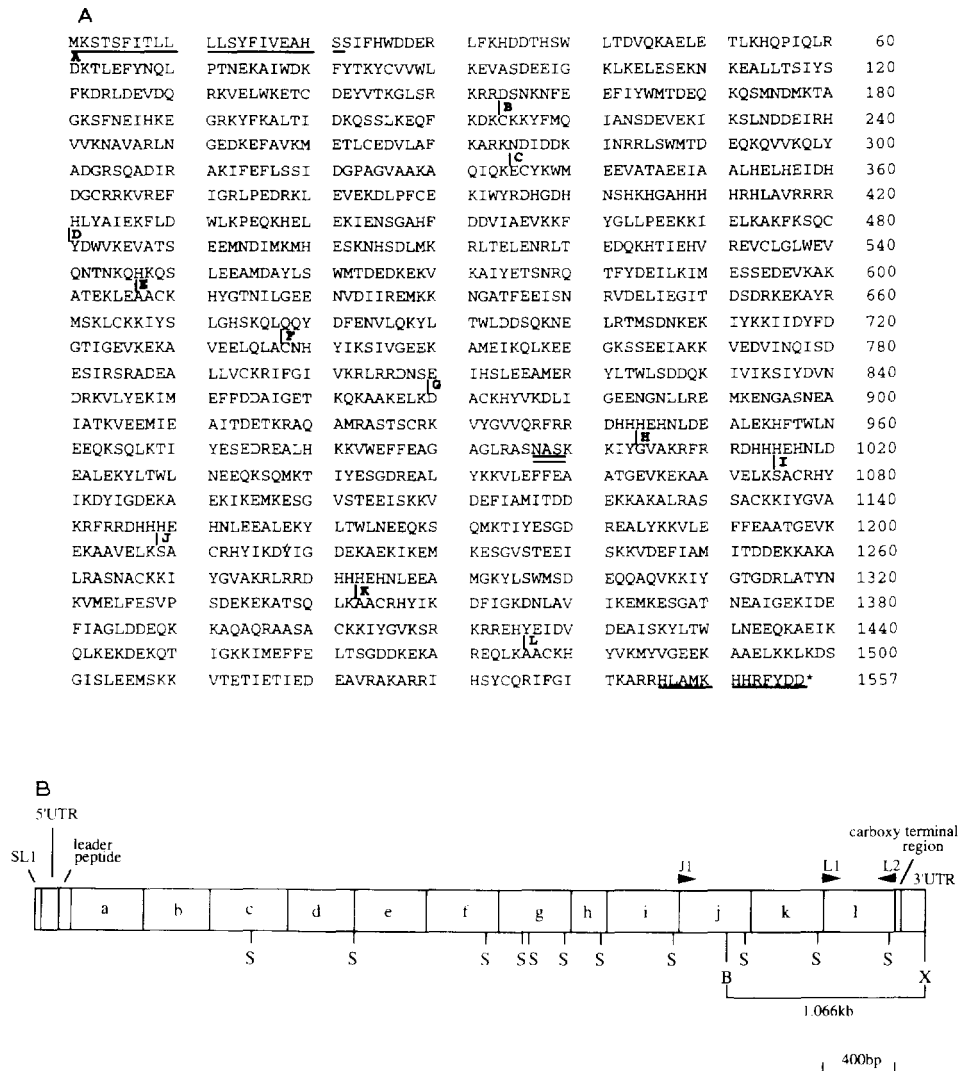


Fig. 1. (A) The predicted amino-acid sequence of the DvA-1 polypeptide of *Dictyocaulus viviparus*. The putative boundaries between repeating units are indicated by vertical lines and are labelled in bold with the letter designation of each repeat unit. The putative leader peptide and C-terminal extension are underscored and the potential N-linked glycosylation site is double underscored. Base differences found between clone pCB10, from the L3 cDNA library and adult cDNA clones were as follows: C(2943)T, C(2949)G, C(2970)G, A(3028)G, G(3076)C, A(4574)G, C(4735)T, A(4736)G, T(4784)C, G(4841)A, G(4847)A, A(4861)C, G(4893)T and the resultant amino-acid changes as follows: H(945)Y, L(947)V, H(954)D, E(973)G, G(989)A, A(1543)V. (B) Diagrammatic representation of the cDNA encoding DvA-1. The 5' and 3' UTR, the SL1 sequence and the regions encoding the putative hydrophobic leader and C-terminal extension are indicated, as are the positions of primers used in PCR amplification of repeat unit *l* (primers L1 and L2) and the continuum of repeats *j-k-l*, (primers J1 and L2) from pCB10. Also shown is the restriction map of *dva-1* with *Sfa*NI (S) and the 1066-bp *Bst*XI (B)-*Xho*I (X) fragment used in the hybridisation experiments.

2.4. Sequencing of cDNA inserts

Plasmid DNA, excised using helper phage R408 (Stratagene), was purified using Magic Miniprep DNA purification system (Promega A7100) and insert DNA was sequenced using Sequenase 2.0 (US Biochemical, Cleveland, OH, USA). To enable complete sequencing of pCB10, sequencing was also performed on Exonuclease III nested deletions of insert DNA, generated using the Erase-a-Base system (Promega E5850). The 5' RACE product was sequenced by a combination of subcloning of restriction fragments and use of gene-specific oligonucleotide primers.

2.5. 5'-Rapid amplification of cDNA ends (5'-RACE)

5'-RACE was carried out to enable sequencing of the 5' end of *dva-1*. First-strand cDNA was prepared from 1 µg of adult *D. viviparus* RNA using the gene-specific primer P1 (gcggccgcTTATAATATCAACAT; lower case indicates *NotI* site; upper case indicates nt 1983–1969 of *dva-1* antisense strand) and Superscript reverse transcriptase (Gibco). The cDNA was then dC-tailed using terminal deoxynucleotide transferase (Gibco) and PCR was carried out using either an anchor primer (Gibco) or a spliced leader (SL1) primer (GgaattcGGTTTAATTACCAAGTTTGAG; lower case indicates *EcoRI* site) in combination with a second gene-specific primer P2 (gcggccgcTTCCATAATTTTAGAATTCA; lower case indicates *NotI* site; upper case indicates nucleotides 1880–1859 of *dva-1* antisense strand), for 30 cycles at 94°C/1 min, 55°C/1 min and 72°C/2 min. The products were visualised on an ethidium bromide-stained gel which showed that priming with the SL1 or anchor primer gave products of the same size. PCR products primed with SL1 and P2 were cloned into pT7Blue (Invitrogen) according to the conditions recommended by the manufacturer.

2.6. Hybridisation of nucleic acids

Genomic DNA (3 µg), isolated by standard procedures [11], was digested with restriction endonucleases, fractionated by agarose gel electrophoresis and transferred to Hybond-N nylon membrane (RPN

303N, Amersham) in 20 × standard saline citrate (SSC) according to standard procedures [11]. The DNA was then UV fixed and hybridised overnight at 65°C in Church buffer [13], with a random primed ³²P-labelled *BstXI-XhoI* (1066 bp, see Fig. 1B) fragment of pCB10. Following hybridisation, blots were washed at 65°C by standard procedures [11] and exposed to Fuji RX X-ray film (Technical Plot Systems, Cumbernault, UK).

Northern blot analysis was performed using 10 µg total RNA isolated from L3 or adult worms using the TRIzol™ reagent according to the manufacturer's instructions (Gibco BRL). This was denatured and fractionated on 1.2% agarose-formaldehyde gels. The RNA was then transferred onto Hybond-N (Amersham), UV fixed, baked at 80°C for 30 min and hybridised at 65°C overnight in Pipes buffer (50 mM Pipes, pH 6.8/100 mM NaCl/50 mM Na₂HPO₄/50 mM NaH₂PO₄/1 mM EDTA/5% SDS) with random primed ³²P-labelled *BstXI-XhoI* pCB10 fragment. Filters were washed as described for Southern blot analysis.

2.7. PCR amplification and *Dva-1* recombinant protein expression

DNA encoding *dva-1* repeat unit *l* was amplified from pCB10 by PCR using primers designed to the 5' and 3' ends of repeat *l* (primer L1, AC-GACGgatccTTACGAAATAGACGTG; primer L2, ATGTTggatccCCTAATGACGTCGTGCTTTTCG; lower case denotes *BamHI* sites; primer positions indicated in Fig. 1B). DNA encoding the three contiguous repeats J-K-L was also amplified using a primer designed to the 5' end of repeat *j* (primer J1: GAGGgatccTCATGAACATAATCTAGAA; lower case denotes *BamHI* site; position indicated in Fig. 1B) and primer L2. Amplification was carried out for 30 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 1 min. The DNA fragments were purified from agarose gel, then digested with *BamHI* and ligated into *BamHI* digested and dephosphorylated pET15b expression vector (Novagen 69729-1, AMS Biotechnology, Witney, UK). Transformants containing insert DNA in the correct orientation, as determined by DNA sequencing, were grown to an OD₆₀₀ of 0.6 then induced with 1 mM IPTG at 37°C for 2

h. Cells were collected by centrifugation at $4000 \times g$, resuspended in binding buffer (5 mM imidazole/0.5 M NaCl/20 mM Tris · HCl, pH 7.9) containing 0.1% Nonidet P-40 (NP-40; Sigma) and sonicated on ice at 18 μ m amplitude for five 1-min bursts. Following centrifugation at $15000 \times g$, the supernatant was applied to His-Bind metal chelation resin (Novagen 69670-1) packed in a 2-ml column. The washing procedure and elution of the fusion protein were carried out according to the manufacturer's instructions. The N-terminal histidine tag was cleaved by thrombin (Invitrogen 69671-1) at a concentration of 1 unit per mg protein and the cleaved tag was then removed by filtration through a Centricon 10 device (Amicon, Stonehouse, UK).

To raise antiserum, repeat protein L (designated DvA-1L) was electrophoresed on a 20% acrylamide-SDS gel and electroeluted using the Biorad electroelutor apparatus (Bio-Rad). Following elution, SDS was removed by dialysis against 10% methanol for 48 h followed by dialysis against PBS for 48 h. Eluted protein (80 μ g) was emulsified in Freund's Complete adjuvant and injected intramuscularly into a rabbit, followed 4 and 8 weeks later with intramuscular boosts of 80 μ g protein in Freund's Incomplete adjuvant. Blood was collected 10 days after boosting. Antibody was also raised in the same way against the contiguous repeats J-K-L polypeptide (designated DvA-1JKL). DvA-1JKL was not electroeluted from preparative gels, but contaminating low-molecular-mass material was removed by filtration through a Centricon 30 membrane (Amicon).

2.8. Immunoblot analysis

D. viviparus adult soluble extracts were prepared in Tris/proteinase inhibitor buffer as previously described [9] and proteins (70 μ g per track) were electrophoretically separated on 5–25% acrylamide-SDS gels using a Pharmacia GE-2/4LS slab gel apparatus. Molecular masses of parasite proteins were estimated by reference to pre-stained marker proteins (15–216 kDa; Gibco). Separated proteins were blotted onto Hybond-C super nitrocellulose membrane (Amersham) using a Tris/glycine/SDS buffer then screened with antibody as described for library immunoscreening.

2.9. Radioiodination and immunoprecipitation

D. viviparus ES products (10 μ g protein) or purified DvA-1 repeats J-K-L (1 μ g protein) were radioiodinated by the Bolton-Hunter method [14] as previously described for ES antigens [9]. Radiolabelled protein ($(1.5\text{--}2) \times 10^5$ cpm per assay tube) was incubated overnight with 2.5 μ l rabbit antiserum to DvA-1 repeat L or repeats JKL or with bovine serum (normal bovine serum or pooled antisera from three calves experimentally infected on three occasions with 5000 *D. viviparus* L3). IgG immune complexes were precipitated using Protein A as previously described [9] and precipitated protein was fractionated by SDS-PAGE and the gels analysed by autoradiography.

2.10. Passive cutaneous anaphylaxis (PCA)

Duncan-Hartley outbred guinea pigs were shaved on the back and injected intradermally with 100 μ l of test serum diluted in the range 1:2 to 1:512. As we have previously reported a limited and genetically restricted antibody response to *D. viviparus* adult antigens in the guinea-pig host [15], this test serum was a pool from eighteen previously *D. viviparus* infected Duncan-Hartley and inbred (Strains 2 and 13) guinea pigs in an effort to reduce the possibility of genetic unresponsiveness to DvA-1.

24 h after injection of this serum, the guinea pigs were given an intravenous injection of either 50 μ g recombinant protein, 250 μ g *D. viviparus* ES or 250 μ g *D. viviparus* adult PBS-soluble extract each mixed with 5 mg Evans Blue (BDH, Poole, UK) in PBS. The guinea pigs were examined 30 min and 1 h later for the development of reaction spots. Normal serum controls were used in each guinea pig and this assay was carried out on two occasions.

3. Results

3.1. The *dva-1* cDNA sequence

The L3-stage cDNA expression library was screened with rabbit antibody to adult ES products and five of the strongest positive plaques identified

were purified through three rounds of immunoscreening. Restriction mapping indicated that there were two pairs of sibling clones and one unique clone (data not shown), the largest of which, designated pCB10, was chosen for DNA sequencing. The insert DNA, which is 3103 bp in length, was sequenced on both strands either directly on the plasmid or by generating a series of exonuclease III deletions. Translation of the cDNA sequence revealed an open reading frame of 2982 bp followed by a 3' untranslated region (UTR) consistent with oligo(dT) priming having occurred at the poly(A) tail of the mRNA. The common polyadenylation signal sequence AAUAAA was not present in the 3'-UTR, but the sequence AAUGAA is present which has been shown to act as a functional polyadenylation signal in the free-living nematode *Caenorhabditis elegans* [16]. No initiator methionine is encoded,

suggesting that this cDNA does not contain the 5' end of the *dva-1* gene.

In order to obtain information on the 5' region of the cDNA, 5'-RACE was carried out on RNA from adult parasites as described in Materials and methods. This provided a DNA fragment of 1880 bp which was cloned into pT7Blue (designated pJMD36) and sequenced. The 3' sequence of the pJMD36 insert overlaps with 86 nucleotides of the 5' end of pCB10 and no differences were found between them. Together with pCB10, this provides the complete sequence of the DvA-1 cDNA.

A search of the GenEMBL DNA database using the programme FASTA [17] indicated that *dva-1* is similar in sequence to cDNAs *Di5* [2] and *Di22-16* [1] encoding a 15-kDa antigen of *Dirofilaria immitis*, cDNA encoding the gp15/400 protein of *Brugia* species [3], and the ABA-1 allergens of the human



Fig. 2. Alignment of the predicted amino-acid sequence of repeat units of DvA-1 protein. The sequences were aligned using the PILEUP [33] programme, which shows progressive divergence from the first sequence shown, with dots denoting gaps introduced to optimise the alignments. The computer assisted alignments did not align the conserved second cysteine of repeat unit A and some of the putative endoproteinase cleavage sites, for which adjustments were made by hand. The amino acids are numbered relative to the known N terminus of the ABA-1 protein obtained by direct protein sequencing of parasite-derived material [7]. The ABA-1 sequence was not included in the PILEUP alignment but is shown in the bottom line for comparison. The arginine tetrads, presumed to be cleavage sites for subtilisin-like endoproteinases, are shown in bold, as are the conserved tryptophan and cysteine residues.

and pig roundworms *Ascaris suum* and *Ascaris lumbricoides* [4,5]. Like these, *dva-1* encodes a polypeptide composed of tandemly repeated units most of which are approx. 134-amino-acids long (Fig. 1A). The translated sequence of *dva-1* comprises 12 such units, one of which is a short unit of only 71 amino acids. With the exception of two repeat units, there is a consensus proteinase cleavage site at the COOH end of each repeat. The repeated units of the cDNA have been named *a* to *l* in the 5' to 3' direction, as shown diagrammatically in Fig. 1B, and the encoded polypeptides are accordingly named A to L in the N- to C-terminal direction. An alignment of the predicted amino-acid sequence of all the repeat units of the DvA-1 protein with *Ascaris* ABA-1 is shown in Fig. 2. Notable conserved residues include the tryptophan of each repeat unit (amino acid 15, numbered according to the known N-terminal sequence of ABA-1 [7]), the two cysteines (amino acids 64 and 120) and the similar abundance of lysine/arginine and glutamate/aspartate residues in all the aligned sequences.

What is most striking about the comparison of DvA-1 to its homologues in other species is the degree of variation between the units. Table 1 shows the percent identity of DvA-1 polypeptide repeats to one another. Repeats I and J are the most alike, being approx. 98% identical at both the amino acid and DNA levels, and these are also closely related to the short repeat H. At the other extreme are adjacent repeats D and E, which are only 17.6% identical at the amino-acid level.

Each putative repeat unit of the DvA-1 polypeptide, with the exception of repeats D and E, ends in a basic amino-acid-rich tetrad which is a consensus cleavage site for endoproteases of the subtilisin family (K/R.X.K/R.R.) [18]. The amino-acid sequence downstream from these tetrads aligns with the known N terminus of the ABA-1 protein isolated from *A. suum* [7]. Only one consensus N-glycosylation site is present, occurring in the short repeat H, in which serine substitutes for a cysteine expected at this position from the alignment with the other units (indicated in Fig. 2). The open reading frame continues beyond the final repeat unit, extending for a further 36bp of open reading frame before the putative stop codon.

Finally, following the initiator methionine, there is a short hydrophobic-rich sequence of amino acids preceeding the first repeat unit. The bond between the serine residues at positions 21 and 22 in this hydrophobic leader sequence may act as a cleavage site according to the predictions of von Heijne [19].

3.2. Isolation of DvA-1-encoding cDNA clones from an adult *D. viviparus* expression library

To determine whether *dva-1* is expressed in the adult stage of the parasite in the same form, eluted antibody to pCB10 was used to immunoscreen adult clones which had previously been selected by screening with antibody to *D. viviparus* adult ES products. Of 15 clones which were positive with anti-ES antibody, 13 reacted with antibody to the recombinant

Table 1
Identity (%) between individual repeat units of DvA-1 polypeptide

A	B	C	D	E	F	G	H	I	J	K	L
B	18.6										
C	28.8	25.6									
D	26.8	21.7	28.0								
E	18.6	21.7	22.2	17.6							
F	18.4	21.2	19.2	16.3	31.5						
G	24.2	20.8	22.5	19.2	40.5	41.7					
H	25.4	16.2	20.3	19.7	22.5	37.7	49.3				
I	21.9	22.5	24.0	25.6	43.4	46.2	56.0	85.9			
J	21.9	22.5	24.0	25.6	45.0	47.7	56.0	84.5	97.8		
K	22.2	24.4	26.6	20.0	43.5	32.1	49.6	55.7	59.4	59.4	
L	19.7	22.5	28.6	20.8	32.8	46.2	42.1	33.8	49.3	48.5	37.6
A	B	C	D	E	F	G	H	I	J	K	L

Amino-acid sequences of repeat units were compared using the GAP [32] programme.

pCB10 polypeptide. Restriction enzyme digestion (not shown) indicated that these included sibling clones and sequencing of the 3' end of three discrete clones showed that cDNA priming had occurred at the poly(A) tail. The nucleotide sequence of the 3' end of these clones differed from the pCB10 sequence with five non-conservative substitutions being found in one clone over a 280-bp region (see legend to Fig. 1A). In another clone, 340 bp of the 3' end of which were sequenced, there were eight nt differences to pCB10, five of which were in the 3'-UTR. The most likely explanation for these sequence differences is allelic variation in the parasite population although the possibility that there is more than one *dva-1* gene cannot be ruled out (see below).

One clone from the adult parasite cDNA library was sequenced in the region in which the short repeat *h* was expected from pCB10, and an identical short repeat was found to be present.

3.3. Organisation of *dva-1*

The organisation of the coding sequence of *dva-1* in genomic DNA was analysed by Southern blot analysis using a 1066-bp *Bst*XI-*Xho*I fragment from the 3' region of the pCB10 insert as a probe (see Fig. 1B). Cleavage of genomic DNA with *Bam*HI or *Sma*I, which neither cut within the pCB10 insert nor in the 5'-RACE fragment, gave single bands of > 21 kb on hybridisation (Fig. 3A). This could indicate the presence of only one copy of the *dva-1* gene although it is also possible that there are two adjacent genes.

A series of partial digests of genomic DNA was carried out using *Sfa*NI, which cleaves at conserved sites in repeats *h* to *l* of the cDNA, as indicated in Fig. 1B. A ladder-like array of five ascending units with approx. 400-bp increments was observed on hybridisation (Fig. 3B), suggesting that the repeat units which contain the conserved *Sfa*NI cleavage site are arranged in tandem in the genomic DNA, with no intervening sequences. The observation of five repeating units on hybridisation, while four would be predicted from Fig. 1B could be due to the presence of an additional *Sfa*NI site downstream from the *dva-1* gene. Complete digestion with *Sfa*NI gave one band of approx. 400 bp (not shown).

To examine the expression pattern of the *dva-1*

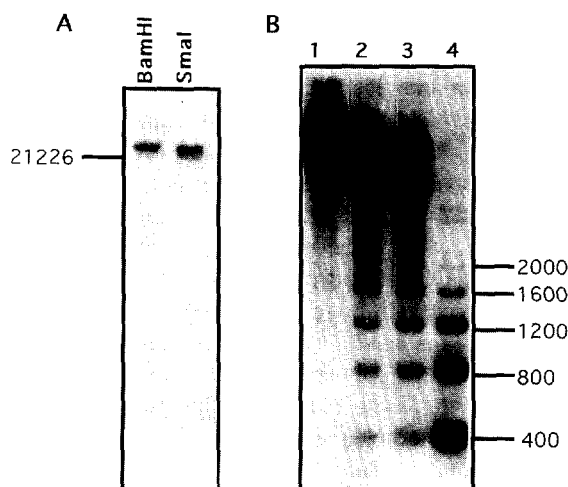


Fig. 3. Restriction analysis of *dva-1*. *D. viviparus* genomic DNA (3 µg) was digested with (A) *Bam*HI or *Sma*I, both at 1 U µg⁻¹ for 4 h or with (B) *Sfa*NI (1 U µg⁻¹) for 1 min (lane 1), 15 min (2), 25 min (3) and 2 h (4). Digested DNA was separated on a 1% agarose gel, blotted onto a nylon membrane and hybridised with a ³²P-labelled 1066-bp *Bst*XI-*Xho*I fragment of pCB10. Approximate sizes are shown in bp.

transcript, RNA from the L3 or adult stages of *D. viviparus* was probed with the 1066-bp *Bst*XI-*Xho*I pCB10 fragment. A band of approx. 5 kb was observed in each blot (data not shown), the intensity of hybridisation being consistently greater in adult RNA than in L3 material for the same quantity of RNA.

3.4. Detection of DvA-1 in somatic and ES compartments of adult *D. viviparus*

DNA encoding polypeptide unit L (DvA-1L) or units J, K and L contiguously (DvA-1JKL) were amplified from clone pCB10 by PCR and expressed in *E. coli* using the pET15b expression vector. Antisera to the two recombinant polypeptides were then used to probe Western blots of somatic extract of adult parasites, or in immunoprecipitation of radio-iodinated ES. The results were confounded by heterophile antibody produced by the rabbits used for the immunisations, but the cross-reactive specificities were identified by comparison with pre-immunisation sera. The *D. viviparus* components most subject to this were those at 20 and 50 kDa (see Fig. 4A for an example of this). Probing adult extract with antibody to the DvA-1L or DvA-1JKL poly-

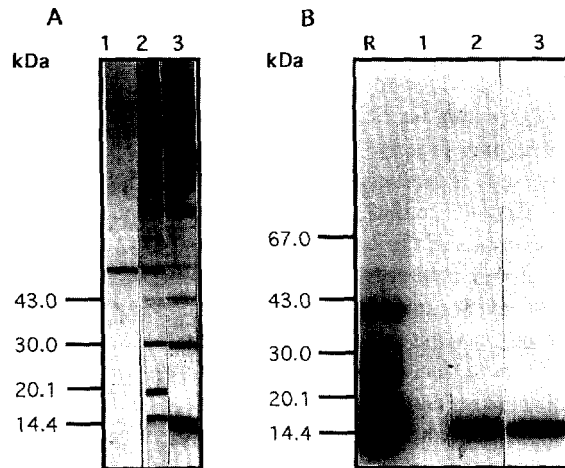


Fig. 4. Antigenicity of DvA-1 polypeptides. (A) Adult extract (70 µg protein per track) was electrophoresed on 5–25% acrylamide gradient-SDS gels under reducing conditions. Proteins were blotted onto nitrocellulose and probed with normal rabbit serum (lane 1), antiserum to DvA-1L (lane 2) or antiserum to DvA-1JKL (lane 3). (B) Radioiodinated adult ES (lane R) was immunoprecipitated with normal rabbit serum (lane 1), anti-DvA-1L (lane 2) or anti-DvA-1JKL (lane 3) antisera. The molecular masses of marker proteins are shown in kDa. All tracks were autoradiographed for the same length of time.

peptides revealed prominent specific bands of approx. 14, 28 and 42 kDa. Antibody to repeats JKL produced doublet bands at these positions while only single bands were observed with antibody to repeat L only.

Immunoprecipitation of radio-iodinated *D. viviparus* adult ES with antiserum to polypeptide DvA-1JKL also identified two antigens of approx. 14–15 kDa, while antiserum to repeat L only identified a single band of 14 kDa (Fig. 4B). Faint precipitation of ES components of approx. 28 and 42 kDa was observed with both antisera.

Using anti-DvA-1JKL serum in immunoblotting, reactivity was also found with 15-kDa components of adult homogenates of the related ruminant strongyles *Haemonchus contortus* and *Cooperia oncophora* (not shown).

3.5. Antigenicity of recombinant DvA-1

We have previously demonstrated an antibody response to a wide range of *D. viviparus* ES prod-

ucts in immunoprecipitation studies using antisera from calves exposed to lungworm infection [9]. In the present study, immunoprecipitation demonstrated strong recognition of recombinant DvA-1 protein by antisera from *D. viviparus*-infected calves, but not by sera from uninfected animals (data not shown). Sera from calves vaccinated with irradiated *D. viviparus* larvae showed only weak recognition of the DvA-1 protein (not shown).

Passive cutaneous anaphylaxis (PCA) studies were carried out in guinea pigs in order to assess whether DvA-1 is allergenic. Normal guinea pigs were injected with serum from previously infected guinea-pig hosts and challenged 24-h later with recombinant DvA-1L, or DvA-1JKL, or control preparations of *D. viviparus* adult extract or ES products. A response was observed with both ES material and parasite extract (both at serum dilutions of 1:512) but with neither of the DvA-1 preparations, using serum dilutions in the range 1:2 to 1:512.

4. Discussion

We present here the complete sequence of the cDNA encoding an immunodominant antigen, DvA-1, of *D. viviparus* which is produced by both infective larvae and adult stages of the parasite, and is detectable in both somatic and ES material. Twelve units of the DvA-1 polypeptide are encoded, whereas approx. 20 NPA units are estimated to be encoded in the genomes of *B. pahangi* [3] and *A. suum* [5], although in these cases the complete genes have not been sequenced. This large size has imposed limitations on the analysis of NPA genes in that only partial cDNAs have been isolated from cDNA libraries to date [1–5]. For DvA-1, a complete description of the transcribed region required a combination of cDNA cloning and 5' RACE analysis. This revealed characteristics common to all known NPAs, but several novel features emerged which are pertinent to our understanding of this class of proteins in general, their processing by the parasites, the immune response to them and their biological function.

The most immediate differences between DvA-1 and its counterparts in other nematode species is that the sequences of its units differ substantially from one another, the identity between complete repeats

ranging from 17.6 to 97.8%. Diversity of this degree has not been found for any other NPA. In those cases where differences have been found, however, the lowest levels of similarity are 80.5% between two repeats in the *D. immitis* gene [1], and approx. 49% between a single unit and all the other units found in *Ascaris* [20]. The only other nematode exhibiting a similar degree of heterogeneity is the free-living nematode *Caenorhabditis elegans* [21].

The NPA polypeptides show no significant sequence similarity to any protein of documented function but multiple alignments of all known NPA sequences have been used to predict a structure for the proteins (Ref. 22 and Kennedy et al., data not shown). It is predicted that the proteins form a four helix bundle similar to carrier proteins such as cytochrome *b*₅₆₂ and the myohemerythrin of marine worms [23]. Circular dichroism has confirmed the predicted degree of alpha-helix in both parasite-derived ABA-1 (Kennedy et al., data not shown) and recombinant DvA-1L expressed in *E. coli* [34].

Apart from *D. viviparus*, the only nematode species for which there is published information on the 3' end of NPA-encoding genes is *A. suum*, in which the open reading frame extends beyond the end of the last unit. A consensus endoproteinase cleavage site immediately precedes this C-terminal extension polypeptide in each case, which may mean that it is removed during post-translational processing. The function of the C-terminal extension of DvA-1 could not be inferred from its sequence, but it is notable that the 12-amino-acid extension of *D. viviparus* bears no resemblance in sequence to the 34-amino-acid extension of *A. suum* [4], making it difficult to argue for a common function. The persistence of the extension polypeptide during processing may explain the doublet bands observed in immunoblotting experiments, although this could also be explained by differential glycosylation or incomplete trimming of peptides following cleavage at the endoproteinase cleavage sites.

Another unexpected feature of the DvA-1 protein is the occurrence of a short repeat encoding only 71 amino acids (unit H) which is itself similar to units I and J. Unit *h* was found in cDNAs isolated from two different libraries, which were constructed from different life cycle stages of the parasite, so it is unlikely to have resulted from an aberration occur-

ring during the cloning procedures. It has been suggested that the central region of the NPA units may be the product of a duplication event [4]. Future analysis and further sequence information on other NPAs may indicate whether the short repeat *h* could represent a primordial NPA protein. Interestingly, the first 49 amino acids of unit H align with the N-terminus of the other DvA-1 repeats, while the last 22 amino acids align most favourably with the C-terminal region of the repeats, although the conserved cysteine residue (Cys 120) is replaced in unit H by a serine, which then forms part of the only consensus *N*-glycosylation site encoded within the DvA-1 sequence. Assuming that unit H polypeptide is functional, it is conceivable that the polypeptide might, like the ROP protein of *E. coli* [24], dimerise to form a functional four helix bundle protein similar to that predicted for the other repeat units. Alternatively unit H may be a non-functional product of unequal homologous recombination, as happens in other head to tail arrays [25]. The close sequence similarity between unit H and the complete units I and J would be consistent with this idea.

In addition to providing information on the complete complement of DvA-1 repeat units, the 5'-RACE analysis of *dva-1* revealed a hydrophobic N-terminal leader to the DvA-1 protein. Identification of similar hydrophobic leader sequences in other NPAs would be instructive in understanding the post- and co-translational events of these proteins.

In all NPAs, the polyprotein is thought to be cleaved at a tetrad of predominately basic amino acids similar to the consensus K/R.X.K/R.R cleavage site for endoproteinases of the subtilisin family [18]. Such sites were found at the expected positions between most of the putative DvA-1 units, although their composition varied within the constraints of the consensus. This variability may not, therefore, have any meaningful effect on the enzymatic cleavage of the polyprotein at these sites. More enigmatic is the apparent lack of a recognised consensus cleavage site between units D, E and F. It may be pertinent here that the gp15/400 protein in *Brugia* species appears to be cleaved from the 400-kDa precursor to a polypeptide of 200 kDa soon after synthesis, followed by further processing to produce 15-kDa subunits [3]. It may be, therefore, that the complete polyprotein is post-translationally processed in dis-

crete stages, first into large segments, then into the single units, and that the amino-acid sequence between DvA-1 polypeptides D, E or F may thus represent the first stage cleavage sites. If so, then different enzymes are likely to be responsible for the discrete stages in the processing. If units D, E and F are not cleaved, then this would be an unprecedented feature amongst known NPA proteins, but might explain the detection of a 42-kDa entity in the immunoblotting using antiserum to DvA-1L (Fig. 4A).

NPA proteins are likely to be important in the immunopathology of nematode infections. The ABA-1 protein of *Ascaris* has been shown to be a major allergen of the parasite [7], while there exists a positive correlation between IgE antibody reactivity to the *Brugia* gp15/400 antigen and disease pathology [8]. We did not detect an IgE response to recombinant DvA-1L and DvA-1JKL polypeptides in infected guinea pigs, but this could be due to genetic restriction of the antibody repertoire to *D. viviparus* antigens in guinea pigs [15], or to a failure of this host (in which the infection terminates at an early developmental stage [26]) to produce IgE to the protein. A proper analysis of the allergenicity of DvA-1 is likely, therefore, to require experimentation in the definitive bovine host, and would be particularly relevant to the pulmonary pathology elicited by the infection [27,28]. Moreover, given that responses to the NPAs of *Ascaris* and *Brugia* are under strong genetic control [7,29,30], then the diversity in DvA-1 polypeptide units may have implications for the immune repertoire to individual units amongst members of outbred host populations.

Finally, no indication of the biological function of the NPAs has been published and data searches for similarities with proteins of documented function have proved fruitless. We have, however, recently obtained evidence that the NPAs of *B. malayi* and *A. suum* can act as high affinity fatty acid and retinoid binding proteins (Ref. 31 and Kennedy et al., data not shown) and that this is also true of the recombinant DvA-1L described here [34]. The amino-acid sequence diversity of DvA-1 units may, therefore, be reflected in qualitative and quantitative differences in their binding function. The production of functional recombinant polypeptides from individual units, as described here, can then provide a

means by which to examine their particular binding properties.

Acknowledgements

This work was supported by a grant from the Wellcome Trust to MWK and from Hoechst Animal Health to M.W.K. and Professor J.L. Duncan of the Department of Veterinary Parasitology, University of Glasgow. J.S.G. was supported by an AFRC Veterinary Fellowship. We are also indebted to Don Donald and Caroline Hunt of Intervet (UK) for the provision of infective larvae of *Dictyocaulus viviparus*, to Heather Spence and Philippe Neuville for technical advice, to Brian Shiels for advice on Northern blot analysis and comments on the manuscript, to David McLaughlin and Ken Bairden for technical assistance, to Chris Franklin for assistance with the figures and to Peter Rickus for the photographic work.

References

- [1] Culpepper, J., Grieve, R.B., Friedman, L., Mika-Grieve, M., Frank, G.R. and Dale, B. (1992) Molecular characterisation of a *Dirofilaria immitis* cDNA encoding a highly immunoreactive antigen. *Mol. Biochem. Parasitol.* 54, 51–62.
- [2] Poole, C.B., Grandea, A.G., Maina, C.V., Jenkins, R.E., Selkirk, M.E. and McReynolds, L.A. (1992) Cloning of a cuticular antigen that contains multiple tandem repeats from the filarial parasite *Dirofilaria immitis*. *Proc. Natl. Acad. Sci. USA* 89, 5986–5990.
- [3] Tweedie, S.A.R., Paxton, W.A., Ingram, L., Maizels, R.M., McReynolds, L.A. and Selkirk, M.E. (1993) A surface-associated glycoprotein (gp15/400) from *Brugia* filarial parasites is composed of multiple tandemly repeated units and processed from a 400kDa precursor. *Exp. Parasitol.* 76, 156–164.
- [4] Spence, H.J., Moore, J., Brass, A and Kennedy, M.W. (1993) A cDNA encoding repeating units of the ABA-1 allergen of *Ascaris*. *Mol. Biochem. Parasitol.* 57, 339–344.
- [5] Spence, H.J. (1994) Molecular cloning and characterisation of the gene encoding the *Ascaris* allergen ABA-1. Ph.D. thesis, University of Glasgow, UK.
- [6] McReynolds, L.A., Kennedy, M.W. and Selkirk, M.E. (1993) The polypeptide allergens of nematodes. *Parasitol. Today* 9, 403–406.
- [7] Christie, J.F., Dunbar, B., Davidson, I. and Kennedy, M.W. (1990) N-terminal amino acid sequence identity between a major allergen of *Ascaris lumbricoides* and *Ascaris suum*

- and MHC-restricted IgE responses to it. *Immunology* 69, 596–602.
- [8] Paxton, W.A., Yazdanbakhsh, M., Kurniawan, A., Partono, F., Maizels, R.M. and Selkirk, M.E. (1993) Primary structure and IgE response to the repeat subunit of gp15/400 from human lymphatic filarial parasites. *Infect. Immun.* 61, 2827–2833.
- [9] Britton, C., Canto, G.J., Urquhart, G.M. and Kennedy, M.W. (1993) Characterisation of excretory-secretory products of adult *Dictyocaulus viviparus* and the antibody response to them in infection and vaccination. *Parasite Immunol.* 15, 163–174.
- [10] Scherrer, K. and Darnell, J.E. (1962) *Biochem. Biophys. Res. Commun.* 7, 486–490.
- [11] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- [12] Ozaki, L.S., Mattei, D., Jendoubi, M., Druible, P., Blisnick, I., Guillotte, M., Puijalon, O. and Da Silva, L.P. (1986) Plaque antibody selection: rapid immunological analysis of a large number of recombinant phage clones positive to sera raised against *Plasmodium falciparum* antigens. *J. Immunol. Methods* 89, 213–219.
- [13] Church, G.M. and Gilbert, W. (1984) Genomic sequencing. *Proc. Natl. Acad. Sci. USA* 81, 1991–1995.
- [14] Bolton, A.E. and Hunter, W.M. (1973) The labelling of proteins to high specific radioactivities by conjugation to a ^{125} I-containing acylating agent. *Biochem. J.* 133, 529–539.
- [15] Britton, C., Canto, G.J., Urquhart, G.M. and Kennedy, M.W. (1992) The antibody repertoire in infection and vaccination with *Dictyocaulus viviparus*: heterogeneity in infected cattle and genetic control in guinea-pigs. *Vet. Immunol. Immunopathol.* 31, 313–322.
- [16] Blumenthal, T. (1993) The *C. elegans* cleavage and polyadenylation signal. *Worm Breeder's Gazette* 13, 62–63.
- [17] Pearson, W.R. and Lipman, D.J. (1988) Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* 85, 2444–2448.
- [18] Barr, P.J. (1991) Mammalian subtilisins: the long-sought dibasic processing endoproteases. *Cell* 66, 1–3.
- [19] von Heijne, G. (1986) A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* 14, 4683–4690.
- [20] Moore, J., Spence, H.J. and Kennedy, M.W. (1993) Extensive diversity in the sequence between individual repeats of the tandemly arrayed gene encoding the ABA-1 allergen of *Ascaris suum*. *J. Cell. Biochem.* 17C, 113 (Suppl.).
- [21] Moore, J., Kennedy, M.W. and Blaxter, M.L. (1993) A homolog of a major class of allergens of parasitic nematodes is encoded in the *Caenorhabditis elegans* genome. *Worm Breeder's Gazette*, 13, 84–85.
- [22] Kennedy, M.W., McCrudden, A.B., Brass, A., Cooper, A., Price, N.C., Kelly, S.M., Christie, J.F., Dunbar, B., Spence, H.J. and Moore, J. (1993) Structural features of the ABA-1 allergen of *Ascaris*. *J. Cell. Biochem.* 17C, 110 (Suppl.).
- [23] Harris, N.I., Presnell, S.R. and Cohen, F.E. (1994) Four helix bundle diversity in globular proteins. *J. Mol. Biol.* 136, 1356–1368.
- [24] Steif, C., Weber, P., Hinz, H.-J., Flossdorf, J., Cesareni, G. and Kokkinidis, M. (1993) Subunit interactions provide a significant contribution to the stability of the dimeric four- α helical bundle protein ROP. *Biochemistry* 32, 3867–3876.
- [25] Branden, C. and Tooze, J. (1991) *Introduction to protein structure*. Garland, New York, NY.
- [26] Wade, A.E., Fox, L.E. and Swanson, L.E. (1960) Studies on infection and immunity with the cattle lungworm *Dictyocaulus viviparus* (Bloch). I. Infection in laboratory animals. *Am. J. Vet. Res.* 21, 753–757.
- [27] Jarrett W.F.H., Jennings F.W., McIntyre W.I.M., Mulligan W., Sharp N.C.C. and Urquhart G.M. (1960) BVA Congress Symposium on Husk. I. The disease process. *Vet. Rec.* 72, 1066–1068.
- [28] Jarrett, W.F.H., McIntyre, W.I.M. and Urquhart, G.M. (1954) Husk in cattle – a review of a year's work. *Vet. Rec.* 66, 665–676.
- [29] Kennedy, M.W., Gordon, A.M.S., Tomlinson, L.A. and Qureshi, F. (1987) Genetic (major histocompatibility complex?) control of the antibody repertoire to the secreted antigens of *Ascaris*. *Parasite Immunol.* 9, 269–273.
- [30] Kwan-Lim, G.-E. and Maizels, R.M. (1990) MHC and non-MHC-restricted recognition of filarial surface antigens in mice transplanted with adult *Brugia malayi* parasites. *J. Immunol.* 145, 1912–1920.
- [31] Kennedy, M.W., Allen, J., Wright, A.S. and Cooper, A. (1995) The gp15/400 polyprotein antigen of *Brugia malayi* binds fatty acid and retinoids. *Mol. Biochem. Parasitol.* 71, 41–50.
- [32] Needleman, S.B. and Wunsch, C.D. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48, 443–453.
- [33] Feng, D.F. and Doolittle, R.F. (1987) Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* 25, 351–360.
- [34] Kennedy, M.W., Britton, C., Price, N.C., Kelly, S.M. and Cooper, A. (In press) The DVA-1 polyprotein of the parasitic nematode *Dictyocaulus viviparus*. *J. Biol. Chem.* 270.