Identification and Characterization of Moca-cyp

A DROSOPHILA MELANOGASTER NUCLEAR CYCLOPHILIN*

Received for publication, April 18, 2002, and in revised form, July 24, 2002
Published, JBC Papers in Press, August 1, 2002, DOI 10.1074/jbc.M203757200

Laurent Cavarec§§, Thilo Kamphausen, Bérangère Dubourg‡‡, Isabelle Cellebaut**, Françoise Lemeunier‡‡, Didier Métivier‡, Jean Feunteun‡, Gunter Fischer‡, and Nazanine Modjtabahid §§

From the §Laboratoire de Génétique Oncologique, UMR1599, Institut Gustave Roussy-PR1, 39 rue Camille Desmoulins, Villejuif 94805 cedex, France, the §Max-Planck Forschungsstelle "Enzymologie der Proteinfaltung," Weinbergweg 22, Halle/Saale D-06120, Germany, the ‡‡Systèmes moléculaires et Biologie structurale, Laboratoire de Minéralogie et Cristallographie de Paris, CNRS UMR7590, Universités Paris 6 & 7, case 115, 4 place Jussieu, Paris 75005 cedex 05, France, and the §§Laboratoire Populations, Génétique et Évolution, CNRS, Gif-sur-Yvette 91198 cedex, France

Cyclophilins are enzymes catalyzing the cis-trans isomerization of peptidyl-prolyl bonds and belong to the enzyme class of peptidyl-prolyl cis-trans isomerasers (PPIases), which includes two more families (FK506 binding proteins and parvulins). We report the characterization of a novel cyclophilin (Moca-cyp) isolated from Drosophila melanogaster. The single-copy Moca-cyp gene, which is localized on chromosome 3R, was cloned and sequenced. The sequence alignment of the gene against Moca-cyp cDNA allowed us to define its intron/exon structure and to identify a variant cDNA corresponding to an alternatively spliced mRNA. By embryo in situ RNA hybridization and immunostaining, we show that the expression of Moca-cyp is regulated during embryogenesis of Drosophila. The 120-kDa nuclear Moca-cyp protein belongs to a subfamily of large cyclophilins sharing structural and enzymatic features: their highly conserved N-terminal PPIase domain is extended by a positively charged and divergent C-terminal tail. Compared with cyclophilin 18, the enzymatic activity carried by the PPIase domain of Moca-cyp is low, exhibiting its characteristic substrate specificity, and shows a reduced sensitivity to the drug cyclosporin A (CsA). The reduced affinity for CsA is one of the typical features linking members of this subfamily and is probably the consequence of two amino acid substitutions within their active site. Another structural feature shared by members of this subfamily is a conserved polypeptidic segment ("moca" domain) that we report for the first time. The moca domain is located within the C-terminal tail and is the exclusive hallmark of a group of large cyclophilins found in multicellular organisms of the animal kingdom.

Peptidyl-prolyl cis-trans isomerasers (PPIase1, EC 5.2.1.8) are enzymes catalyzing the cis-trans isomerization of peptidyl-prolyl bonds and gather three families of proteins: cyclophilins, FKBP (FK506 binding proteins), and parvulins (for reviews see Refs. 1 and 2). The prototypic cyclophilin (cyp18) was isolated as the cellular receptor for the immunosuppressant drug cyclosporin A (CsA) (3) and shown to be a PPIase (4, 5). In T cells, CsA exerts its immunosuppressive effect by inhibiting the calcium-dependent phases of interleukin-2 production (6). The molecular mechanism underlying this immunosuppressive effect involves the binding of the CsA-cytoplasm complex to the phosphatase calcineurin (7, 8), which in turn leads to the inhibition of the translocation of the nuclear factor of activated T cells (NFAT) transcription factor to the nucleus and prevention of interleukin-2 gene transcription (9). Studies have shown that inhibition of the PPIase activity of cyp18 does not mediate immunosuppressive effects (10).

Since this first discovery, many other family members have been isolated from different tissues and organisms. From bacteria to humans, they are ubiquitously, abundantly expressed and show a very strong conservation of their PPIase domain (11). Cyclophilins have been localized in different cell compartments, and their poorly understood functions have been so far linked to multiple cellular processes, such as protein folding (12), trafficking (13), chaperone activity (14), oxidative stress response (15), and functional organization of components within receptor complexes (16). The abundance and ubiquitous expression of these proteins have suggested that their function might be pleiotropic, but a growing body of evidence points toward the fact that different cyclophilins function in a restricted and specific manner. For instance, their specific localization in different cellular compartments might be an indication of their specificity of substrate and function.

The specific function of a given cyclophilin will depend not only on its PPIase domain structure but also on the composition of the surrounding segments, which will dictate its cellular localization and interaction with partners. Structural and experimental studies of these polypeptidic domains will be of great interest in the understanding of the functional regulation of cyclophilins. This may make it possible to link a particular cyclophilin to a given signal transduction pathway.

For example, cyp18, which does not span behind the PPIase

S-transferase; NLS, nuclear localization signal; BP-NLS, bipartite NLS; pol II, polymerase II; TPR, tetratricopeptide; RT, reverse transcription; Tricine, N-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine; RACE, rapid amplification of cDNA ends; HMM, Hidden Markov Model; HCA, Hydrophobic Cluster Analysis; MBF1, multiprotein bridging factor 1, S6K, S6 kinase; pNA, p-nitroanilide; ORF, open reading frame; CTD, C-terminal domain; FKBP, FK506 binding protein.
domain, is predominantly cytoplasmic (11). Some others, like human
Cyp23 and Cyp22 possess, respectively, secretory and
mitochondrial targeting signal peptides (1). Another cyclophilin,
*Drosophila* ninaA, a membrane protein localized in the
secretory pathway, is responsible of the correct folding of rho-
dopsins in the compound eye (13). The structure of ninaA is in
coherence with its localization, because to its PPIase domain is
attached an N-terminal signal peptide and a very hydrophobic
C-terminal tail. Another subfamily of cyclophilins gather
cyto40-like proteins (17), which are conserved from yeast to
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
humans. Their PPIase domain is extended by a protein-protein
interaction domain containing tetratricopeptide (TPR) repeats
and a calmodulin binding sequence. Through its TPR domain,
sequences. HCA combines sequence comparison with secondary structure predictions and is particularly efficient at low levels of sequence identity (typically below 20–25%). Guidelines to the use of HCA were described elsewhere (32). Secondary structure predictions were performed using the profile neural network prediction PHD program (33).

Three-dimensional model was constructed using Modeler-3 (34), using the experimental structure of the uncomplexed form of Brugia malayi Cytip as template (PDB identifier: 1A58).

**Northern Blot Analysis—**Cytoplasmic RNAs were extracted from *Drosophila* S2 cells as described previously (35). Poly(A+)* RNAs were purified (mRNA purification kit, Amersham Biosciences), resolved by electrophoresis in a 1.2% agarose-formaldehyde gel and transferred to a ZetaProbe membrane (Bio-Rad) in 0.15 M NaH2O4 acetate buffer. After transfer, the membrane was prehybridized for 1 h in Church buffer (36) and then hybridized for 16 h at 65°C in the same buffer supplemented with the 32P-labeled probe (Mega prime kit, Amersham Biosciences). Then the filter was washed twice with 0.5× SSC/0.1% SDS (1× SSC is 0.15 M NaCl, 0.015 M sodium citrate) for 15 min at 65°C and exposed to an autoradiographic film.

**In Vitro Protein Binding Assay—**GST fusion proteins were expressed in bacteria and recovered on glutathione-Sepharose 4B beads (26) (Amersham Biosciences). pMef184 and pBLSuecript532 were in vitro transcribed and translated (TNt-coupled reticulocyte lysate, Promega) in the presence of [35S]methionine (PerkinElmer Life Sciences). Radiolabeled translation products were incubated with the indicated GST fusion protein (400 µl of Net.N supplemented with protease inhibitors (1 mM phenylmethylsulfonyl fluoride, 20 µg/ml each of aprotinin, leupeptin, and pepstatin) for 45 min at 4°C. After five washes in the same buffer, precipitated proteins were boiled in 1× SB, resolved by SDS-PAGE, and visualized by autoradiography, after ENHANCE (PerkinElmer Life Sciences) treatment of the gel.

**Cell Culture and Transfection—**S2 cells were grown at 20°C in Schneider media (Invitrogen) supplemented with 10% heat-inactivated fetal bovine serum (Sigma) and transfected using an Invitrogen calcium phosphate transfection kit and protocols. 24 h after transfection, transfected S2 cells expressing pMTV5-His or pMTV5-His/LacZ or pMTV5-His/Moca-cyp was induced by the addition of 0.5 mM CuSO4 to the growth medium for 24 h. Pools of stably transfected S2 cells expressing pMTV5-His or pMTV5-His/LacZ or pMTV5-His/Moca-cyp were seeded onto chamber slides (LAB-TEK) in the growth media and induced with 0.5 mM CuSO4. After 24 h of induction, cells were fixed with 2% paraformaldehyde and then incubated for 3 min with methanol at room temperature. V5-tagged proteins were detected by incubation with an anti-V5 mouse monoclonal antibody at 1 µg/ml (in phosphate-buffered saline/Tween 0.5%/bovine serum albumin 3%) followed by a goat anti-mouse antibody conjugated to fluorescein isothiocyanate (Jackson ImmunoResearch Laboratories). Samples were DNA-stained with 0.5 µg/ml 4,6-diamidino-2-phenylindole, mounted using Vectashield mounting medium H-1000 (Vector Laboratories) and observed by confocal microscopy (Leica Microsystems SA, HX2 PLAPO CS 63/1.32–0.60 oil).

**In Situ Hybridization—**Whole-mount RNA in situ hybridization experiments were performed according to Gavis and Lehmann (39). Antisense and sense digoxigenin-labeled RNA probes were prepared using the pSKPCR2–502 plasmid (labeling kit, Roche Molecular Biochemicals, according to the supplier’s protocol). Embryos were mounted in Spurr embedding medium (Sigma). In situ hybridizations on polytene chromosomes were prepared as described previously (40).

**Whole-mount Antibody Staining—**All steps of embryo preparation and incubation with the antibody were carried out as described in Gavis and Lehmann (39). The affinity-purified anti-Moca-cyp antibodies were conjugated with fluorescein isothiocyanate (Amersham Biosciences) and used at a 1:200 dilution, and detection was realized with horseradish peroxidase-conjugated secondary antibody using the Vectastain Elite ABC Kit (Vector Laboratories). Embryos were mounted in Spurr embedding medium (Sigma) and analyzed.

**Peptidyl-prolyl cis-trans Isomerase (PPIase) Assay—**PPIase activity was measured according to Hani et al. (41). Inhibition of PPIase activity was measured by preincubating Ca2+ with the enzyme in the mixture for 5 min at 10°C before starting the reaction by the addition of the substrate and the α-chymotrypsin. The inhibition constant KI was calculated according to a slow tight binding model with the following formula,

\[
f = f_{\text{MAX}} \times \frac{([E] + [I])}{([E] + [I] - [I])} + f_{\text{MAX}} \times \frac{[E]}{([E] + [I])} + f_{\text{MAX}} \times [E] 
\]

where [E] is enzyme concentration, [I] is inhibitor concentration, and KI is enzymatic acceleration of the followed reaction.

**RESULTS**

**Moca-cyp cDNA Cloning—**With the aim of searching for nuclear proteins involved in the regulation of p300/CREB-dependent transcription, a protein fusion of glutathione S-transferase (GST) to the CH3 domain of *Drosophila* CREB-binding protein (DCBP) (amino acids 2366–2540) (25) was radiolabeled and used as a probe to screen a *Drosophila* embryo cDNA expression library. Thus, the clone 814 (1.2 kb) was isolated and analyzed. A fragment of the isolated cDNA was subcloned (p814ES) and used as a probe for Northern blot analysis of poly(A+) mRNAs from 814 cells (Fig. 1B). Two poly(A)+ mRNAs of about 3.5 and 4.3 kb were detected. Because the cloned cDNA 814 had a poly(A)+ tail and was shorter than the two mRNA detected by Northern blot analysis, we concluded that it was not complete and the corresponding 5’ end was missing. Another round of cDNA library screening was realized using a p814ES radiolabeled DNA probe. Among multiple cDNA clones isolated, the longest one (clone 532) was fully sequenced and further analyzed (Fig. 1A). The clone 532 contained a poly(A)+ tail and its 5’ end corresponded to the posi-
**Fig. 1. Moca-cyp cDNA cloning.** A, full-length Moca-cyp cDNA sequence and its predicted translation product. Nucleotides and amino acids are numbered on the right. PPlase domain (dark gray), moca domain (brackets), EVES motif (boxed, light gray) are shown. SR repeats are underlined. DSP repeats are boxed. B, Northern blot analysis of Moca-cyp mRNA using poly(A+) mRNA purified from Drosophila S2 cells (lanes 1–3 correspond respectively to 2.5, 5, and 10 μg of poly(A)+). Relative positions of ribosomal RNA (28 S and 18 S) and RNA markers (kb) are, respectively, shown at the left and right side of the figure. C, in vitro association of Moca-cyp with dCBP/CH3. In vitro translated and 35S-labeled partial (814) (lanes 1–3) and full-length (532) (lanes 4–6) Moca-cyp polypeptidic products were incubated with GST (lanes 2 and 5) or with GST-dCBP/CH3 (lanes 3 and 6) immobilized on glutathione beads. After multiple washes, labeled proteins, retained on the beads, were resolved by SDS-PAGE and revealed by autoradiography. Lanes 1 and 4 correspond to the translated products analyzed directly as reference.
tion 40 of the nucleotide sequence presented in figure 1A. The 5' end of the cDNA (Fig. 1A) was obtained by PCR amplification (see “Experimental Procedures”).

The full-length composite cDNA (named Moca-cyp), which spans 3177 nucleotides, contains an open reading frame of 2907 nucleotides (nucleotides 98–3004) and encodes a 968-amino acid protein (Fig. 1A). Sequence analysis of this polypeptide revealed several structural features, which point toward its function. The N-terminal segment, corresponding to the first 180 amino acids, shares the characteristics of globular domains (i.e., approximately one-third of strong hydrophobic amino acids) and shows a strong similarity to the peptidyl-prolyl cis-trans isomerase (PPIase) domain of cyclophilins (11) (Fig. 1, A and D) (see below). The remainder of the polypeptide has a very low complexity, is highly charged (50%), and is mostly hydrophilic. The identification of eight potential bipartite nuclear localization signals (BP-NLS) (profileScan program: hits.isb-sib.ch/cgi-bin/PFSCAN, data not shown) within this C-terminal segment suggested that Moca-cyp could be a nuclear protein (see below for experimental data).

Several potential phosphorylation sites were predicted within Moca-cyp (Phosphobase version 2.0 program, www.cbs.dtu.dk/databases/Phosphobase). For instance, the identification of multiple consensus sequences for each of the following kinases: protein kinase A, protein kinase C, CKI, CKII, CaMKII, S6 kinase, and cdc2, within the low complexity segment of the protein (data not shown), suggested that the potential enzymatic activity of Moca-cyp is targeted by several signal transduction pathways and by the cell cycle control machinery.

The low complexity segment of the protein contains 18% arginine (R) and 12% serine (S) residues. The presence of SR and RS dipeptide repeats (Fig. 1A) suggested that Moca-cyp could belong to SR protein family. Members of this family are arginine/serine (RS)-rich nuclear phosphoproteins involved in the splicing and maturation of mRNA (42). An original hallmark of Moca-cyp is a repeat of a doublet of polar uncharged amino acids glutamine (Q) and asparagine (N) (Fig. 1A, residues 554–572). Between amino acids 482 and 485 (Fig. 1, A and D), we identified a segment composed of amino acids EVES, which motif was initially found within the C-terminal negative regulatory domain of c-Myc transcription factor that interacts with its N-terminal DNA-binding domain. This intramolecular interaction has a negative regulatory effect on c-Myc function (43).

The segment, which spans between amino acids 600 and 800 (Fig. 1A), is highly hydrophilic and corresponds to a repeat region (DSP repeat, Figs. 1A, 1D, and 2B). The core motif of the repeat could be defined as an octapeptide corresponding to the following consensus sequence: (K/R)(K/R)(K/R/X/D/S/P)(XIX) (however, only the serine residue is completely invariant). Many of the phosphorylation consensus sequences, mentioned earlier, were identified within this segment. The analysis of the DSP repeat region, in relation with these potential phosphorylation sites, is noteworthy because it shows that the serine residue in the repeat unit could be targeted not only by the cyclin-dependent kinase (cdc2) but also by protein kinase A, the calcium/calmodulin dependent kinase II, and by the S6 kinase (S6K). Six out of eight BP-NLSs identified in Moca-cyp are localized within this repeat region.

Moca-cyp protein was isolated by virtue of its binding to the CH3 domain of dCBP. The latter shows a high degree of conservation among different members of the p300/CBP family and encompasses the segment of the protein necessary for its binding with the adenoviral oncoprotein E1A (44, 45). Moreover, it is known that the region of E1A involved in the binding with p300/CBP spans the N-terminal and CR1 domain of the viral oncoprotein (46). We observed that the C-terminal region of Moca-cyp (amino acids 861–951) shares a predicted secondary structure organization similar to that of the first 80 amino acids of E1A encompassing the binding region to p300/CBP family members (data not shown). By GST pull-down experiments, we tested the binding capacity of in vitro translated full-length or partial Moca-cyp proteins to the fusion protein pGEX-dCBP/CH3 (Fig. 1C). One can see that both the full-length Moca-cyp and the partial protein that contains the E1A similarity region bind specifically to pGEX-dCBP/CH3. Further mapping of the C-terminal segment of Moca-cyp will allow the assessment of the role of the E1A similarity region in the in vitro binding with pGEX-dCBP/CH3. So far, our searches for the in vivo Moca-cyp-dCBP complex have not been successful. It is possible that the association between Moca-cyp and dCBP needs to be triggered by a specific signal transduction pathway. Indeed, examples of inducible interactions between CBP and other proteins have been reported (47, 48).

Moca-cyp Belongs to a Family of Large Cyclophilins—Moca-cyp belongs to a group of large cyclophilins that share several structural features: their highly conserved N-terminal PPIase domain is extended by a highly charged and divergent C-terminal tail (Fig. 1D). However, this tail is distinguished from that of the cyp40 type cyclophilin, which contains a TPR domain (18). Moca-cyp type cyclophilins do not have any TPR domain, but instead, their C-terminal extensions share a conserved small segment, which we report here for the first time (the “moca” domain, see hereafter).

The combination of features listed below is exclusively characteristic of a small group of large cyclophilins. Therefore, with the aim of distinguishing these proteins from other large cyclophilins, we propose to designate them by the Moca family terminology. The Moca family includes rat matrincyp (23), human srcyp (21, 22), human and mouse NK-TR (49); cyp1 from Caenorhabditis elegans (51) and Moca-cyp isolated from D. melanogaster (this report).

PPIase domain of Moca family cyclophilins cited above were aligned against representatives from other group(s) of cyclophilins (Fig. 2A). The comparison with the prototypic human cyp18 protein shows that, out of the 13 amino acids participating in the PPIase active site (2) (Fig. 2A, star symbols), 2 are not conserved in Moca family members. Ala103 and Trp121 from cyp18 are instead substituted, respectively, by Arg/Lys and His in members of Moca family (Fig. 2A, white arrow symbols). The comparison between the three-dimensional structures of Brugia malayi cyp1 solved without or in complex with CsA (52–54) and those of the well-studied human cyp18 has shed light on the influence of the two amino acid substitutions on the structure of the PPIase domain and its sensitivity to CsA. In Fig. 3, the structure of the uncomplexed form of cyp18 is superimposed to the model of D. melanogaster Moca-cyp PPIase domain, constructed on the basis of uncomplexed B. malayi cyp1 structure. From this comparison, we conclude that, like B. malayi cyp1 (54), the binding site of D. melanogaster Moca-cyp may require substantial changes to accommodate CsA, explaining therefore

Relative positions of molecular weight markers (kDa) are shown at the left side of the figure. D, schematic representation of the characteristic regions of Moca-cyp: PPIase, moca, EVES, and DSP repeat.

2 In this report, human srcyp and its rat homolog matrin cyp will be discussed as one protein matrin/srcyp.
the low affinity of this protein for the latter (see below, Moca-cyp PPIase activity and "Discussion").

Another major structural feature differentiating Moca family members from cyp18 type cyclophilins corresponds to the insertion of eight amino acids (52) in the CD loop of the PPIase domain (Figs. 2A and 3, Moca-cyp, residues 57–64). The function of which is still to be discovered.
Drosophila Nuclear Cyclophilin: Moca-cyp

Moca-cyp is a peptidyl-prolyl cis-trans isomerase

| Substrate, Xaa- | D. melanogaster Moca-cyp $k_{cat}/K_m$ | Human cyp18 $k_{cat}/K_m$ | B. malayi cyp1a $k_{cat}/K_m$ |
|----------------|----------------------------------------|--------------------------|-------------------------------|
| -Ala-          | 56 (100%)                             | 17,600 (100%)            | 130 (100%)                    |
| -Phe-          | 11 (20)                               | 2,700 (15)               | 41 (32)                       |
| -Gly-          | 4.7 (8)                               | 1,800 (10)               | 11 (9)                        |
| -Val-          | 68 (121)                              | 5,400 (31)               | 117 (90)                      |
| -Leu-          | 50 (90)                               | 5,000 (29)               | 690 (530)                     |
| -Ile-          | 49 (85)                               | 4,300 (24)               | 259 (150)                     |
| -Gln-          | 18 (32)                               | 1,000 (59)               | 110 (85)                      |
| -Lys-          | 9.5 (17)                              | 1,600 (9)                | 35 (27)                       |
| -His-          | 4.1 (7)                               | 1,600 (9)                | 2 (1)                         |
| -Glu-          | 2.1 (4)                               | 4,300 (24)               | 60 (46)                       |

* Reported by Ma et al. (56).

Finally, for the first time, we report the identification of a new conserved polypeptidic segment, which characterizes all members of the Moca family. This segment, which we call the moca domain, is situated within the divergent C-terminal tail of the protein (Fig. 1, A and D). The limits of moca domain were highlighted in Drosophila Moca-cyp amino acid sequence (Fig. 1A, amino acids 328–373), using hydrophobic cluster analysis (HCA), because this globular domain was clearly encircled by sequences with a low content of hydrophobic amino acids.

The moca domain was used as a query in PSI-BLAST search (non-redundant data base at NCBI). Significant similarities were revealed with B. malayi cyp1, NK-TR (human and mouse), matrix/cyp (rat and human), and cyp8 (C. elegans). These similarities were further confirmed using HCA and reciprocal PSI-BLAST searches. Because we noticed that the moca domain was only shared by members of the large cyclophilin family, we searched for its presence specifically in these proteins using HCA. This additional search allowed us to find a similar but more divergent motif in C. elegans cyp9 (51). This additional member was further confirmed through PSI-BLAST and profile-based analysis. It is noteworthy that, for all members of the family, the moca domain is located at a comparable distance relative to the PPIase domain and in the case of cyp9 the location of the motif corresponds to the very C-terminal end of the protein.

A multiple alignment of moca domains characterizing all members of Moca family is presented in Fig. 2C. This comparison shows clearly that the domain is composed of two highly conserved motifs (GRG) and (TPPHW) separated by a linker sequence. The composition of the linker is not the same among different members (compare scryp and NK-TR) but is conserved between homologs isolated from different organisms (compare human scryp and rat matrincyp). Secondary structure prediction indicates an all β-fold (Fig. 2C, arrows), without linking the moca domain to an already known fold. Additional searches, using HMMER, with the multiple alignment of all moca domains did not reveal similarities with any other proteins.

The PPIase Domain of Moca-cyp Is Functional and Responsive to Cyclosporin A—To analyze the PPIase activity of Moca-cyp, we subcloned the PPIase homology region into the vector pet28a. After overexpression in bacteria, the PPIase domain was purified and its enzymatic activity was measured according to Hani et al. (41). The substrate specificity was measured using 10 different synthetic peptides having the structure suc-cinyl-Ala-Xaa-Pro-Phe-pNA, where Xaa corresponds to any of the amino acids listed in Table I. We first observed that the PPIase activity of Moca-cyp ($k_{cat}/K_m = 5.6 \times 10^4 \text{M}^{-1}\text{s}^{-1}$), measured with the substrate suc-Ala-Pro-Phe-pNA, was more than two orders of magnitude lower compared with cyp18.

This low activity seems to be a common characteristic of Moca family members. Indeed, published data show that members of this family exhibit PPIase activities, which could be measured in the range of $1.2 \times 10^6 \text{M}^{-1}\text{s}^{-1}$ and $1.3 \times 10^5 \text{M}^{-1}\text{s}^{-1}$ (23, 51, 55, 56).

To compare measurements and appreciate potential substrate specificity, the value obtained with suc-Ala-Pro-Phe-pNA was set to 100% as the reference for both PPIases Moca-cyp and cyp18. We observed that in contrast to cyp18, Moca-cyp showed a preference for large hydrophobic, but non-aromatic, residues at the position preceding the proline (Table I). Again, published data analysis suggested that this specificity could be a common feature shared by Moca family members. For instance, the latest work published by Ma et al. (56) shows that like D. melanogaster Moca-cyp, B. malayi cyp1 exhibits a preference toward large hydrophobic, non-aromatic residues.

The inhibition of the PPIase activity of Moca-cyp was measured in the presence of the substrate suc-Ala-Pro-Phe-pNA and increasing amounts of CsA. The inhibition constant of $K_i = 450 \text{nM}$ was determined by non-linear regression (Fig. 4). It is noteworthy that this value is >100-fold higher than the $K_i = 2.6 \pm 0.4 \text{nM}$ determined for cyp18 (4). Published inhibition constants for other members of the Moca family have been described in an even higher range of concentration than 450 nM (23, 52, 55).

Moca-cyp cDNA Encodes a Nuclear Protein—Amino acid sequence analysis of Moca-cyp, revealing eight BP-NLS, suggested its nuclear localization. To verify this possibility, a polyclonal serum raised against Moca-cyp (serum SE2382) was first
tested by Western blot analysis of S2 cell extracts (Fig. 5A). The preincubation of the serum with the competitor antigenic peptide showed clearly that, among multiple cellular proteins recognized by this one, two were specific (Fig. 5A, compare lanes 1 and 2). The two proteins with apparent molecular masses of 120 and 85 kDa, respectively, were shown to be present in the nuclear fraction of *Drosophila* S2 cell extract (Fig. 5A, lane 3).

Then, to establish the relationship between the cloned protein and those revealed specifically in the cell extract, the Moca-cyp cDNA was cloned into an *in vivo* expression vector (plasmid pMTV5-His/Moca-cyp) and transfected into S2 cells. After immunoprecipitation, either with SE2382 polyclonal serum (Fig. 5B, lane 1) or an anti-V5 monoclonal antibody (Fig. 5B, lane 3), the overexpressed V5-tagged protein was revealed by SE2382 in a Western blot experiment. The cloned Moca-cyp protein, which was specifically recognized by the SE2382 polyclonal, comigrated with the abundant 120kDa endogenous protein (Fig. 5B, lane 5). We concluded that the translation product of the cloned cDNA corresponded to the 120-kDa polypeptide detected *in vivo*. So far, given its very low amount and detection limits, we don’t have further information regarding the 85-kDa protein, but this polypeptide could correspond to the translation product of a splice variant transcribed from Moca-cyp locus (see below).

To confirm the nuclear localization of Moca-cyp protein, in addition to nuclear extract Western blot analysis (Fig. 5A, lane 3), we realized indirect immunofluorescence experiments using *Drosophila* S2 cells overexpressing the Moca-cyp cDNA. For this purpose, stably transfected pools of S2 cells expressing a V5-tagged version of Moca-cyp polypeptide under the control of an inducible metallothionein promoter were analyzed by indirect immunofluorescence (see “Experimental Procedures”), and we observed that the V5-tagged Moca-cyp protein was localized in the nucleus (Fig. 5C). In parallel, as a control, we stained and analyzed a pool of stably transfected S2 cells overexpressing the LacZ gene (Fig. 5C, pMTV5-His/LacZ). Clearly, within these cells the V5-tagged protein did not co-localize with the DNA and was associated with the cytoplasmic compartment. In conclusion, we confirmed that Moca-cyp was a nuclear protein.

**Moca-cyp Gene Is Alternatively Spliced—*In situ* hybridization to polytene chromosomes showed that Moca-cyp was encoded by a single-copy gene at 98D1–2 on the polytene map (Fig. 6A). Later, by screening a *Drosophila* P1 phage genomic library (P1 high density filter, Berkeley *Drosophila* Genome Project, Genomic Systems Inc.) with radiolabeled probes we identified a genomic fragment containing the Moca-cyp locus (BDGP, recombinant P1 phage DS04617, Genomic Systems Inc.). Subsequently, a 5.2-kb fragment, which contained all Moca-cyp cDNA sequence, was subcloned and sequenced (data not shown). Sequence alignment against the gene showed that the isolated Moca-cyp cDNA was composed of eight exons (Fig. 6B). The PPlase domain of the translated protein (residues 1–185, Fig. 1A) was contained within the first four small exons (respectively, 401, 87, 29, and 139 nucleotides). The C-terminal tail (residues 185–968) was encompassed within three small exons (147, 129, and 191 nucleotides) and a very large one (1880 nucleotides).

The existence of at least two mRNA (Fig. 1C) suggested that the Moca-cyp gene could be alternatively spliced. With the aim of characterizing new splice products, using *Drosophila* S2 cell mRNA and gene-specific primers encompassing Moca-cyp cDNA ORF (Fig. 6B, primers 5′Ex and 3′Ex), we reverse-transcribed, amplified, cloned, and sequenced a variant cDNA.

![Fig. 4. Inhibition of Moca-cyp peptidyl-prolyl cis-trans isomerase activity by cyclosporin A. PPIase activity was measured as described in Table I using the substrate Succinyl-Ala-Ala-Pro-Phe-pNA and increasing amounts of CsA Data were fitted using a slow tight binding model.](image)
mRNA was analyzed during Drosophila embryogenesis. For this purpose a digoxigenin-labeled antisense Moca-cyp RNA probe was synthesized and used for in situ hybridization of whole Drosophila embryos (Fig. 7A). After fertilization, at very early stages of embryogenesis preceding the cellularization (stages 1–4, Fig. 7, A1–A4), a homogeneously distributed maternal Moca-cyp mRNA stock could be detected throughout the egg. During the cellularization, we did not detect any staining in the pole cells, suggesting that Moca-cyp mRNA was absent from these cells or its accumulation was below the detection level (Fig. 7A4). At the onset of gastrulation, as the ventral furrow begins to invaginate, Moca-cyp mRNA expression became more confined to the mesoderm (stage 6, Fig. 7A5). During the germ-band elongation stages (stages 8–10, Fig. 7, A6 and A7), the staining was more evident in the mesoderm, the central nervous system, and in the invaginating anterior and posterior midgut. When the germ-band retracts (stages 12 and 13, Fig. 7, A8 and A9), mRNAs were still present in the central nervous system, and in the imaginating anterior and posterior midgut. Finally, when the dorsal closure and epidermis are completed (stage 15, Fig. 7A10), the probe stained mainly the hindgut and the ventral nerve cord, which begins to contract. The specificity of staining with Moca-cyp antisense RNA probe was confirmed by in situ hybridization experiments realized with a sense Moca-cyp RNA probe (data not shown).
Moca-cyp protein distribution during the embryogenesis was analyzed by whole-mount immunostaining with the anti-Moca affinity-purified SE2382 antibodies (see “Experimental Procedures”). At very early stages of embryogenesis, Moca-cyp protein was present as a maternal stock (Fig. 7, B1–B4). At stage 3, when somatic nuclei begin to arrive to the egg surface, the nuclear localization of the protein could already be detected (Fig. 7B3). In keeping with RNA expression analysis (Fig. 7A), we did not detect any staining in the pole cells, suggesting one more time that Moca-cyp was absent or its accumulation was below the detection level (Fig. 7B4). At the onset of gastrulation (Fig. 7B5) and beginning of germ band extension (Fig. 7B6) Moca-cyp was distributed throughout the embryo. Between stages 9 and 11, where the germ band becomes more elongated (Fig. 7, B7 and B8), the staining became more pronounced in the invaginating anterior and posterior midgut, the central nervous system, and in the mesoderm. Between the time that the germ band completes its retraction (stage 13, Fig. 7B9) and the ventral nerve cord continues its contraction (stage 16, Fig. 7B10), in addition to the staining of the nerve cord, those of the pharynx, proventriculus, and hindgut could be detected. The specificity of whole-mount immunostaining was confirmed by experiments realized with the purified anti-Moca-cyp SE2382 antibodies preincubated with a free competitor antigenic peptide (data not shown).

The temporal expression profile of Moca-cyp protein, during different developmental stages, was analyzed by Western blot (Fig. 7C). The 120-kDa Moca-cyp could be detected as a low maternal stock between 0 and 2 h of embryogenesis (Fig. 7C, lane 2, arrow). The amount of this protein was increased, after the cellularization until stage 16 (between 2 and 15 h of embryogenesis, Fig. 7C, lanes 3–7), and decreased to a much lower level at the end of embryogenesis (between 15 and 24 h, Fig. 7C, lanes 8–10). The 85-kDa protein, specifically detected in S2 cells using SE2382 polyclonal serum (Fig. 5A), was also recognized by the affinity-purified anti-Moca-cyp SE2382 in extracts prepared from S2 cells (Fig. 7C, lane 1, black arrowhead) and from embryos aged between 0 and 15 h (Fig. 7C, lanes 7, white arrowhead). Two additional proteins of, respectively, 90 and 82 kDa could also be detected during embryogenesis (Fig. 7C). The first one was specifically found in the embryos aged between 2 and 5 h (Fig. 7C, lanes 3 and 4, white arrowhead), whereas the other one of 82 kDa accumulated during late stages (Fig. 7C, lanes 2–10, star). All detected proteins were specifically recognized by the affinity-purified anti-Moca-cyp SE2382, because we observed that antibodies preincubated with a free competitor antigenic peptide (data not shown) did not reveal any of them. Quality and quantity of the loaded S2 cell and embryo extracts were controlled by Western blot using the polyclonal serum anti-MBF1 (Fig. 7C). Further characterization will be required to establish the relationship between all the single protein bands detected by anti-Moca-cyp SE2382 antibodies and alternative splicing events controlling the expression of Moca-cyp locus. In summary, this set of experiments shows that the spatio-temporal expression of Moca-cyp is controlled during the embryogenesis of Drosophila.

DISCUSSION

Despite the fact that the Drosophila whole genome sequence analysis has predicted 19 cyclophilin-type peptidyl-prolyl cis-trans isomerases (58), only ninaA, a membrane cyclophilin involved in phototransduction, has been described previously (13). In the present study, we report the cloning of a large cyclophilin called Moca-cyp, which is the first Drosophila nuclear cyclophilin to be characterized.

Data base searches and structural feature comparisons allowed us to define a subfamily of cyclophilins, to which Moca-cyp and a few other proteins belong. This family is called “Moca family” and includes: Moca-cyp isolated from D. melanogaster (this report); matrin/sreyp and NK-TR isolated from rat, human, and mouse (21–23, 40); cyp1 isolated from the parasite B. malayi (50); and cypA and cypB from C. elegans (51). The highly conserved N-terminal PPlase domain of these cyclophilins is extended by a positively charged and divergent C-terminal tail. Members of this family share at least four structural features: two amino acid substitutions in the active site, the enlargement of the CD loop in the PPlase domain, the conservation of a polypeptidic segment called moca domain, and finally the presence of RS domains in their C-terminal tail.

Compared with the prototypic cyp18, the PPlase activity of Moca-cyp could be distinguished for three aspects: first, a lower PPlase activity; second, a preference for large, non-aromatic, hydrophobic residues for the amino acid preceding the proline; and finally, a reduced affinity for CsA. Published data analyses suggest that these three enzymatic activity features are a common hallmark for Moca family members. So far, there is no explanation for the first two characteristics, which are most probably the reflection of the structural particularities shared by these proteins. Nevertheless, the comparison between the three-dimensional structures of human cyp18 and cyp1 isolated from B. malayi (52–54) in addition to the mutational analysis of the latter (59) have shown that the two amino acid substitutions (Ala103 → Arg/Lys and Trp121 → His) could be responsible for the reduced affinity of Moca family members for CsA. The reduced affinity of B. malayi cyp1 for CsA, compared with that of human cyp18, has been explained by an incomplete pre-organization of the binding site in the vicinity of the His132, which is substituted by a Trp121 in cyp18. More precisely, Ellis et al. (54) have shown that the formation of a hydrogen bond between the His132 and CsA seems to necessitate a shift in the backbone in this region and an associated displacement of the side chain of Phe71. In contrast, in the case of cyp18, the binding of the CsA does not need any displacement of Trp121. The other substitution, which corresponds to cyp18 Ala103 → Arg/Lys in the Moca family, has also been shown to affect the access of CsA to the active site. Briefly, the side chain of B. malayi cyp1 Lys114 limits the access of CsA to the active cleft by forming a hydrogen bond with the carbonyl oxygen of Gly89 and thus needs to be rotated, with the disruption of the hydrogen bond, in the complexed state (54). Residues Phe71, Gly83, Lys114 and His132 from B. malayi cyp1, discussed above, correspond, respectively, to Phe77, Gly99, Arg120, and His138 in D. melanogaster Moca-cyp.

A specific feature of Moca family members is a segment called moca domain and situated within the divergent C-terminal tail of each protein. A comparative analysis of predicted proteins encoded by different genomes has been undertaken to assess the evolutionary conservation of this polypeptidic segment. This analysis shows that the moca domain is a specific signature of this subfamily of cyclophilin identified only in multicellular organisms of the animal kingdom and could not be found in plants or in unicellular eukaryotic organisms such as yeast.3 One could speculate that the appearance of this polypeptidic domain reflects a specific and complex regulation required for the function of Moca family cyclophilins in the context of multicellular animal organisms. The relevance of this domain to the function of Moca family cyclophilins is so far unknown and will await further investigations.

Others and we have observed that the C-terminal tail of Moca family cyclophilins is extremely basic and contains multiple BP-NLS (23, 52). This observation suggested that the

3 I. Callebaut, P. Dessen, and N. Modjtahedi, unpublished data.
unknown function of these long cyclophilins could be confined predominantly to the nuclear compartment. Indeed, Moca-cyp polypeptide (this report) and matrin/srcyp (22, 23) are nuclear. NK-TR, another member of the family, was found localized in the nucleus and cytoplasm of myeloid cells (60).

What could be the possible functions of the Moca family cyclophilins and how could their PPIase activity influence nuclear events? All members of this family harbor serine/arginine-rich regions (RS domains) within their C-terminal degenerate tail (this report and Refs. 22, 23, 49, and 52). RS domains have been found in SR proteins mostly localized in nuclear speckles where the vast majority of RNA maturation factors are stored and/or assembled (42). The actual model supporting the co-transcriptional splicing of mRNA suggests that polypeptides bearing RS domains mediate a network of protein-protein interactions resulting in an intimate interconnection between the control of transcriptional elongation and splicing (61). The human SR protein matrin/srcyp (22) has been isolated through its direct interaction with the hyperphosphorylated C-terminal domain (CTD) of the large subunit of RNA polymerase II (pol II). Although this Moca family cyclophilin has been found associated with the nuclear matrix and co-localized with splicing factors at nuclear speckles (22, 23), so far its direct function in RNA splicing has not been proved. Besides, given its interaction with the hyperphosphorylated form of the large subunit of RNA pol II, the involvement of matrin/srcyp in the elongation of transcription is also an appealing possibility. Awaiting experimental data, it has been speculated that the PPIase activity of matrin/srcyp could influence formation of macromolecular complexes linking RNA pol II to the splicing machinery, either by modifying the conformation of splicing factors or the CTD of the large subunit of RNA pol II itself (22) or eventually by acting as a molecular chaperone involved in the dynamic regulation of the nuclear speckle domains (23). Indeed, it has been shown in living cells that speckle domains are highly dynamic and respond specifically to the transcriptional activation of nearby genes (62). At the onset of transcription, splicing and SR-domain factors leave the speckles and accumulate at the sites of active transcription. It has also been shown that the phosphorylation of serine residues in the RS domains is required for their recruitment to sites of transcription in vitro (63). One could speculate that phosphorylation events targeting RS domains could influence the molecular function of the Moca family cyclophilins by controlling their association and dissociation from nuclear structures such as speckles and/or from their potential substrate. The comparison of two Moca family members, matrin/srcyp and NK-TR isolated from the embryonic extracts of Drosophila, with respect to their functions and expression patterns in the Drosophila CNS, may provide new insights into the potential functions of cyclophilins in the CNS.
42. Tacke, R., and Manley, J. L. (1999) Curr. Opin. Cell Biol. 11, 358–362
43. Dash, A. B., Orrico, F. C., and Ness, S. A. (1996) Genes Dev. 10, 1858–1869
44. Eckner, R., Ewen, M. E., Newsome, D., Gerdes, M., DeCaprio, J. A., Lawrence, J. B., and Livingston, D. M. (1994) Genes Dev. 8, 869–884
45. Arany, Z., Sellers, W., Livingston, D. M., and Eckner, R. (1999) Cell 77, 799–800
46. Bayley, S. T., and Mymryk, J. S. (1994) Int. J. Oncol. 5, 425–444
47. Zhong, H., Voll, R. E., and Ghosh, S. (1998) Mol. Cell 1, 661–671
48. Nakajima, T., Fukamizu, A., Takahashi, J., Gage, F. H., Fisher, T., Bellen, J., and Montminy, M. R. (1996) Cell 86, 465–474
49. Arany, Z., Sellers, W., Livingston, D. M., and Eckner, R. (1994) Cell 77, 799–800
50. Page, A. P., Landry, D., Wilson, G. G., and Carlow, C. K. (1995) Biochemistry 34, 11545–11550
51. Page, A. P., MacNiven, K., and Hengartner, M. O. (1996) Biochem. J. 317, 179–185
52. Taylor, P., Page, A. P., Kontopidis, G., Holger, H., and Walkinshaw, M. D. (1998) FEBS Lett. 425, 361–366
53. Mikol, V., Ma, D., and Carlow, C. K. S. (1998) Protein Sci. 7, 1310–1316
54. Ellis, P. J., Carlow, C. K. S., Ma, D., and Kuhn, P. (2000) Biochemistry 39, 592–598
55. Rinftet, A., Collins, C., Menard, R., and Anderson, S. K. (1998) Biochemistry 33, 1668–1673
56. Ma, D., Nelson, L. S., LeCoz, K., Poole, C., and Carlow, C. K. (2002) J. Biol. Chem. 277, 14925–14932
57. Cavener, D. R. (1987) Nucleic Acids Res. 15, 1353–1361
58. Rubin, G. M., Yandell, M. D., Wortman, J. R., Gabor Miklos, G. L., Nelson, C. R., Hariharan, I. K., Fortini, M. E., Li, P. W., Apweiler, R., Fleischmann, W., Cherry, J. M., Henikoff, S., Skupski, M. P., Misra, S., Ashburner, M., Birney, E., Boguski, M. S., Brody, T., Brudzynski, P., Celniker, S. E., Chervitz, S. A., Coates, D., Cravechik, A., Gabrielan, A., Galle, R. F., Gelbart, W. M., George, R. A., Goldstein, L. S., Gong, F., Guan, P., Harris, N. L., Hay, B. A., Hoskins, R. A., Li, J., Li, Z., Hynes, R. O., Jones, S. J., Kasht, P. M., Lemaitre, B., Littleton, J. T., Morrison, D. K., Mungall, C., O’Farrell, P. H., Pickeral, O. K., Shue, C., Vosshoph, L. B., Zhang, J., Zhao, Q., Zheng, X. H., and Lewis, S. (2000) Science 287, 2204–2215
59. Ma, D., and Carlow, C. K. S. (1998) Mol. Biochem. Parasitol. 92, 361–365
60. Giardina, S. L., Coffman, J. D., Young, H. A., Potter, S. J., Frey, J. L., Ortaldo, J. R., and Anderson, S. K. (1996) Blood 87, 2269–2274
61. Proudfoot, N. (2000) Trends Biochem. Sci. 25, 290–293
62. Misteli, T., Caceres, J. F., and Spector, D. L. (1997) Nature 387, 523–526
63. Misteli, T., Caceres, J. F., Clement, J. Q., krainer, A. R., Wilkinson, M. F., and Spector, D. L. (1998) J. Cell Biol. 143, 297–307
64. Teigelkamp, S., Achsel, T., Mund, C., Goeth, S. F., Cronshagen, U., Lane, W. S., Marahiel, M., and Luhrmann, R. (1998) RNA (N. Y.) 4, 127–141
65. Horowitz, D. S., Lee, E. J., Mabon, S. A., and Misteli, T. (2002) EMBO J. 21, 470–480
66. Hunter, T. (1998) Cell 92, 141–143
67. Yang, W.-M., Inouye, C. J., and Seto, E. (1995) J. Biol. Chem. 270, 15187–15193
68. Leverson, J. D., and Ness, S. A. (1998) Mol. Cell 1, 203–211
69. Fair, K., Anderson, M., Bulanova, E., Mi, H., Tropschug, M., and Diaz, M. O. (2001) Mol. Cell. Biol. 21, 3589–3597
70. Whitmarsh, A. J., and Davis, R. J. (2000) Cell Mol. Life Sci. 57, 1172–1183
71. Banister, A. J., and Mis, A. A. (2000) Cell Mol. Life Sci. 57, 1184–1192

Drosophila Nuclear Cyclophilin: Moca-cyp