Cloning and biochemical characterization of the cyclophilin homologues from the free-living nematode *Caenorhabditis elegans*

Antony P. PAGE*‡, Kenneth MACNIVEN* and Michael O. HENGARTNER†

*Wellcome Unit of Molecular Parasitology, The Anderson College, Glasgow University, Glasgow G11 6NU, Scotland, U.K., and †Cold Spring Harbor Laboratory, P.O. Box 100, Cold Spring Harbor, New York 11724, U.S.A.

Cyclosporin A (CsA) is the most widely used immunosuppressive agent, whose properties are exerted via an interaction with cyclophilin, resulting in down-regulation of signal-transduction events in the T-cell. Cyclophilin is identical with peptidylprolyl *cis–trans* isomerase (PPI; EC 5.2.1.8), an enzyme which catalyses the isomerization between the two proline conformations in proteins, thereby acting as a catalyst in protein-folding events. Several reports indicate that CsA has potent anti-parasitic activity, effective against both protozoan and helminth species. In order to understand the various biological roles that cyclophilins play we have initiated a study of these proteins in the genetically tractable nematode *Caenorhabditis elegans*. Here we describe the cloning and characterization of 11 cyclophilin genes (*cyp-1* to *-11*) derived from this nematode; this is currently the greatest number of isoforms described in a single species. Southern blotting and physical mapping indicated that these genes are dispersed throughout the nematode genome. A high degree of conservation exists between several isoforms, which also share characteristics with the ubiquitous isoforms previously described. The remaining isoforms are divergent, having altered CsA-binding domains and additional non-cyclophilin domains, which may impart compartmental specificity. Ten of these isoforms have been expressed in *Escherichia coli*, and the resultant fusion proteins have been examined biochemically for PPI activity, which they all possess. Isomerase activity is highest in the conserved and lowest in divergent isoforms, perhaps indicating a more specific substrate for the latter. Analysis of the *C. elegans cyp* genes will provide answers as to the roles played by cyclophilins in protein folding and signal transduction.

INTRODUCTION

The cyclophilins (CYPs) are a large family of ubiquitous proteins of which six different isoforms have been described in humans [1,2]. As their name implies, these proteins bind to the immunosuppressive agent cyclosporin A (CsA) via a central highly conserved CsA-binding domain (CBD) [3]. The N- and C-terminal domains are relatively divergent, and many isoforms posses secretory or mitochondrial targeting sequences. Biochemical studies have shown that CYPs possess peptidylprolyl *cis–trans*-isomerase (PPI: EC 5.2.1.8) activity, and act both as catalysts and chaperones in protein-folding events, speeding up slow rate-limiting steps in the folding of various proteins, including the proline-rich collagens [4,5].

CsA behaves as a CYP substrate analogue and binds to the active site of the protein. CsA-binding results in the competitive inhibition of CYP isomerase activity, and thereby can slow down the folding of certain proteins [5]. Interestingly, a second structurally unrelated PPI, FKBP (FK-506-binding protein), has also been identified, whose association with the bacterial macrolides FK506 and rapamycin also resulted in immunosuppression [6]. The potent human T-helper-cell-specific immunosuppression of these structurally unrelated compounds is not due to inhibition of their respective immunophilin PPI activities, but results from the complex formed between them binding to and inhibiting the protein phosphatase calcineurin [2]. This CsA/CYP–calcineurin interaction is evolutionarily conserved, also being described in yeast [7].

Cyclophilins have been identified in a large number of species ranging from bacteria to vertebrates, including several parasitic species [8]. Sub-immunosuppressive levels of CsA are toxic to a wide range of parasites: this effect has been been noted in both helminth and protozoan species, and to date, remains uncharacterized [9]. Non-immunosuppressive CsA analogues have also been described which have potent effects against malaria parasites [10] and HIV-1 virions [11].

Divergent forms of cyclophilins, which have altered CBDs and extended C- and N-terminal domains have recently been described in vertebrates [12,13], in *Drosophila melonagaster* [14] and in the parasitic nematode *Brugia malayi* [15]. Such divergent forms may possess a more specific function, and the altered CBD has been hypothesized to correlate with substrate specificity [12]. Consistent with this hypothesis, ninaA, a divergent cyclophilin isoform identified in *Drosophila*, was shown to be responsible for the specific folding of the rhodopsins in the compound eye [14], acting as a specific chaperone [16]. Recent studies also indicate that CYPs play a role in the stress response. The yeast *cyp-1* gene contains a heat-inducible promoter, and yeast *cyp-1* and *cyp-2* mutants have a reduced survival rate following heat shock [17].

We have undertaken an analysis of the cyclophilin homologues from the genetically tractable nematode *Caenorhabditis elegans*, with the goal of determining their biological roles in nematodes. As a first step, we have set out to identify and clone the cyclophilin isoforms in this organism, and this has been achieved using a combination of techniques: (1) PCR on genomic DNA using degenerate primers designed to hybridize to highly

Abbreviations used: PPI, peptidylprolyl *cis–trans*-isomerase; CsA, cyclosporin A; CYP, cyclophilin; CypA, cyclophilin A isoform; MBP, maltose-binding protein; CBD, cyclosporin A-binding domain; RT, reverse transcriptase; EST, expression sequence tags.

Nucleotide sequence data reported in this paper have been deposited in the GenBank databases under the accession numbers: U30943, U34354, U31077, U36187, U31948, U27354, U27559, U31078, U36581, U34954 and U34955.

‡ To whom correspondence should be addressed.

conserved regions of the protein; (2) searching the *C. elegans* expression sequence tags (EST) database; and (3) searching of the genomic DNA sequences generated by the *C. elegans* genome project [18].

So far, we have identified, cloned and sequenced 11 cyclophilin genes from this nematode, representing the largest number of homologues from a single species, and undoubtedly reflecting the progress of the genome sequencing project. We predict that a similar high number of isoforms will be present in other eukaryotic species, including humans. These genes have been termed *cyp-1* to *-11* (**cy**clo**p**hilin-related). Where necessary, the physical location of these genes has been determined, and their genomic organization and degree of cross-reactivity has been examined by Southern blotting. We have overexpressed ten of these isoforms in *Escherichia coli* as maltose-binding fusion proteins, and analysed them for PPI activity. All ten isoforms are active PPIs and their activity corresponds to their degree of identity with the ubiquitous isoforms, with the highly conserved proteins being the most active enzymes.

MATERIALS AND METHODS

The standard molecular biology techniques used in this study are described in detail elsewhere [15,19].

Cloning of the *C. elegans* cyclophilin genes: degenerate PCR

Degenerate primers were designed from regions of cyclophilins conserved among vertebrate and yeast isoforms (DCypF and DCypR; Table 1). These primers were used to amplify *C. elegans* homologues from genomic DNA. Two distinct bands were noted on agarose gels, and the degenerate PCR products corresponding to these bands were subcloned into pBluescriptII (Stratagene, La Jolla, CA, U.S.A.). Sequencing of the inserts showed that both corresponded to bona fide *C. elegans* homologues, which we named *cyp-1* and *cyp-2*. The cloned PCR fragments were used to screen a *C. elegans* lambdaZAP cDNA library (a gift from R. Barstead, Oklahoma Medical Research Foundation, Oklahoma, OK, U.S.A.). From approximately 500000 plaques, two independent *cyp-1* and three independent *cyp-2* cDNAs were isolated. These clones were sequenced to completion in both directions. The *cyp-2* clone was incomplete at the 5′ end and a corresponding genomic clone was isolated and sequenced to determine the 5′ sequence. This genomic sequence was then further confirmed by reverse transcriptase (RT)-PCR with cyp2F and cyp2R primers (Table 1) on *C. elegans* cDNA.

Table 1 Primers used in the subcloning of the *C. elegans* cyclophilins

Underlined sequences refer to restriction sites. i = inosine, y = c/t, r = a/g, m = a/c.

Primer name	Primer sequence	Restriction site
DCypF	cciaaracigcigaraayttymgigciyt	–
DCypR	ccraaiaciacitgytticcrtciarcca	–
RISL1	gcggaattcggtttaattacccaagtttgag	*Eco*RI
cyp1F	gcggaattcatgaaatttctactccgtgcctcctc	*Eco*RI
cyp1R	gcgtctagattactcgctcttcaactctccgc	*Xba*I
cyp2F	cgaattcatgccacgtgtcaaagtgttcatc	*Eco*RI
cyp2R	cgtctagatcacttcatttctccgcagtc	*Xba*I
cyp3F	gcggaattcatgagccgctcaaaggtcttt	*Eco*RI
cyp3R	ggcaagctttatgccttgagttgtcgacagtc	*Hin*dIII
cyp4F	gcgtctagaatggctccagttacaagcaat	*Xba*I
cyp4RT	cgaagctttacacacaaatacttcagctctcata	*Hin*dIII
cyp5F	gcgaattcatgaagtcgttcttgttgtggcg	*Eco*RI
cyp5R	gcgtctagattagacgacggcctcgcgttc	*Xba*I
cyp6F	gcggaattcatgagccgtgctcttctttttc	*Eco*RI
cyp6R	gcggatcctcatggaatatgtccagcgttggc	*Bam*HI
cyp7F	ggaaggatttcaatgagccgcccaagagtcttcttc	*Xmn*I
cyp7R	gctctagattagagctgtccgcagtcggc	*Xba*I
cyp8F	cggaattcatgcctccagaggtccgtggtaac	*Eco*RI
cyp8R	gcgtctagattacggctcagaagcagctgc	*Xba*I
cyp9F	gcgaattcatggccgctgaaaaaagagtgttcc	*Eco*RI
cyp9RT	gctctagatcatttcatgcttcttctcattg	*Xba*I
cyp10F	cggaattcatgtcagtcacacttcacac	*Eco*RI
cyp10R	gctctagatcacttttgctgaacgagag	*Xba*I
cyp11F	gcgaattcatgactgaatacgacaagtttg	*Eco*RI
cyp11R	ccggatccttagagttggccacattgaac	*Bam*HI

RT-PCR cloning from EST clones

The genes *cyp-3*, *cyp-5* and *cyp-7* were all originally identified by the genome sequencing project as partial ESTs: CEESH95F, CEESH33F and CEESN11F respectively, which all had significant BLAST hits against cyclophilins [20,21]. *Cyp-3* was also independently isolated by M. Krause, National Institutes of Health, Bethesda, MD, U.S.A. (personal communication). Full-length cDNAs were isolated by SL-1 RT-PCR using a primer sequence corresponding to the 5′-most stop codon of the corresponding EST (Table 1: cyp3R, cyp5R and cyp7R respectively) and the nematode transpliced leader sequence primer, SL-1 (Table 1; RISL1), on mixed-stage cDNA. Poly(A)$^+$ mRNA was obtained using the Micro-Fast Track Kit (InVitrogen), and cDNA was synthesized using a cDNA synthesis kit (Amersham, Little Chalfont, U.K.). RT-PCR conditions were as follows: 10 pmol of each primer, 1 × buffer (45 mM Tris, pH 8.8/11 mM ammonium sulphate/4.5 mM $MgCl_2$/6.7 mM 2-mercaptoethanol/4.4 μM EDTA, pH 8), 1 mM of each dNTP, 113 μg/ml non-acetylated BSA, 2 μl of mixed-stage cDNA and 1 μl of Taq polymerase (Advanced Biotechnologies), cycled 20 times at 94 °C for 30 s, 60 °C for 1 min and 72 °C for 1 min in 100 μl final volume.

RT-PCR cloning from cosmid clones

Cyp-4 was originally identified by J. Waddle and R. Waterston (Washington University, St. Louis, MO, U.S.A.) during the characterization of a neighbouring gene, *cap-2* ([22]; J. Waddle, personal communication). The remaining genes, *cyp-6*, *cyp-8*, *cyp-9*, *cyp-10* and *cyp-11*, were all originally identified by the genome sequencing project as sequenced cosmid clones which possessed significant BLAST scores against cyclophilins. These were contained in the cosmids C51B6, D1009, C14B1, B0252 and T01B7 respectively. Primers were designed which corresponded to the predicted open reading frames of these genes (Table 1), and cDNA clones were generated as above from mixed-stage cDNA by RT-PCR.

Physical mapping

The genes *cyp-1*, *cyp-2*, *cyp-3*, *cyp-5* and *cyp-7* were mapped to the *C. elegans* genomic YAC grid [23] by hybridization of gel-purified inserts following random priming with [^{32}P]dATP. This method localizes the clone to a 100–200 kb region. The position of the clones can be further refined by testing for hybridization to cosmid clones from the relevant region (Table 2).

Genomic Southern hybridization

Southern blots of *Hin*dIII-restricted *C. elegans* genomic DNA were probed with [^{32}P]dATP-labelled inserts corresponding to the genes *cyp-1* to *cyp-11* in hybridization buffer [23] at 65 °C overnight. Blots were subsequently washed extensively at

Table 2 The family of *C. elegans* cyclophilins

The *C. elegans* isoforms are listed CYP-1 to -11. Letters in parentheses refer to: A, cyclophilin A or cytosolic CYP-like isoform; B, cyclophilin B or secreted CYP-like isoform; C, cyclophilin C or mitochondrial CYP-like isoform; D, cyclophilin D or divergent CYP-like isoform. The accession numbers of the various isoforms are listed, and the sizes of predicted proteins include the signal peptide where applicable (aa, amino acid). Names of genes and accession numbers in square brackets refer to available genomic sequences. Mapping locations (Map) define the physical chromosomal location in the worm's genome, with precise cosmid location listed where applicable; chromosomes are listed I–X [23]. The CsA domain refers to the residues in the human CypA isoform which bind CsA, with changes in this conserved domain depicted in single-letter amino acid code, and numbering referring to the position in the human cyclophilin. The presence or absence of a signal peptide is indicated, as is the presence of a central 7–8 amino acid insert. The presence and type of C-terminal domain is also indicated, and pI refer to isoelectric points for the isoforms minus any signal peptides.

Isoform	Accession number	Size (aa)	Map (cosmid)	CsA-domain	Signal peptide	Insert	C-extension	pI
CYP-1 (C)	U30943	192	VC	Conserved	Yes (19 aa)	Yes	No	6.7
CYP-2 (A)	U34354	171	IIIR (ZK526)	A103C	No	Yes	No	8.1
CYP-3 (A)	U31077	173	VR	Conserved	No	Yes	No	8.8
CYP-4 (D) [F59E10.2]	U36187 [Z46935]	524	IIRC (M106)	T73H, A103K, W121Y	No	No	Yes (hydrophobic)	7.9
CYP-5 (C)	U31948	204	IR (B0467)	Conserved	Yes (28 aa)	No	Yes (short)	9.5
CYP-6 (B) [F42G9.2]	U27354 [U00051]	201	IIIL (C51B6)	Conserved	Yes (25 aa)	No	Yes (short)	5.8
CYP-7 (A)	U27559	171	VR	Conserved	No	Yes	No	8.5
CYP-8 (D) [D1009.2]	U31078 [U40938]	466	XRC (D1009)	A103R, W121H	No	Yes	Yes (hydrophilic)	10.7
CYP-9 (D) [T27D1.1]	U36581 [Z48245]	309	IIIL (C14B1)	T73R, A103K, W121H, L122C	No	Yes	Yes (hydrophilic)	8.6
CYP-10 (D) [B0252.4]	U34954 [U23453]	147	IIC (B0252)	T73K, A103N, W121H, H126Y	No	No	No	5.3
CYP-11 (D) [T01B7.4]	U34955 [Z66499]	183	IIC (T01B7)	W121F	No	Yes	No	6.5

65 °C in 0.5% SDS/0.5 × SSC (SSC = 0.15 M NaCl/0.015 M sodium citrate) and exposed for autoradiography.

Sequencing and overexpression of the *C. elegans* cyclophilin isoforms

All cDNA clones were digested with the appropriate restriction enzymes (Table 1) and ligated into similarly cut pMalc2 vector (New England Biolabs). Plasmid DNA was isolated (Plasmid Kit, Qiagen) and the inserts were sequenced in both directions by thermal-cycle sequencing on an Applied Biosystems Automated Sequencer, and where appropriate were compared with the available corresponding EST and cosmid DNA sequences. Sequence alignments and comparisons were done using the University of Wisconsin Genetics Computer Group software. Database searches were performed using the BLAST program [24]. Protein sequences were searched for conserved functional motifs using PROSITE [25].

Maltose-binding fusion proteins were obtained for all isoforms, except CYP-7, following the manufacturer's instructions (New England Biolabs; Protein Fusion and Purification System Manual). CYP-7 was excluded due to its extremely high degree of identity with the cyclophilin A isoform CYP-3 (87% identical, 91% similar), which rises to 100% identity over the PPI- and CsA-binding domains. Initial insolubility problems with the large two-domain isoforms CYP-8 and -9, and the three-domain protein CYP-4, were overcome by expressing the cyclophilin domains only (Figures 3A and 3B).

PPI assay

The recombinant maltose-binding protein (MBP)–cyclophilin fusion proteins were analysed for proline isomerase activity to a synthetic Ala-Pro-containing peptide (*N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanalide; Bachem AG), following published methods [15] based on the assay originally described by Fischer et al. [26]. Reactions were performed at 6 °C/400 nm using a Beckman DU640 spectrophotometer. First-order rate kinetics were observed with a rate constant $k_{obs} = (K_{cat}/K_m)[E]$.

RESULTS AND DISCUSSION

We report in this paper the identification and characterization of 11 separate cyclophilin genes from *C. elegans*, ten of which have been expressed in *E. coli* and have been determined to posses PPI activity. All of these genes are expressed, as we have been able to isolate full-length cDNAs for each of them. The genes *cyp-1* and

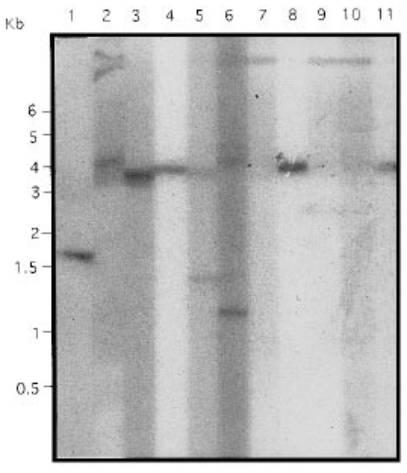


Figure 1 Southern blot analysis of the *C. elegans* cyclophilin genes

Genomic DNA was restricted with *Hin*dIII and separated on a 1% curtain agarose gel. The blot was cut into individual strips which were subsequently probed with the individual *cyp* cDNA clones. Lane numbering refers to *cyp* genes (*cyp-1* to *-11*). Size markers are in kilobases (Kb).

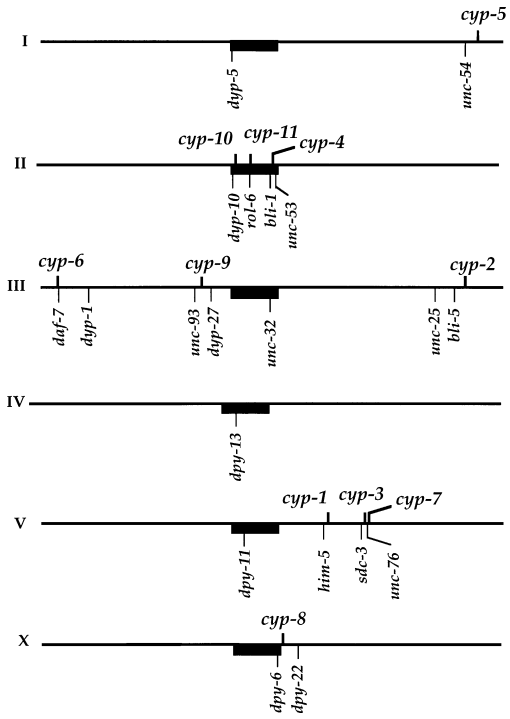


Figure 2 Map position of *C. elegans* cyclophilins

The *C. elegans* cyclophilin genes *cyp-1* to *-11* were placed onto the genetic map based on their physical position. Clones were mapped to the physical map by hybridization to the YAC polytene filter (See Materials and methods section). Genetic markers are denoted under each chromosome; for a full description refer to [38].

cyp-2 were isolated by degenerate PCR and, as expected, closely resemble the ubiquitous cyclophilins in the databases. Partial clones of *cyp-3*, *-5* and *-7* were originally isolated as EST clones [20,21] and, not surprisingly, these are amongst the most abundantly expressed of the 11 genes (A. P. Page, unpublished work). The remaining cyclophilins, *cyp-4*, *-6*, *-8*, *-9*, *-10* and *-11,* were all originally identified during the *C. elegans* genome sequencing project, which is still underway. In total, this represents the greatest number of isoforms identified in a single species, and may not be a complete list. Southern blotting and physical mapping indicate that these genes are separate (Figure 1) and are distributed throughout the nematode genome (Figure 2).

The proteins encoded by the *cyp* genes are shown in Figure 3(A), aligned to the well-characterized human cyclophilin A isoform (CypA). The *C. elegans* CYP proteins are highly varied in sequence, many isoforms being more closely related to human CypA than to some of the more divergent isoforms (Table 3), suggesting that the common ancestor of *C. elegans* and mammals already possessed a number of cyclophilin isoforms, from which various subfamilies originated [27]. This analysis indicated that CYP-4, -8, -9, -10 and -11 are the most divergent isoforms, both when compared with the other *C. elegans* CYPs and with human CypA. All isoforms, except CYP-9, -10 and -11, have the 11 amino acid PROSITE [25] consensus sequence for PPI, with either 1 or 2 amino acid differences being present (Figure 3A). The PROSITE consensus sequence does however fall outside the conserved PPI active domain, and the PPI domains of these isoforms are relatively conserved. The side-chains of human CypA, which have been shown by NMR and X-ray crystallographic analysis to be associated with CsA binding, are denoted by asterisks in Figure 3A: R55, I57, F60, M61, Q63, G72, T73, A101, N102, A103, Q111, F113, W121, L122 and H126 [3,28]. The NMR data suggest the presence of three hydrogen-bonds between *N*-methyl-L-leucine-10 (MeLeu) (CsA) and R55, MeLeu-9 (CsA) and W121 and L-α-aminobutyric acid-2 (Abu) (CsA) and N102 [28]. Site-directed mutagenesis of the CsA-insensitive *E. coli* cyclophilin demonstrated the importance of W121 in CsA binding, since changing the corresponding residue from F to W in this isoform resulted in an increased sensitivity to CsA [29]. As can be seen from Figure 3(A) and Table 2, many of the *C. elegans* CYPs have altered CBDs.

The properties of the CYP isoforms are summarized in Table 2: the majority of the isoforms, like other cyclophilins from higher eukaryotes, are basic proteins, and only CYP-1, -6, -10 and -11 have acidic pIs. CYP-2, -3 and -7 most closely resemble the well-characterized CypA in structure and sequence, and thus are likely to be members of the cytosolic CsA-binding cyclophilin subfamily. Like human CypA, these three genes are highly expressed (A. P. Page, unpublished work). Three of the cyclophilins contain potential N-terminal signal peptides. The 25 amino acid hydrophobic signal peptide found in CYP-6 is characteristic of secreted or cyclophilin B isoforms, and this protein also has a highly conserved CBD. CYP-1 and -5 also possess signal peptides, however, these more closely resemble mitochondrial targeting sequences, since they are both amphipathic and positively charged. This suggests that these isoforms localize to cellular organelles, and are members of the cyclophilin C subfamily. Both isoforms have the conserved CsA-binding motifs. CYP-5 probably represents the most abundant CYP isoform, since it was independently isolated five times during the course of the *C. elegans* cDNA sequencing project [20,21].

Based both on overall amino acid divergence with the other CYPs (Table 3), and a comparison with the central conserved region of human CypA (Figure 3A), the remaining cyclophilin isoforms, CYP-4, -8, -9, -10 and -11, all represent divergent or 'cyclophilin D isoforms'. These isoforms posses one or more residue changes in the well-characterized CBD [3] (Figure 3A, residues denoted by an asterisk). Similarly altered CBDs have been hypothesized to correlate with substrate specificity [12]. Identification of CsA derivatives, with specific binding to the divergent forms and not the canonical forms, would result in further functional dissection of the cyclophilins. The isoforms CYP-8, CYP-9 and CYP-4 are especially divergent proteins: all three have altered CBDs and extended C-terminal domains (Figure 3B). The CYP-8 and CYP-9 isoforms have a W121 to H change in the major CsA-binding residue and extremely hydrophilic tail domains. Both attributes are also found in the *B. malayi* CsA-insensitive isoform, presently the only other reported nematode cyclophilin [15]. The CYP-4 isoform also has two changes in the CBD, including a change in the major CsA-binding residue W121 to Y. This isoform also possesses an extended, extremely hydrophobic C-terminal domain, which is most closely related to the tail of the well-characterized *Drosophilia ninaA* gene product [14]. This similarity suggests that CYP-4 may also be an integral membrane protein and, like ninaA, have a more stringent substrate specificity. CYP-4 is actually a three-domain protein (Figure 3B), as it also possesses a large (254 amino acid) N-terminal domain which has weak homology to heat-shock protein-70. Similar multi-domain PPI-containing proteins have been described previously, in both the cyclophilins [30] and the FKBP families [31].

The cyclophilins CYP-1, -2, -3, -7, -8, -9 and -11 all possess a central 7–8 amino acid insert not commonly found in other cyclophilins (Figure 3A). Similar inserts in this region have however been found in cyclophilins from plants [32], *B. malayi*

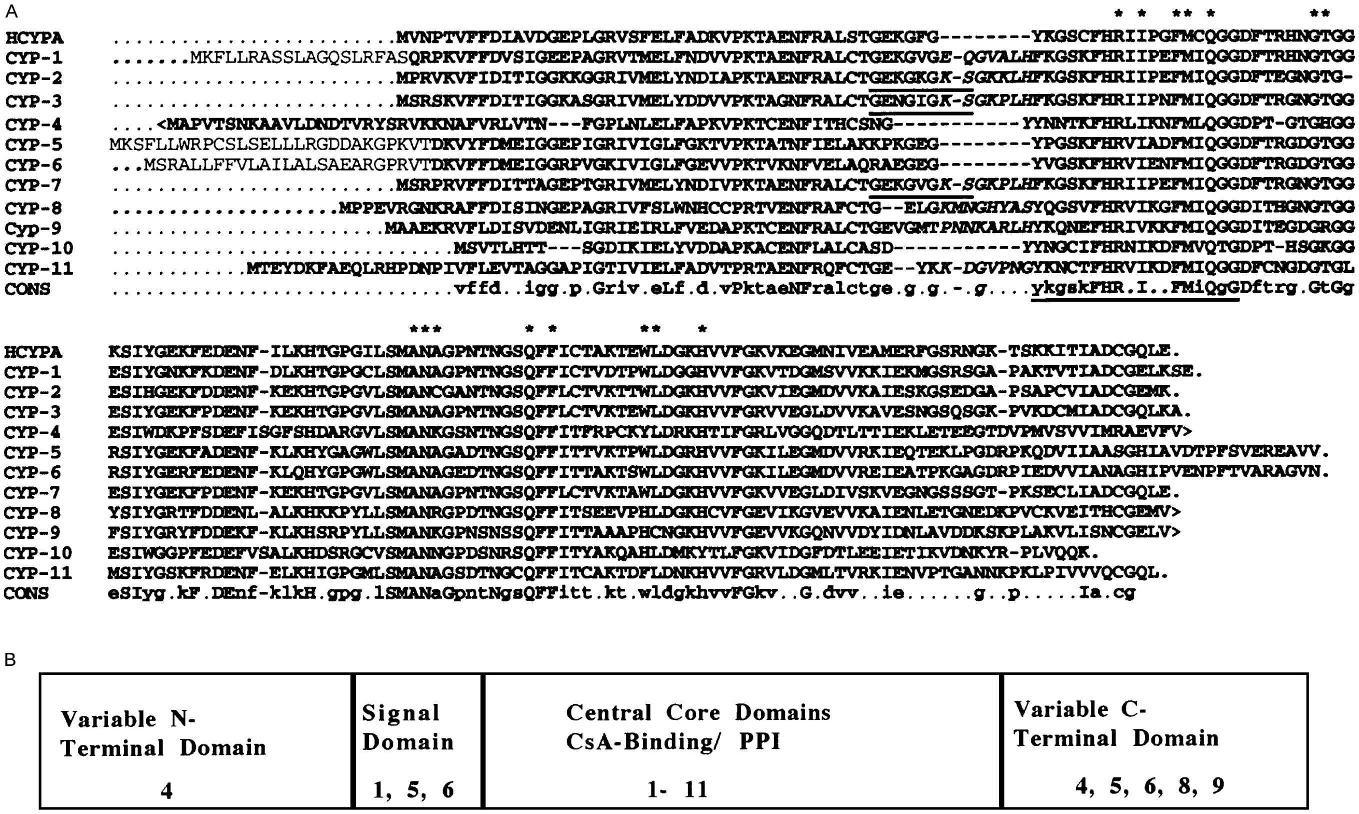


Figure 3 Structure and sequence of the *C. elegans* cyclophilin isoforms

(**A**) Alignment of CYP-1 to -11 with human CypA. Broken lines (–) are introduced for maximal alignment; *, residues which constitute the CBD in human CypA [3]. The consensus line (Cons) is for the *C. elegans* isoforms only. Lower case letters indicate > 5/11 resides identical; upper case letters indicate 11/11 residues identical. The central 7–8 amino acid insert is indicated in italics. The PROSITE [25] PPI consensus sequence ([F/Y]-X2-[S/T/L/V]-X-F-H-[R/H]-[L/I/V/M]2-X2-F-[L/I/V/M]-X-Q-[A/G]-G) is underlined in the consensus. The PROSITE conserved ATP/GTP-binding site motif A (A/G-x4-G-K-S/T) [33] is underlined in CYP-2, -3 and -7. Complete sequences are shown for all isoforms, except CYP-4, -8 and -9 (< and > indicating that more sequence is missing at the N- or C-terminal end respectively), and include signal peptides where applicable (CYP-1, -5 and -6). (**B**) Schematic representation of the various domains of the *C. elegans* cyclophilins. Numbers represent cyclophilin isoforms possessing these features.

Table 3 Percentage identity between the *C. elegans* CYP isoforms

All predicted amino acid sequences were compared using the GCG Gap program (University of Wisconsin Genetics Computer Group), and adjusted manually to provide the best alignment. Percentage similarities are indicated in parentheses. HCYP-A, human cyclophilin A.

Isoform	C-like CYP-1	A-like CYP-2	A-like CYP-3	D-like CYP-4	C-like CYP-5	B-like CYP-6	A-like CYP-7	D-like CYP-8	D-like CYP-9	D-like CYP-10	D-like CYP-11
CYP-1 [C]	–										
CYP-2 [A]	71 (80)	–									
CYP-3 [A]	72 (80)	81 (87)	–								
CYP-4 [D]	25 (51)	39 (61)	38 (60)	–							
CYP-5 [C]	55 (69)	59 (73)	60 (74)	31 (60)	–						
CYP-6 [B]	52 (67)	57 (73)	59 (78)	32 (57)	70 (78)	–					
CYP-7 [A]	73 (81)	79 (85)	87 (91)	40 (59)	60 (72)	56 (71)	–				
CYP-8 [D]	47 (70)	53 (71)	48 (69)	35 (61)	45 (70)	48 (68)	51 (70)	–			
CYP-9 [D]	46 (67)	50 (67)	45 (65)	31 (56)	41 (65)	42 (64)	50 (66)	51 (69)	–		
CYP-10 [D]	41 (66)	42 (62)	40 (63)	48 (69)	36 (61)	37 (63)	42 (63)	42 (64)	44 (64)	–	
CYP-11 [D]	48 (66)	49 (67)	51 (66)	35 (60)	49 (66)	51 (69)	55 (70)	48 (68)	45 (64)	41 (64)	–
HCYP-A	65 (79)	58 (66)	65 (77)	34 (59)	56 (68)	55 (70)	71 (83)	53 (71)	49 (66)	38 (67)	54 (70)

[15] and divergent vertebrate cyclophilin-like proteins [12,13]. The function of these inserts remains to be established. The insert is relatively conserved among CYP-1, -2, -3 and -7 but is divergent in the others. It is of interest to note that CYP-2, -3 and -7 all posses a PROSITE [25] conserved ATP/GTP-binding site motif A (A/G-x4-G-K-S/T)[33] at the start of this insert (Figure 3A, underlined).

The cyclophilin domains of *cyp-1*, *-2*, *-3*, *-4*, *-5*, *-6*, *-8*, *-9*, *-10*

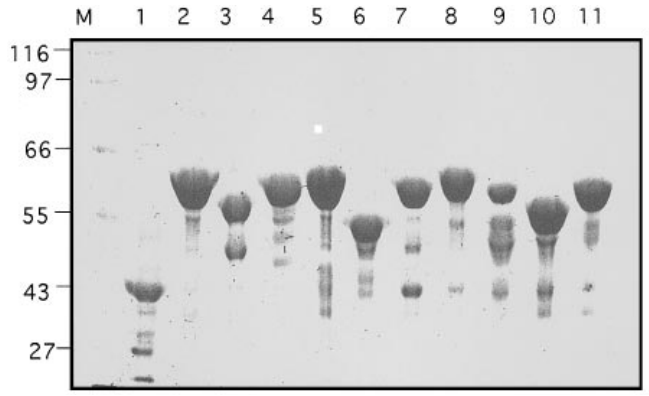
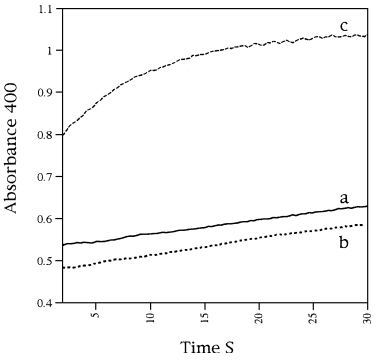


Figure 4 Expression of recombinant *C. elegans* cyclophilin

A 10 μg amount of each isoform was separated on a 10%- SDS/PAGE gel, and the proteins were visualized by Coomassie Brilliant Blue staining. Protein molecular mass markers (M, in kDa) are indicated on the left (New England Biolabs). Lanes: 1, MBP; 2, CYP-1; 3, CYP-2; 4, CYP-3; 5, CYP-4; 6, CYP-5; 7, CYP-6; 8, CYP-8; 9, CYP-9; 10, CYP-10; 11, CYP-11.



Isoform	Kc/Km [μM/min]
CYP-6 (B)	502.15
CYP-2 (A)	36.95
CYP-3 (A)	21.38
CYP-5 (C)	4.44
CYP-1 (C)	4.21
CYP-8 (D)	1.17
CYP-10 (D)	1.16
CYP-4 (D)	1.12
CYP-9 (D)	0.95
CYP-11 (D)	0.95

Figure 5 Enzyme activity of *C. elegans* cyclophilin isoforms

Typical PPI progression curves: curve a, thermal isomerization of Ala-Pro synthetic peptide in the absence of enzyme; curve b, thermal isomerization of Ala-Pro synthetic peptide in the presence of MBP; curve c, increase in isomerization of Ala-Pro synthetic peptide in the presence of 10 nM CYP-6. Enzyme activity is expressed as K_{cat}/K_m (μM/min). Isoforms are listed in order of activity.

and *-11* were expressed as soluble fusion proteins in *E. coli* (Figure 4), and in a typical experiment yielded 18–26 mg/l. The recombinant cyclophilins were analysed using the documented PPI assay [15,26]. A typical progress curve is depicted in Figure 5, where the increased catalytic isomerization of a synthetic Ala-Pro peptide is demonstrated by its subsequent increased susceptibility to chymotrypsin-release of the attached *p*-nitroaniline chromophore. This Figure also demonstrates the background or thermal isomerization of substrate in the absence of enzyme, and lack of PPI activity of the unfused MBP. Analysis of the PPI activity groups the enzymes into their various classes, with the most highly conserved enzymes being most active and the divergent isoforms being least active (Figure 5). This result supports the contention that the divergent isoforms may indeed have more specific substrates. The CYP-6 or -B isoform is the most active PPI, being 14–22-fold more active than the CYP-A homologues CYP-2 and -3 respectively. The CYP-C isoforms CYP-1 and -5 are grouped next, being 5–8 times less active than the CYP-A isoforms, and finally the least active are the divergent isoforms CYP-8, -10, -4, -9 and -11, which in turn are 4–5-fold less active than the CYP-C isoforms.

The large number of cyclophilin isoforms in *C. elegans* may reflect a general redundancy or regulatory flexibility for CYPs with the same or similar function. This may be particularly true for the highly conserved isoforms (CYP-1, -2, -3, -5, -6 and -7), which might provide interchangeable functions in different cell types, cellular compartments or organelles. It is interesting to speculate that the more divergent isoforms may have specific or restricted functions, as has been described for the *ninaA* gene in *Drosophila*, which encodes a cyclophilin involved in the specific folding of rhodopsins in the compound eye [14]. Excluding this case, evidence that cyclophilins play a physiological role in protein folding is still more suggestive than conclusive. Nonetheless, several reports support such a link. For example, *in vitro* and *in vivo* experiments have demonstrated that CsA specifically inhibits collagen folding in chicken fibroblasts [5]. Cyclophilin B has recently been implicated in association with heat-shock protein-47, as aiding the export of procollagen from the endoplasmic reticulum [34]. A human CypA isoform has also been shown to associate with HIV-1 virions [11], this interaction being specific to a single proline in the Gag polyprotein of this virus [35]. More recently, mitochondrial cyclophilin isoforms have been implicated in protein folding in this organelle [36,37].

The function of such a large number of isoforms in a single species may not be restricted to promoting *de novo* protein folding, and we predict that numerous isoforms will be present in other eukaryotes, including man. The natural ligands for cyclophilins in general are practically unknown, although evidence accumulated during the study of the immunosuppressive action of CsA suggests that many cyclophilins (perhaps the highly conserved isoforms) are tightly associated with specific target molecules, rather than acting as general protein-folding catalysts.

Why are so many cyclophilins necessary? Do they have independent functions? What are their roles in protein folding and signal transduction? What is their involvement in the antiparasitic effects of CsA? Further analysis of this important group of enzymes in *C. elegans* may elucidate the many roles played by cyclophilins.

We thank M. Krause (National Institutes of Health, Bethesda, MD, U.S.A.) and J. Waddle (Washington University, St. Louis, MO, U.S.A.) for sharing unpublished information and providing clones. We are also grateful to A. Coulson (MRC, Cambridge, U.K.) for supplying the cosmid clones and YAC grids used in this study. Some early aspects of this work were performed in the laboratory of H. R. Horvitz (MIT, Boston, MA, U.S.A.) in collaboration with A. Abeliovich and S. Tonegawa, and all three are thanked for their support and encouragement. We also thank I. Johnstone and D. Barry (WUMP, Glasgow, Scotland, U.K.) for encouragment and critical reading of this manuscript. A. P. P. is funded by the Medical Research Council of Great Britain through a Career Development Fellowship, and M. O. H. is a Rita Allen Foundation Scholar.

REFERENCES

1 Heitman, J., Movva, N. R. and Hall, M. N. (1992) New Biologist **4**, 448–460
2 Galat, A. and Metcalfe, S. M. (1995) Prog. Biophys. Mol. Biol. **63**, 67–118
3 Pfugl, G., Kallen, J., Schirmer, T., Jansonius, J. N., Zurini, G. M. and Walkinshaw, M. D. (1993) Nature (London) **361**, 91–94
4 Bachinger, H. P. (1987) J. Biol. Chem. **262**, 17144–17148
5 Steinmann, B., Bruckner, P. and Superti-Furga, A. (1991) J. Biol. Chem. **266**, 1299–1303
6 Siekierka, J. J., Hung, S. H. Y., Poe, M., Lin, S. C. and Sigal, N. H. (1989) Nature (London) **341**, 755–757
7 Foor, F., Parent, S. A., Morin, N., Dahl, A. M., Ramadan, N., Chrebet, G., Bostian, K. A. and Nielsen, J. B. (1992) Nature (London) **360**, 682–684
8 Page, A. P., Kumar, S. and Carlow, C. K. S. (1995) Parasitol. Today **11**, 385–388
9 Chappell, L. H. and Wastling, J. L. (1992) Parasitology **105**, S25–S40

10 Bell, A., Wernli, B. and Franklin, R. M. (1994) Biochem. Pharmacol. **48**, 495–503
11 Thali, M., Bukovsky, A., Kondo, E., Rosenwirth, B., Walsh, C. T., Sodroski, J. and Gottlinger, H. G. (1994) Nature (London) **372**, 363–365
12 Kieffer, J. L., Thalhammer, T. and Handschumacher, R. E. (1992) J. Biol. Chem. **267**, 5503–5507
13 Anderson, S. K., Gallinger, S., Roder, J., Young, H. A. and Ortaldo, J. R. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 542–546
14 Schneuwly, S., Shortridge, R. D., Larrivee, D. C., Ono, T., Ozaki, M. and Pak, W. L. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 5390–5394
15 Page, A. P., Landry, D., Wilson, G. G. and Carlow, C. K. S. (1995) Biochemistry **34**, 11545–11550
16 Baker, E. K., Colley, N. J. and Zucker, C. S. (1994) EMBO J. **13**, 4886–4895
17 Sykes, K., Gething, M.-J. and Sambrook, J. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 9853–9857
18 Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J. et al. (1994) Nature (London) **368**, 32–38
19 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
20 Waterston, R., Martin, C., Craxton, M., Huynh, C., Coulson, A., Hillier, L., Durbin, R., Green, P., Shownkeen, R., Halloran, N. et al. (1992) Nature Genet. **1**, 114–123
21 McCombie, R. W., Adams, M. D., Kelly, J. M., Fitzgerald, M. G., Utterback, T. R., Khan, M., Dubnick, M., Kerlavage, A. R., Venter, J. C. and Fields, C. (1992) Nature Genet. **1**, 124–131
22 Waddle, J. A., Cooper, J. A. and Waterston, R. H. (1993) Mol. Biol. Cell **4**, 907–917
23 Coulson, A., Waterston, R., Kiff, J., Soulston, J. A. and Kohare, Y. (1988) Nature (London) **335**, 184–188
24 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) J. Mol. Biol. **215**, 403–410
25 Bairoch, A. and Bucher, P. (1994) Nucleic Acids Res. **22**, 3583–3589
26 Fischer, G., Wittmann-Liebold, B., Lang, K., Kiefhaber, T. and Schmid, F. X. (1989) Nature (London) **337**, 476–478
27 Trandahl, C. D., Pao, G. M. and Saier, Jr., M. H. (1992) FASEB J. **6**, 3410–3420
28 Theriault, Y., Logan, T. M., Meadows, R., Yu, L., Olejniczak, E. T., Holzman, T., Simmer, R. L. and Fesik, S. W. (1993) Nature (London) **361**, 88–91
29 Liu, J., Chen, C.-M. and Walsh, C. T. (1991) Biochemistry **30**, 2306–2310
30 Kieffer, J. L., Seng, T. W., Li, W., Osterman, D. G., Handschumacher, R. E. and Bayney, R. M. (1993) J. Biol. Chem. **268**, 12303–12310
31 Wiederrecht, G., Hung, S., Chan, H. K., Marcy, A., Martin, M., Calaycay, J., Boulton, D., Sigal, N. H., Kincaid, R. L. and Siekierka, J. J. (1992) J. Biol. Chem. **267**, 21753–21760
32 Bartling, D., Heese, A. and Weiler, E. (1992) Plant Mol. Biol. **1**, 529–530
33 Walker, J. E., Saraste, M., Runswick, M. J. and Gay, N. J. (1982) EMBO J. **1**, 945–951
34 Smith, T., Ferreira, L. R., Herbert, C., Norris, K. and Sauk, J. J. (1995) J. Biol. Chem. **270**, 18323–18328
35 Billich, A., Hammerschmid, F., Peichl, P., Wenger, R., Zenke, G., Quesniaux, V. and Rosenwirth, B. (1995) J. Virol. **69**, 2451–2461
36 Rassow, J. M., Mohrs, K., Koidl, S., Barthelmess, I. B., Pfanner, N. and Tropschug, M. (1995) Mol. Cell. Biol. **15**, 2654–2662
37 Matouschek, A., Rospert, S., Schmid, K., Glick, B. S. and Schatz, G. (1995) Proc. Natl. Acad. Sci. U.S.A. **92**, 6319–6323
38 Hodgkin, J., Durbin, R. and Martinelli, S. (1995) The Worm Breeders Gazette **13**, 1–109

Received 22 December 1995; accepted 27 February 1996