

Smith ScholarWorks

Biological Sciences: Faculty Publications

Biological Sciences

6-15-1991

A Proline-Rich Structural Protein of the Surface Sheath of Larval Brugia Filarial Nematode Parasites

Murray E. Selkirk University of London - Imperial College of Science, Technology and Medicine

Maria Yazdanbakhsh University of London - Imperial College of Science, Technology and Medicine

Danie Freedman Smith College

Mark L. Blaxter University of London - Imperial College of Science, Technology and Medicine

Edith Cookson University of London - Imperial College of Science, Technology and Medicine

See next page for additional authors

Follow this and additional works at: https://scholarworks.smith.edu/bio_facpubs Part of the <u>Biology Commons</u>

Recommended Citation

Selkirk, Murray E.; Yazdanbakhsh, Maria; Freedman, Danie; Blaxter, Mark L.; Cookson, Edith; Jenkins, Rosalind E.; and Williams, Steven A., "A Proline-Rich Structural Protein of the Surface Sheath of Larval Brugia Filarial Nematode Parasites" (1991). Biological Sciences: Faculty Publications, Smith College, Northampton, MA. https://scholarworks.smith.edu/bio_facpubs/36

This Article has been accepted for inclusion in Biological Sciences: Faculty Publications by an authorized administrator of Smith ScholarWorks. For more information, please contact scholarworks@smith.edu

Authors

Murray E. Selkirk, Maria Yazdanbakhsh, Danie Freedman, Mark L. Blaxter, Edith Cookson, Rosalind E. Jenkins, and Steven A. Williams

A Proline-rich Structural Protein of the Surface Sheath of Larval *Brugia* Filarial Nematode Parasites*

(Received for publication, August 20, 1990)

Murray E. Selkirk‡§, Maria Yazdanbakhsh‡¶, Daniel Freedman||, Mark L. Blaxter‡, Edith Cookson‡, Rosalind E. Jenkins**, and Steven A. Williams||

From the ‡Wellcome Research Centre for Parasitic Infections, Department of Biochemistry, Imperial College of Science, Technology and Medicine, London SW7 2AZ, United Kingdom, *Department of Biological Sciences*, Smith College, Northampton, Massachusetts 01063, and the **Electron Microscopy Unit, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 6BT, United Kingdom

Both cDNA and genomic DNA sequences have been isolated which encode a proline-rich precursor protein of the sheath from microfilariae, the first stage larvae of the filarial nematode parasites Brugia pahangi and Brugia malayi. This 22-kDa protein is soluble only under reducing conditions and is extensively crosslinked by both disulfide and nonreducible bonds. Immunogold electron microscopy shows that the protein is localized exclusively in the sheath, a vestigial remnant of the eggshell, which is retained by and encloses the mature microfilaria. Analysis by Western blotting confirms that the protein is expressed only in microfilariae and adult female worms, although transcripts are detectable only in adult females. The deduced amino acid sequence contains a short N-terminal hydrophobic putative leader sequence, a central repetitive domain that contains 14 copies of a degenerate 5amino acid repeat with the consensus sequence Met-Pro-Pro-Gln-Gly, and a C-terminal proline-rich domain flanked by clusters of cysteine residues. These clusters can be aligned with cysteine residues implicated in cross-linking of a family of cuticular collagens originally identified in Caenorhabditis elegans but which extends to other nematodes.

The majority of free-living nematodes such as *Caenorhabditis elegans* and many parasitic species (*e.g. Ascaris*) give rise to progeny encased in eggshells, one of the most resistant and impermeable biological structures (Croll and Matthews, 1977). In contrast to this general plan, filarial nematode parasites are ovoviviparous, giving birth to living first stage larvae termed microfilariae. The reason for this phenomenon lies in the fact that progeny are not voided into the environment but are released into the bloodstream, where they are taken up by feeding arthropods which act as vectors for the parasite. Within the filariae, a dichotomy exists in which some species (*e.g. Brugia*) release microfilariae which are enclosed by a loose baglike structure termed a sheath. Ultrastructural studies indicate that the sheath is derived from the embryonic eggshell (Rogers *et al.*, 1976; Zaman, 1987), but the reasons for the conservation of this structure in larvae of some species and not others are unclear.

Although mechanisms of immunity against nematode parasites are relatively obscure, there is a clear positive correlation of antibodies to the microfilarial sheath and clearance of larvae from the circulation (Wong and Guest, 1969; McGreevy *et al.*, 1980; Pinder *et al.*, 1988). Interest in the detailed structure of the sheath is thus not purely academic, but little is known about its biochemical composition other than that it contains protein, carbohydrate (Furman and Ash, 1983; Sayers *et al.*, 1984), sulfated proteoglycans (Simpson and Laurence, 1972), and chitin (Fuhrman and Piessens, 1985).

The general plan of a nematode eggshell is that of a multilamellate structure, which contains a central chitin-protein complex, an inner lipid layer, and outer layers derived from the vitelline membrane of the fertilized oocyte and/or secretions from uterine epithelial cells (Wharton, 1980). Given this general knowledge of the morphology of the eggshell, it is surprising that the primary structure of an eggshell protein has not been described in any nematode species, even in one so comprehensively studied as *C. elegans*. In this report, we describe a major proline-rich protein of the sheath/eggshell of *Brugia pahangi* and highlight homologies in primary sequence that point to a similar pattern of disulfide crosslinking utilized by cuticular collagens of *C. elegans*.

EXPERIMENTAL PROCEDURES

Parasites

Adult Brugia malayi of the zoophilic biotype (Partono and Purnomo, 1987) were recovered from the peritoneal cavity of jirds (Mongolian gerbils; Meriones unguiculatus) infected more than 3 months previously with 200 infective larvae from Aedes aegypti mosquitoes. Microfilariae were harvested from peritoneal cavities and separated from host cells by passage through Sephadex G10 (PD10) columns equilibrated with RPMI 1640 medium preheated to 37 °C (Taylor et al., 1984). Infective larvae (L3) were dissected from mosquitoes and washed free of host contaminants. Adult B. pahangi were a kind gift from Dr. D. A. Denham (London School of Hygiene and Tropical Medicine).

Antiserum Preparation

The β -galactosidase fusion protein expressed by cDNA clone Bpa22/7 was induced and purified as previously described by electroelution from preparative SDS-polyacrylamide gels (Selkirk *et al.*, 1989a). A polyclonal antiserum was raised in rabbits by intramuscular inoculation of 200 μ g of purified protein emulsified in Freund's

^{*} This work was supported by the Medical Research Council, the Wellcome Trust, the Filariasis component of the United Nations Developmental Program/World Bank/World Health Organization Special Program for Research and Training in Tropical Diseases, and the Commission of the European Communities. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EMBL Data Bank with accession number(s) X58063.

[§] To whom correspondence should be addressed.

[¶] Current address: Dept. of Parasitology, University of Leiden, Postbus 9605, 2300 RC Leiden, Netherlands.

complete adjuvant. Four weeks later, the rabbit was boosted via a subcutaneous injection of 50 μ g of protein emulsified in Freund's incomplete adjuvant, and the rabbit was bled 7 days later.

One- and Two-dimensional SDS-PAGE and Western Blotting

Discontinuous (7-25%) SDS-PAGE was performed according to standard protocols. Two-dimensional electrophoresis was performed according to O'Farrell et al. (1977) and as previously described (Selkirk et al., 1989b). For one-dimensional gels, the solubility of Mf22 was examined by preparing parasite extracts via sequential homogenization in buffer a (TBS,¹ 1.5% n-octyl glucoside plus a mixture of protease inhibitors: 1 mM EDTA, 1 mM EGTA, 0.2 mM 1-chloro-3tosylamido-7-amino-2-heptanone, 0.1 mM L-1-tosyl-amido-2-phenylethyl chloromethyl ketone, 1 mM N-ethylmaleimide, 2 mM phenylmethanesulfonyl fluoride). Following extraction on ice for 1 h, homogenates were centrifuged at $10,000 \times g$ for 30 min. Insoluble pellets were extracted in buffer b (the same buffer, with the addition of 5% 2-mercaptoethanol (2-ME) for 1 h at 37 °C) before centrifugation under the same conditions and recovery of the supernatant. For twodimensional gels, parasites were homogenized in 100 mM DTT, the first dimensions run in 50 mM DTT, 6 M urea, and the second dimensions in 100 mM DTT.

Western blotting was essentially according to Towbin *et al.* (1979), with details as described previously (Selkirk *et al.*, 1989b). Specifically, following electrophoretic transfer of proteins to nitrocellulose, the blot was blocked via a 1-h incubation in 20 mM Tris-HCl, pH 8.0, 0.15 M sodium chloride (TBS) plus 5% dried milk powder. The primary antiserum (rabbit anti-Bpa 22/7 β -galactosidase fusion protein) was used at a dilution of 1/1000 in TBS, 5% milk powder, 0.1% Tween 20 overnight at 4 °C, and horseradish peroxidase-conjugated anti-rabbit IgG (Bio-Rad; 172-1013) was used at 1/2000 for 2 h at room temperature in TBS alone. Binding was visualized by 4-chloro-1-naphthol (Sigma, C 8890) at a final concentration of 0.5 mg/ml.

Northern Blotting and Hybridization

RNA was isolated by sedimentation through cesium chloride following lysis of parasites in guanidinium thiocyanate (Chirgwin *et al.*, 1979), and separated on formaldehyde-based gels ($15 \ \mu g$ /track; Maniatis *et al.*, 1982); DNA probes were labeled by random priming (Feinberg and Vogelstein, 1983). Hybridizations were performed in 5 × SSC at 65 °C with 10⁶ cpm/ml, and blots were washed at a final stringency of 0.1 × SSC at 65 °C.

Immunoelectron Microscopy

Fixation and Embedding-Fixation and embedding were carried out via a modification of the method of McPhail et al. (1987). Adult female parasites were cut into 0.5-cm lengths and were subsequently treated in an identical manner to microfilariae. The parasites or parasite pieces were fixed in 2% paraformaldehyde, 0.1% glutaraldehyde in 0.1 M phosphate buffer for 1 h at 4 °C, followed by an overnight wash in 0.1 M phosphate buffer at the same temperature. The tissue was then dehydrated in methanol by successive incubations of 30 min in 50% methanol at 4 °C and 60 min each in 70, 90, and 100% methanol at -25 °C. The material was then infiltrated with LR Gold (London Resin Gold, London Resin Co.), first by incubation for 30 min in 1:1 methanol:LR Gold, 0.1% benzoin methyl ether (LRG/BME), then by a 1-h incubation in 3:7 methanol:LRG/BME, and finally by a 72-h incubation in LRG/BME. After embedding in gelatin capsules with fresh resin, polmerization was effected by exposure to UV light for 72 h. All infiltration and polymerization steps were carried out at -25 °C.

Immunogold Localization—Ultrathin sections (80–90 nm) of embedded worms (adult females or microfilariae) were cut on a Reichert Jung ultracut ultramicrotome and mounted on Formvar-coated nickel grids. Tissue sections were blocked for 30 min at room temperature in PBS, 1% bovine serum albumin, 0.01% Tween 20 (PBS/BT). Grids were then incubated in rabbit anti-Bpa 22/7 β -galactosidase, diluted 1/1000 in PBS/BT for 2 h at room temperature, washed three times in PBS/BT, and incubated for 30 min in goat anti-rabbit IgG-gold conjugate (Seralab, 10 nM gold particles) diluted 1/10 in the same solution. Samples were washed once in PBS/BT, once in PBS, 0.01% Tween, and once in distilled water prior to counterstaining in 4% uranyl acetate for 10 min, followed by Reynold's lead citrate for 2 min. Control grids were treated with normal rabbit serum diluted 1/ 30 in PBS/BT, in place of the primary antibody. Processed samples were examined in a Jeol 100CX transmission electron microscope.

RESULTS

Mf22: Stage Specificity, Solubility, and Localization-We have recently described the physical properties, localization, and turnover of Gp29, the major surface glycoprotein of adult stage parasites of the genus Brugia, and raised a polyclonal antiserum to it via preparative SDS-PAGE of a 29-kDa fraction (Maizels et al., 1989; Selkirk et al., 1990). In using this antiserum to screen a cDNA library derived from mixed adult B. pahangi RNA and propagated in $\lambda gt11$, we isolated 28 clones, of which 14 represented a related family judged by cross-hybridization. The largest cDNA from this family (Bpa22/7; 649 base pairs) was lysogenized in E. coli Y1089 (Young and Davis, 1983), and the β -galactosidase fusion protein induced and purified by preparative SDS-PAGE as previously described for other gt11 clones (Selkirk et al., 1989a). A polyclonal antiserum was raised in rabbits by immunization with the purified fusion protein in order to determine the identify of the native protein which this class of cDNAs encoded.

Surprisingly, the antiserum bound to a series of proteins expressed in first stage larvae (microfilariae) and adult female worms but not in third stage (infective) larvae or adult males (Fig. 1). Thus the cDNAs clearly did not code for Gp29, but for a protein or series of proteins which, in addition to their distinct stage specificity, were insoluble in detergent (n-octyl glucoside) and only soluble in reducing agents such as 2-ME or DTT. The antiserum bound to a number of proteins in microfilariae which resolved as apparent multimers of 22 kDa, and this was confirmed by Western blots of parasite extracts separated by two-dimensional gels. A series of proteins could now be visualized whose molecular mass resolved as multiples of 22 kDa up to 88 kDa, and had a number of distinct isoelectric points (Fig. 2). One obvious interpretation was that a monomeric 22-kDa protein was being cross-linked and modified to form a network of higher order products in microfilariae, and this was supported by the resolution of a major species at 22 kDa in adult females. The higher molecular mass (approximately 29 kDa) band seen in female worms in Fig. 1 was a product of incomplete reduction and was abolished

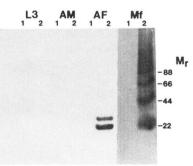


FIG. 1. Stage specificity of expression of Mf22. Extracts from third stage larvae (L3), adult males (AM), adult females (AF), and microfilariae (Mf) of B. malayi were run on discontinuous (7-25%) polyacrylamide gels prior to transfer to nitrocellulose and reaction with a rabbit antiserum generated to the β -galactosidase fusion protein expressed by cDNA clone Bpa22/7. Extracts were prepared by sequential solubilization of parasites in n-octyl glucoside (1), and 2-ME (2) (see "Experimental Procedures" for details). The position of migration of the 22-kDa monomer and polymers is indicated.

¹The abbreviations used are: TBS, Tris-buffered saline; PBS, phosphate-buffered saline; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; 2-ME, 2-mercaptoethanol; DTT, dithiothreitol; PRP, proline-rich protein; EGTA, [ethylene bis(oxyethylenenitrilo)]tetraacetic acid.

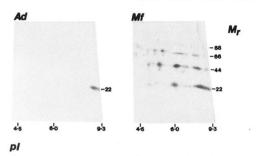


FIG. 2. Detection of cross-linked Mf22 in microfilariae. Extracts of adult females (Ad) or microfilariae (Mf) of *B. malayi* were separated by two-dimensional polyacrylamide gel electrophoresis, blotted, and reacted with the Bpa22/7 β -galactosidase fusion protein antiserum. Extracts were prepared by boiling in loading buffer (see "Experimental Procedures"). The migration of apparent monomers, dimers, trimers, and tetramers of Mf22 is indicated (Mr). pI, isoelectric point.

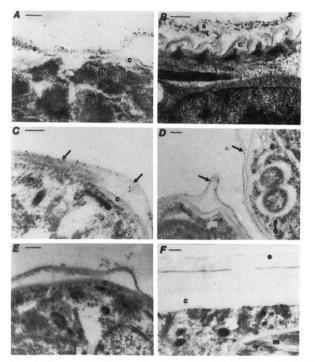


FIG. 3. Immunogold localization of Mf22 in *B. malayi*. Bar = $0.25 \ \mu$ m. *A*, mature microfilaria stained with Bpa22/7 β -galactosidase antiserum, showing exclusive localization of the gold particles to the sheath; *B*, lack of staining of mature microfilaria with control (normal) rabbit serum; *C* and *D*, positive staining of the sheath in uterine microfilaria with the test serum; *E*, lack of staining of microfilaria *in utero* with control serum; *F*, lack of staining of an adult male worm with the test serum. Section through the body wall of the worm. Abbreviations: *s*, sheath; *c*, cuticle; *e*, epicuticle; *h*, hypodermis; *m*, muscle.

when extracts were boiled extensively (10 min) in 5% 2-ME or 100 mM DTT.

The localization of the native protein or protein complexes in microfilariae and adult worms was examined by utilizing the same antiserum generated to the Bpa22/7- β galactosidase fusion protein in immunogold electron microscopic studies. We were surprised to find that the antiserum bound solely, and at high density, to the surface sheath surrounding microfilariae, either as "mature" first stage larvae isolated from the peritoneal cavity of infected jirds, or as immature larvae developing *in utero* in sections of adult female worms (Fig. 3). Thus, Fig. 3A shows gold particles concentrated on the sheath in mature microfilariae, while C and D depict staining on the same structure in developing larvae *in utero*. No staining was evident in any internal structures such as the cuticle, hypodermis, or body-wall musculature. Additionally, the antiserum did not bind to adult male worms (Fig. 3F), confirming the stage-specificity of expression seen in Western blots (Fig. 1). Fig. 3, *B* and *E* shows the lack of binding of control sera to microfilariae, even when used at a much higher concentration, thus confirming the specificity of the reaction seen with the test antiserum. Because of the molecular mass of the native protein and its localization to the microfilarial surface sheath, we have termed this protein Mf22 for convenience.

A Northern blot using the Bpa22/7 cDNA as a hybridization probe indicates that the Mf22 gene was not transcribed in microfilariae (Fig. 4), despite the fact that the protein was present in this stage at a high concentration on the sheath (Figs. 1–3). It thus appears that Mf22 was presynthesized in adult female worms and did not undergo significant turnover in microfilariae, consistent with a role as a structural protein. The size of the major transcript (0.9 kilobase pairs) is also consistent with a final protein product of 22 kDa, allowing for 5'- and 3'-untranslated regions.

DNA Sequence, Protein Sequence, and Structural Homologies—The cDNA clone Bpa22/7 was used as a hybridization probe in a series of restriction enzyme digests of genomic DNA from B. malayi and B. pahangi, which suggested that it was encoded by a single gene which resided on a HinPI restriction fragment of approximately 1370 base pairs. Fig. 5 illustrates the sequence of the Mf22 gene from B. pahangi, including 5'- and 3'-flanking regions, derived from three partial cDNAs and the complete sequence of the HinPI genomic fragment. The 5' end of the longest cDNA (Bpa22/7) lies at residue 23 (arrow), and this clone contains an uninterrupted reading frame through to an 18-base pair poly(A) stretch at position 749. There is no consensus polyadenylation site within Bpa22/7, but two can be located in tandem approximately 80 base pairs downstream from the 3' end (underlined), indicating that the cDNAs have been primed from the internal poly(A) tract of 20 residues seen from positions 732 to 751.

The putative initiator ATG lies 23 residues upstream from the 5' end of Bpa22/7. The sequence of this and two other shorter cDNA clones precisely matches that of the corresponding region of the genomic HinPI fragment, except for

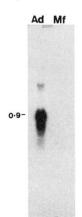


FIG. 4. Stage specificity of transcription of Mf22. Total RNA from adult females (Ad) and microfilariae (Mf) of B. pahangi was size-separated on a 1.5% agarose gel (15 μ g/track), blotted onto nitrocellulose, and hybridized with cDNA clone Bpa22/7. The post-hybridization wash was performed at 65 °C and 0.1 × SSC. The size of the major transcript is given in kilobases. The loadings of RNA were checked by rehybridization of the same filter with a filarial heat shock protein 70 probe (Selkirk *et al.*, 1989a).

HinPI

HinPI	
-390 GCGCATTTTTTCATAAATCAAATAAATGCGTCTACTTAGAGCATGTTAAGTAGAATAGCAGAA -324 GAATATGGTTTGATAGCCAACAATTAAGCGTATCAATGGAAGAAATCATTCAGTAATGAATG	
GAATATGGTTTGATAGCCAACAATTAAGCGTATCAATGGAAGAAATCATTCAGTAATGAATG	
TATTTGTGTTTGATAAATATATATATATAAATTGATAATA	
асттттааасатстоттттаатаааттттсатсттсааттораатттораотттааааастттттттттт	
AAATATT 5' end of longest CDNA	
↓ <u>Hin</u> dIII 1 ATG ATG ATT TCA GTT ATG TGT TGT AAG CTT ATT CCT TCA TTT TGC ATG CTT TTA 54	
1 Met Met Ile Ser Val Met Cys Cys Lys Leu Ile Pro Ser Phe Cys Met Leu Leu 18	
Putative initiation codon * * *	
TCG GTA GTA AAT GCA GTG CAT TTG GGA CAT TAT GGT CCG CAA GAA GAA CAT TCT CAG GAT 114 Ser Val Asn Arg Val His Leu Gly His Tyr Gly Pro Gln Glu Glu His Ser Gln Asp 38	
Ser val val Asn Arg val His Leu Giy His Tyr Giy Pro Gin Giu Giu his Ser Gin Hsp 50	
GTG CAA CCA CAA GAA ATG CAA CGG AAA AGT ATA AAA TTA CTA CGA ATG GGA CCG CAG ATT 174	J
Val Gln Pro Gln Glu Met Gln Arg Lys Ser Ile Lys Leu Leu Arg Met Gly Pro Gln Ile 58	
TE CCA CCC CAN GET ATE CCA CCA CAN GET ATE CCA CCA CAN GET ATE CCA CCA CAN GET 234	1
ATG CCA CCG CAA GGT ATG CCA CCA CAA GGT ATG CCA CCA CAA GGT ATG CCA CCA CAA GGT 234 Met Pro Pro Gln Gly Met Pro Pro Gln Gly Met Pro Pro Gln Gly Met Pro Pro Gln Gly 78	
ATG CCA CCA CAA GGC ATG CCA CCA CAA GGC ATG CCA CCA CAA GGC ATG CAA CCA CAA GGT 294	
Met Pro Pro Gln Gly Met Pro Pro Gln Gly Met Pro Pro Gln Gly Met Gln Pro Gln Gly 98	
ATG CAA CCA CAA GGT ATG CAG CCA CAA GGT ATG CGA CCA CAA GGT ATG CAA CCG CAA TAC 354	1
Met Gln Pro Gln Gly Met Gln Pro Gln Gly Met Arg Pro Gln Gly Met Gln Pro Gln Tyr 118	
ATC GGA CCA CAA G 367	
Ile Gly Pro Gln 122	
aplice donor 78bp intron splice acceptor	
splice donor 78bp intron splice acceptor GTAATTAATTATGATTACACATTTGTATCCACAATGCTTACTTCTTACAGTCTTCAGTCATCAACTGA <u>TTTCAG</u> 44	5
CT GTT GAT AAA AAG TGT GCT GGC TGC ATA ATA AAT ATA AAT TGT GGC 493	
Ala Val Asp Lys Lys Cys Ala Gly Cys Ile Ile Asn Ile Asn Cys Gly 138	,
GGT GCA GAT TGT ATG CCA ACG CTA ACA CAA CAA ACA CCA ACA CCA CCG ATT TGG ACG TTA 55	2
Gly Ala Asp Cys Met Pro Thr Val Thr Gln Gln Thr Pro Thr Pro Pro Ile Trp Thr Leu 15	8
*	
CCA CCT ACG CAG ACA CCA GGA TGG ACT CCT GGA CCA CCG CTA ACT CCA AAA CCA ACG GCA 61	2
Pro Pro Thr Gln Thr Pro Gly Trp Thr Pro Gly Pro Pro Leu Thr Pro Lys Pro Thr Ala 174	
CCT CGC CAT GTG ACG CCA GGA GGT TGT CGT TTG TGT CCA TGT TAC ATA CCT CCT TGT 67	
Pro Pro His Val Thr Pro Gly Gly Cys Arg Leu Cys Pro Cys Tyr Ile Pro Pro Pro Cys 19	8
-	
CAG ATA TGT CAA CCA TGT CAA TGA 696	
Gln Ile Cys Gln Pro Cys Gln *** 205	
* *	
cDNAs terminate in this poly(A) stretch \downarrow GAACTATGACCACTTCATTCAACCATAGCTAAAACAAAAAAAA	13
AAAAAAAAACAATTTTCTATTTCATTATTATTCCTCAGTAAAAGACTTAGTAAAATTTCTTGCAAACATATAGCAAAAGAC 82	-
AGTTTTAATAAAATAAATTGTGAGATTGAAAAATTCAATTCATTACGATCAATGTTTTGGTCTGAAGTCAGCAAAGAT 90)1
AATTGGAAACAAATTTTGCTCATAAAGTATAATATTAGATTTATCTTAGATTATTAAACTATCCTAGAAATGGCGCG 97	18
HinPI	
FIG. 5. Coding sequence and flanking regions of the Mf22 gene from <i>B. pahangi</i> . Composite sequence	
derived from analysis of a <i>Hin</i> PI genomic DNA fragment and the cDNA clone Bpa22/7. Nucleotide and amino acid residues are <i>numbered</i> from the presumed initiator methionine, and the 5' and 3' termini of Bpa22/7 are	
indicated by arrows. The position of splice donor and acceptor sites flanking the single intron are marked and the	

acid residues are *numbered* from the presumed initiator methionine, and the 5' and 3' termini of Bpa22/7 are indicated by *arrows*. The position of splice donor and acceptor sites flanking the single intron are marked, and the 5-residue repeats in the central domain are *boxed*. Cysteine residues are marked with an *asterisk*.

the presence of a 78-base pair intron, which is bounded by consensus splice signals that match those determined for another *Brugia* gene (Perrine *et al.*, 1988). This intron lies precisely at the 3' end of a series of 14 tandem 15-base pair repeats. The repeat is not homogeneous, and the relative identity of the most frequent variant (ATG CCA CCA-CAA GGT) falls off towards the 5' and 3' ends of the repeat block. In addition, two degenerate copies of the repeat can be discerned, in phase, upstream of a 15-base pair region with little homology to the most common repeat variant.

All of the immunoreactive cDNAs are fused with β -galactosidase in the same reading frame, which is also represented in Fig. 5. This reading frame is open for only 18 codons upstream of Bpa22/7, the longest cDNA. The first in-frame ATG (in fact a double ATG) lies 7 codons upstream of the 5' end of 22/7, and the derived sequence predicts a protein with a molecular mass of 21.5 kDa. This protein can be conceptually divided into three domains (Figs. 5 and 6). The N terminus contains a stretch of markedly hydrophobic residues, which constitute a potential signal sequence. The central repetitive domain contains 14 copies of a degenerate repeat with the consensus sequence Met-Pro-Pro-Gln-Gly, but elements of this repeat can be detected from residues 20 to 122, at precisely which point the intron is inserted. Following the intron is a stretch of 82 amino acids that is rich in proline residues (30%).

The derived protein sequence of Mf22 was used to search the NBRF/University of Wisconsin Genetics Computer Group databases using the Lipman and Pearson (1985) algorithm in the FASTA program, and considerable homology was found to a diverse array of proteins which share with Mf22 the property of having proline-rich repetitive elements. The mammalian proline-rich proteins (PRPs) constitute a set of peptides found in saliva and salivary glands. The mouse PRPs are made up of tandemly repeated 14- or 19-amino acid sequences rich in glycine, glutamine, and proline (Clements et al., 1985; Ann et al., 1988). The human PRPs have a more complex repeat structure, with a series of subrepeats made up predominantly of these three residues (Wong et al., 1979; Wong and Bennick, 1980; Maeda et al., 1985). The sequence of Mf22 compared with that of the human PRPs A and C (Wong and Bennick, 1980) showed short stretches of high homology (16/22 identical residues; 72%) ranging to longer stretches of weaker similarity (31/64 identical residues; 48%).

Examples of other proteins with similar proline-rich repetitive elements include giant octopus rhodopsin, which has a 5-amino acid N-terminal repeat (Gln-Gly-Tyr-Pro-Pro) not found in other rhodopsins (Ovchinnikov *et al.*, 1988). Bindin, a protein from the sperm of the purple sea urchin implicated in the recognition of the egg by sulfoglycan binding, has a 7amino acid repeat (Pro-Gln-Gly-Met-Gly-Pro) at its N terminus (Gao *et al.*, 1986). It is interesting to note that all of these repeats are somewhat redundant. Thus individual members may diverge by up to 30%, but the substituted residues are often still glycine, glutamine, or proline. The same situation is observed in the repetitive elements of Mf22, and we elaborate on this in the discussion.

Repetitive sequences are common in many fiber-forming proteins such as keratin (Steinert *et al.*, 1983) and elastin (Bressan *et al.*, 1987), and the primary sequence of the repetitive region of Mf22 is most similar to that of collagens, containing glycine and proline, albeit in arrays of 5 amino acids rather than the 3 amino acids necessary for the triplehelical formation of collagen fibrils (Bornstein and Traub, 1979).

The C-terminal proline-rich domain of Mf22 is flanked by

two regions with clusters of cysteine residues (Fig. 6). This arrangement is reminiscent of the primary structure of cuticular collagens from C. elegans and Haemonchus contortus in which the proline-rich (Gly-Xaa-Yaa) repeat domains are flanked by clusters of cysteines (Kramer et al., 1982; von Mende et al., 1988; Shamansky et al., 1989). The similarity is heightened by the fact that (a) both classes of proteins are disulfide-cross-linked and (b) disregarding residues in potential signal sequences, all of the cysteines in both molecules are clustered at these sites. A closer examination of these cysteine clusters reveals an exact alignment of the 3 residues at site 2 in the C. elegans collagen family, and 3 of the cysteine residues at site 2 in Mf22 (Fig. 6). In addition, several other amino acids show alignment in this region, notably 3 prolines proximal to one of the cysteine residues. The clustered cysteines of cuticular collagens are presumed to be essential for cross-linking, and LRol (left roller) phenotypes of C. elegans have been shown to result from mutations in sqt-1, a collagen gene, which result in the loss of cysteine residues. Conversely, RRol (right roller) phenotypes result from the gain of a cysteine in *sot-1*, and collectively the data indicate that precise disulfide bond formation between collagens is essential for the correct formation of the cuticular exoskeleton (Kramer et al., 1990). The observed alignment of clustered cysteine residues suggests that a similar pattern of intermolecular disulfide bridges are utilized to form a lattice of Mf22 proteins in the sheath of Brugia.

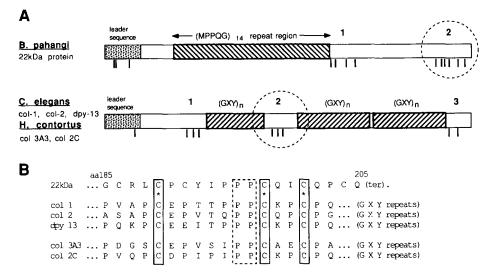
DISCUSSION

When released into the environment, the nematode egg represents a highly impervious self-contained unit which is resistant to dessication, and the success of nematodes in colonizing a wide array of environmental niches must have been aided greatly by the development of the eggshell. The composition of nematode eggshells is highly variable, but a prototype structure has been outlined by Anya (1976) and Wharton (1980), which consists of an inner lipid layer, a central chitinous layer, and an outer lipoprotein or "vitelline" layer. The general impermeability of nematode eggshells is due to the inner layer, which is composed of high molecular weight glycosides termed "ascarosides" due to their isolation from *A. lumbricoides* (Tarr and Fairbairn, 1973).

Filarial nematode parasites clearly have no need for a thick impervious eggshell, and ultrastructural studies have failed to identify an inner lipid layer (McLaren, 1972; Laurence and Simpson, 1974). Tarr (1973) also found no ascarosides in eggshells of the ovoviviparous free-living nematode *Panagrellus redivivus*, and thus the environment in which the embryo develops probably determines whether this layer is formed or not. Consistent with the absence of the inner lipid layer, the permeability of filarial sheaths contrasts sharply with that of "classical" nematode eggshells, and fluorescein-tagging experiments have shown that proteins up to 45 kDa in mass can penetrate the sheath of *B. pahangi* (Devaney, 1985) and *B. malayi* (quoted in Schraermeyer *et al.*, 1987).

The central layer of nematode eggshells is frequently composed of a chitin-protein complex providing a structural backbone through the high tensile strength of chitin dispersed in the form of fibrils in a protein coat. Fibers have been described in *Trichuris suis* (Wharton, 1979) and *Porrocaecum ensicaudatum* (Wharton and Jenkins, 1978) which contain a 2.8 nM chitin microfibril core surrounded by a protein coat, an organization which had been previously observed in arthropod cuticles (Neville, 1975). The presence of chitin in the sheath of *B. malayi* has been inferred from positive staining of microfilariae with fluoresceinated lectins (*Solanum tuberosum*

FIG. 6. Position of clustered cysteine residues in nonrepeat regions of Mf22 and comparison with cuticular collagens from C. elegans (col 1, col 2, dpy 13) and H. contortus (col 3A3, col 2C). A, cartoon of the primary structure of Mf22 and C. elegans collagens, with putative leader sequences and repeat domains shown in stippled and diagonal shading, respectively. The position of cysteines is indicated by vertical bars below the protein. B, alignment of cysteine residues (boxed) from C. elegans and H. contortus collagens (site 2; circled) with those from the C terminus (site 2) of B. pahangi Mf22.



agglutinin and Laburnum alpinum agglutinin) which bind chitobiosyl and chitotriosyl residues, $\beta(1-4)$ linked oligomers of N-acetylglucosamine (Fuhrman and Piessens, 1985). Moreover, the same authors observed that female worms incubated with diflubenzuron, an inhibitor of chitin synthesis, produced microfilariae with truncated sheaths. These data are by no means definitive, however, as the sheath is a very permeable structure, and fluoresceinated lectins might either bind to the cuticle or accumulate in the space between cuticle and sheath. Schraermeyer et al. (1987) confirmed the presence of terminal N-acetylglucosamine (GlcNAc) residues in the sheaths of uterine microfilariae via binding with colloidal gold-conjugated wheat germ agglutinin but observed a progressive depletion of binding such that mature blood stage larvae were negative. The early assumption that chitin is the principal constituent of nematode eggshells (Chitwood and Chitwood, 1950) is clearly not always the case, and some species may possess none at all (Rogers, 1962). Given that the role of chitin is to provide rigidity, there would not seem to be an overwhelming need for it in the filarial sheath, a relatively flexible structure which allows free movement of the enclosed larva.

Although Mf22 shows a superficial similarity to trematode eggshell and silkworm chorion proteins in containing a large domain of amino acid repeats containing glycine, it differs in possessing a very low content of tyrosine, which in the latter proteins are oxidized to highly reactive DOPA quinones implicated in "tanning" or cross-linking of the eggshell (Cordingley, 1987). It is clear that nonreducible cross links are also formed between molecules of Mf22. The nature of these is as yet unknown, but they appear to form progressively as microfilariae mature, judging from the lack of higher molecular mass products in adult female worms (Figs. 1 and 2). This makes sense in that the sheath must expand during development to accommodate the growing larva, which progressively uncoils within it, and thus the sheath must possess a degree of deformability. When the larva (microfilaria) is mature, covalent cross-links appear to form between molecules of Mf22 which presumably stabilize this structure without affecting its permeability. The absence of extensive tanning would allow flexibility, which is necessary as the sheath must accommodate a highly motile larva.

We do not believe that there is a direct functional or evolutionary relationship between Mf22 and the proline-rich proteins highlighted previously, despite the fact that an antiserum to the human PRP-C (a kind gift from Dr. A. Bennick, Department of Biochemistry, University of Toronto) binds weakly to Mf22 in *Brugia*, either in the form of the native protein or the β -galactosidase fusion proteins derived from the family of cDNA clones (data not shown). The mammalian PRPs are calcium-binding proteins involved in regulating calcium concentrations in saliva and keeping the exposed mineralized tissue of the teeth intact under physiological conditions (Bennick, 1976; Bennick and Cannon, 1978). Mf22 does not bind calcium (data not shown), and the cross-linked nature of the protein would suggest that it plays a structural role in determining the morphology of the sheath rather than a function involving calcium binding.

Our interpretation is that the similarities derive from chemical constraints dictated by the effects of multiple proline residues, which are mitigated by the presence of small amino acids such as glycine, resulting in a number of proteins with a high degree of apparent homology. These constraints are evident in collagens, where the presence of glycine in every third position is necessary for the formation of the triple helix. The regularity of prolines and diprolines seen in the central domain of the microfilarial sheath protein suggests some form of multiple turn structure, but little else can be predicted at this stage.

The C-terminal domain of Mf22 is flanked by cysteines which show an identical alignment to the residues found between Gly-Xaa-Yaa repeat blocks of nematode cuticular collagens and implicated in cross-linking (Kramer et al., 1982; Kramer et al., 1990; Shamansky et al., 1989). As Mf22 is only soluble in reducing agents, and all of the cysteine residues outside the putative leader sequence are clustered in these two sites, it seems reasonable to assume that this alignment is structurally significant. The C-terminal domain thus shows similarity to repeat domains of nematode cuticular collagens in that they are both proline-rich and flanked by clusters of cysteine residues. Moreover, this domain is encoded by a distinct exon, which is separated precisely from that which encodes the 5-amino acid repeat region (Fig. 5). These similarities suggest that Mf22 may be incorporated into a higher order fibrous structure like the collagens.

The fertilized ovum of *Brugia* has a clearly visible eggshell, which is closely applied to the oolemma. At the third or fourth division, the eggshell separates from the outer layers of the embryo. Later on, it is reinforced by uterine secretions of an unknown composition, and it is this composite structure that forms the sheath which is retained in mature larvae (Rogers *et al.*, 1976). We are currently attempting to define the precise origin and site of formation of Mf22 in the uterus by immunogold labeling.

The reason why some filarial species retain the sheath and others discard it is unknown. Suggestions that it may serve to protect the microfilariae against immune reactions are unconvincing in that this merely shifts the focus of an effector response onto another surface. Moreover, there is a strong positive correlation between the presence of antibodies which bind the sheath and clearance of microfilariae (Wong and Guest, 1969; McGreevy et al., 1980; Pinder et al., 1988). We are therefore examining the potential of Mf22 as an immunogen which would promote clearance of microfilariae and thus effectively block transmission to the mosquito vector.

REFERENCES

- Ann, D. K., Smith, K., and Carlson, D. M. (1988) J. Biol. Chem. 263, 10887 - 10893
- Anya, A. O. (1976) Adv. Parasitol. 14, 267-351
- Bennick, A. (1976) Biochem. J. 155, 163-169
- Bennick, A., and Cannon, M. (1978) Caries Res. 12, 159-169
- Bornstein, P., and Traub, W. (1979) in The Proteins, Vol. 4, pp. 412-633, Academic Press, New York
- Bressan, G. M., Argos, P., and Stanley, K. K. (1987) Biochemistry 26.1497-1503
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979) Biochemistry 18, 5294-5299
- Chitwood, B. G., and Chitwood, M. B. (1950) An Introduction to Nematology, Monumental Printing Co., Baltimore
- Clements, S., Mehansho, H., and Carlson, D. M. (1985) J. Biol. Chem. 260, 13471-13477
- Cordingley, J. S. (1987) Parasitology Today 3, 341-344
- Croll, N. A., and Matthews, B. E. (1977) Biology of Nematodes, Blackie & Son, Glasgow, Scotland
- Devaney, E. (1985) Trop. Med. Parasitol. 36, 25-28
- Feinberg, A. P., and Vogelstein, B. (1983) Anal. Biochem. 132, 6-13
- Fuhrman, J. A., and Piessens, W. F. (1985) Mol. Biochem. Parasitol.
- 17.93-104
- Furman, A., and Ash, L. R. (1983) Acta Trop. 40, 45-51
- Gao, K., Klein, L. E., Britten, R. J., and Davidson, E. H. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 8634-8638
- Kramer, J. M., French, R. P., Park, E.-C., and Johnson, J. J. (1990) Mol. Cell. Biol. 10, 2081-2089
- Kramer, J. M., Cox, G. N., and Hirsh, D. (1982) Cell 30, 599-606
- Laurence, B. R., and Simpson, M. G. (1974) Int. J. Parasitol. 4, 523-536
- Lipman, D. J., and Pearson, W. R. (1985) Science 227, 1435-1441
- Maeda, N., Kim, H.-S., Azen, E. A., and Smithies, O. (1985) J. Biol. Chem. 260, 11123–11130
- Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Maizels, R. M., Gregory, W. F., Kwan-Lim, G.-E., and Selkirk, M. E. (1989) Mol. Biochem. Parasitol. 32, 213-228

- McGreevy, P. B., Ratiwayanto, S., Tuti, S., McGreevy, M. M., and Dennis, D. T. (1980) Am. J. Trop. Med. Hyg. 29, 553-562
- McLaren, D. J. (1972) Parasitology 65, 317-332
- McPhail, G. D., Finn, T., and Isaacson, P. G. (1987) J. Pathol. 151, 231 - 238
- Neville, A. C. (1975) Biology of the Arthropod Cuticle, Springer-Verlag, Berlin
- O'Farrell, P. Z., Goodman, H., and O'Farrell, P. H. (1977) Cell 12, 1133-1142
- Ovchinnikov, Y. A., Abdulaev, N. G., Zolotarev, A. S., Artamonov, I. D., Bespalov, I. A., Dergachev, A. E., and Tsuda, M. (1988) FEBS Lett. 232, 69-72
- Partono, F., and Purnomo (1987) Trans. R. Soc. Trop. Med. Hyg. 81, 657 - 662
- Perrine, K. G., Denker, J. A., and Nilsen, T. W. (1988) Mol. Biochem. Parasitol. 30, 97–104
- Pinder, M., Dupont, A., and Egwang, T. G. (1988) J. Immunol. 141, 2480-2486
- Rogers, R., Ellis, D. S., and Denham, D. A. (1976) J. Helminthol. 50, 251 - 257
- Rogers, W. P. (1962) The Nature of Parasitism, Academic Press, New York
- Sayers, G., Mackenzie, C. D., and Denham, D. A. (1984) Parasitology 89, 425-434
- Schraermeyer, U., Peters, W., and Zahner, H. (1987) Parasitol. Res. 73, 557-564
- Selkirk, M. E., Denham, D. A., Partono, F., and Maizels, R. M. (1989a) J. Immunol. 143, 299-308
- Selkirk, M. E., Nielsen, L., Kelly, C., Partono, F., Sayers, G., and Maizels, R. M. (1989b) Mol. Biochem. Parasitol. 32, 229-246
- Selkirk, M. E., Gregory, W. F., Yazdanbakhsh, M., Jenkins, R. E., and Maizels, R. M. (1990) Mol. Biochem. Parasitol. 42, 31-43
- Shamansky, L. M., Pratt, D., Boisvenue, R. J., and Cox, G. N. (1989) Mol. Biochem. Parasitol. 37, 73-86
- Simpson, M. G., and Laurence, B. R. (1972) Parasitology 64, 61-88
- Steinert, P. M., Rice, R. H., Roop, D. R., Trus, B. L., and Steven, A. C. (1983) Nature 302, 794-800
- Tarr, G. E. (1973) Comp. Biochem. Physiol. 46, 167-176
- Tarr, G. E., and Fairbairn, D. (1973) Lipids 8, 7-16
- Taylor, D. W., Goddard, J. M., and McMahon, J. E. (1984) Trans. R. Soc. Trop. Med. Hyg. 78, 707-708 Towbin, H., Staehelin, T., and Gordon, J. (1979) Proc. Natl. Acad.
- Sci. U. S. A. 76, 4350-4357
- von Mende, N., Bird, D. M., Albert, P. S., and Riddle, D. L. (1988) Cell 55, 567-576
- Wharton, D. A. (1979) Int. J. Parasitol. 9, 127-131
- Wharton, D. A. (1980) Parasitology 81, 447-463
- Wharton, D. A., and Jenkins, T. (1978) Tissue & Cell 10, 427-440
- Wong, M. M., and Guest, M. F. (1969) Trans. R. Soc. Trop. Med. Hyg. 63, 796-800
- Wong, R. S. C., and Bennick, A. (1980) J. Biol. Chem. 255, 5943-5948
- Wong, R. S. C., Hofmann, T., and Bennick, A. (1979) J. Biol. Chem. 254, 4800-4808
- Young, R. A., and Davis, R. W. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 1194-1197
- Zaman, V. (1987) Parasitol. Res. 73, 281-283