Protein L-Isoaspartyl Methyltransferase from the Nematode *Caenorhabditis elegans*: Genomic Structure and Substrate Specificity†

Ron M. Kagan and Steven Clarke*

Department of Chemistry and Biochemistry and the Molecular Biology Institute, University of California, Los Angeles, California 90095-1569

Received May 4, 1995; Revised Manuscript Received June 14, 1995®

ABSTRACT: We identified a protein L-isoaspartate (D-aspartate) *O*-methyltransferase (EC 2.1.1.77) in the nematode worm *Caenorhabditis elegans*. The methylation of abnormal L-isoaspartyl residues by this enzyme can lead to their conversion to L-aspartyl residues and represents a protein repair step for polypeptides damaged by spontaneous reactions during the aging process. We show that the levels of this enzyme increase 2-fold in *C. elegans* in the dauer larval form, a developmental stage where the organism can survive for extended periods of time. Utilizing degenerate oligonucleotide primers derived from conserved amino acid sequences of mammalian, plant, and bacterial L-isoaspartyl methyltransferases and PCR amplification, we made DNA probes that allowed us to obtain cDNA and genomic DNA clones encoding this enzyme in the nematode. The deduced amino acid sequence is 53% identical to the human enzyme and 29% identical to the *Escherichia coli* enzyme. Overexpression of the cDNA for the *C. elegans* enzyme in *E. coli* gave an active product with micromolar K_m values for L-isoaspartyl-containing peptide substrates and for the methyl donor *S*-adenosyl-L-methionine. No methylation of D-aspartyl-containing peptides was detected under conditions where the human enzyme catalyzed this reaction, suggesting that the ability to methylate D-aspartyl residues in addition to L-isoaspartyl residues was a later evolutionary adaptation of this enzyme. The *C. elegans* gene for the methyltransferase, designated *pcm-1*, was mapped to a single site in a 31 kb region in the central portion of chromosome V. The gene is 3.2 kb in length and includes six introns. Although much smaller, its genomic organization is similar to that of the corresponding mouse gene, with identically positioned intron−exon splice junctions at five of seven sites. We propose that this gene plays an important role in facilitating the long term survival of this organism.

Proteins are susceptible to a wide variety of spontaneous covalent alterations, including oxidation, glycation, deamidation, isomerization, and racemization. These reactions can lead to the aberrant function or the degradation of the modified protein (Stadtman, 1988). Asparagine and aspartate residues are particularly susceptible to deamidation, isomerization, and racemization reactions (Clarke et al., 1992). D-Aspartyl, L-isoaspartyl, and D-isoaspartyl residues have been detected in a variety of cellular proteins, including eye lens crystallins (Masters et al., 1978; Fujii et al., 1994), myelin basic protein (Fisher et al., 1986; Shapira et al., 1988), serine hydroxymethyltransferase (Artigues et al., 1990), calmodulin (Ota & Clarke, 1990), and β-amyloid protein (Roher et al., 1993). Evidence has been presented that calmodulin (Johnson et al., 1987b), epidermal growth factor (George-Nascimento et al., 1990), calbindin (Chazin et al., 1989), and HPr (Brennan et al., 1994) all are less active when altered aspartyl residues are present.

The mammalian protein L-isoaspartate (D-aspartate) *O*-methyltransferase (EC 2.1.1.77) is capable of transferring a methyl group from AdoMet to L-isoaspartyl and D-aspartyl residues, but not to normal L-aspartyl residues (Lowenson & Clarke, 1991a, 1992). Significantly, the enzymatic methylation of these abnormal residues in peptides can lead to their conversion to L-aspartyl forms (Johnson et al., 1987a; McFadden & Clarke, 1987). Such a repair role of this methyltransferase has been observed in proteins as well. Incubation of deamidated calmodulin with the bovine methyltransferase restores its ability to activate calmodulin-dependent protein kinase to 50−68% of that of the native form (Johnson et al., 1987b). Likewise, when the bacterial phosphocarrier protein HPr, deamidated and isomerized at asparagine residues 12 and 38, is incubated with recombinant human L-isoaspartyl methyltransferase, its phosphohydrolytic activity is restored to a level intermediate between that of the double aspartate mutant form and the deamidated form (Brennan et al., 1994). Computer simulations of the age-dependent accumulation of L-isoaspartyl and D-aspartyl residues in human erythrocytes have shown that its methyltransferase can dramatically decrease the levels of these atypical residues (Lowenson & Clarke, 1991b).

L-Isoaspartyl methyltransferase activity has been detected in a broad spectrum of organisms, including Gram-negative bacteria (O'Connor & Clarke, 1985; Fu et al., 1991; Li & Clarke, 1992a), protozoans (R. Kagan, unpublished data), fungi (Johnson et al., 1991; R. Kagan, unpublished data), plants (Johnson et al., 1991; Mudgett & Clarke, 1993),

† This work was supported by National Institutes of Health Grant GM 26020. R.M.K. was supported in part by USPHS Training Grant GM 07185.

* To whom correspondence should be addressed. Telephone: 310-825-8754. Fax: 310-825-1968. E-mail: clarke@ewald.mbi.ucla.edu.

® Abstract published in *Advance ACS Abstracts,* August 15, 1995.

0006-2960/95/0434-10794$09.00/0

invertebrates (Johnson et al., 1991), and a variety of vertebrates including man (O'Connor & Clarke, 1985; Clarke, 1985). The most direct evidence for the biological relevance of this methylation pathway is the significantly diminished stationary phase survival and increased heat-shock sensitivity of *Escherichia coli pcm*$^-$ mutants that lack the protein L-isoaspartyl methyltransferase (Li & Clarke, 1992b). Interestingly, plant seeds, which can often survive for years, have much higher levels of methyltransferase than other parts of the plant (Mudgett & Clarke, 1993).

The widespread distribution of this enzyme and its high degree of cross-kingdom sequence conservation suggest that it plays an important role in the metabolism of damaged proteins that accumulate as cells age. We sought to develop an animal model for the role of this methyltransferase in aging and long-term survival. The nematode *Caenorhadbitis elegans* is well suited for this task. Its short lifespan, well-defined life cycle, and amenability to genetic and biochemical analysis makes it a good model for the study of aging (Johnson et al., 1993). The dauer larval stage of *C. elegans* is a nonaging, developmentally arrested third-stage larva specialized for long-term survival and dispersal (Cassada & Russell, 1975; Riddle, 1988). Dauer larvae are capable of surviving for up to several months without feeding and exhibit reduced metabolic rates (Riddle, 1988; Reape & Burnell, 1990) and RNA polymerase II transcription levels (Dalley & Golomb, 1992). Dauer larvae are also more resistant to chemical, osmotic, thermal, and oxidative stress than other stages (Cassada & Russell, 1975; Larsen, 1993). How do dauer larvae deal with the age-dependent spontaneous damage to proteins? We wished to determine whether this methyltransferase may play an essential role in promoting the long-term survival of both adult worms and dauer larvae analogous to its role in stationary phase survival of *E. coli*. In this report we characterize cDNA and genomic clones for this enzyme in *C. elegans* and determine the kinetic properties of the partially purified nematode methyltransferase enzyme expressed in *E. coli* cells containing an overexpression plasmid.

MATERIALS AND METHODS

Growth and Harvesting of Nematodes. *C. elegans* strain Bristol (N2), kindly provided by Dr. Jorge Mancillas, University of California, Los Angeles, was grown at room temperature on a lawn of *E. coli* strain OP50 cells on NGM agar plates as described (Sulston & Hodgkin, 1988). Liquid cultures of nematodes grown on *E. coli* OP50 cells in S media (Sulston & Hodgkin, 1988) were inoculated from 5-day-old saturated NGM plates and incubated with shaking at 225 rpm at room temperature as described (Sulston & Hodgkin, 1988). Nematodes were harvested at 4 °C by pelleting in 50 mL tubes at 2000*g* for 3−5 min followed by two washes in 0.1 M NaCl. The nematodes were resuspended in 0.1 M NaCl and then sucrose-floated by the addition of an equal volume of ice-cold 60% (w/w) sucrose solution and centrifugation at 1500*g* for 5 min at 4 °C (Sulston & Hodgkin, 1988). The nematodes were then washed in 0.1 M NaCl, resuspended, and agitated for 30 min at room temperature to allow the digestion of any remaining bacteria in their digestive tracts. After an additional wash in 0.1 M NaCl, the nematodes were either used immediately or were stored at −80 °C. Dauer larvae were prepared by growing liquid cultures for 2−3 weeks and then pelleting and washing in 0.1 M NaCl as described above. The nematodes were then resuspended in 1% SDS (w/v) and agitated at room temperature for 30 min to kill any remaining non-dauers (Cassada & Russell, 1975). The dauer nematodes were then pelleted, washed twice with 0.1 M NaCl, and sucrose-floated as described above. After two additional washes, the nematodes were layered on top of a 15% (w/w) ficoll solution and pelleted at 900*g* (Sulston & Hodgkin, 1988). The dauer larvae were then collected, washed twice in 0.1 M NaCl, and either used immediately or stored at −80 °C.

Preparation of C. elegans Cytosol. Nematodes harvested as described above were resuspended in 1 mL of homogenization buffer (50 mM Tris-HCl, 10 mM EGTA, 2 mM EDTA, 0.25 M sucrose, 0.5 mM DTT, 0.5 mM PMSF, 40 μg/mL leupeptin, 40 μg/mL pepstatin, pH 8.0) per 1 g of pelleted nematodes. The nematode suspension was ground in a liquid nitrogen cooled mortar and pestle. The homogenate was thawed and centrifuged at 16000*g* at 4 °C for 10 min to pellet the cellular debris. The recovered cytosol was either used immediately or stored at −80 °C. Protein concentrations were determined after precipitation in 10% trichloroacetic acid using a modification of the Lowry method (Bailey, 1967).

Methyltransferase Assay. The number of methyl groups transferred from AdoMet to exogenous peptide or endogenous substrates was determined using a vapor diffusion assay (Gilbert et al., 1988). The reaction mixture generally consisted of 10 μM *S*-adenosyl-L-[*methyl*-^{14}C]methionine (ICN, 52 mCi/mmol), methyl-accepting substrate, enzyme, and 0.2 M sodium citrate, pH 6.0, to a final volume of 50 μL. Reactions were incubated at 30 °C for 30 min and then quenched with 50 μL of 0.2 M NaOH and 1% (w/v) SDS. A 70 μL aliquot was spotted on a 1.5 cm × 8 cm piece of thick filter paper (Bio-Rad 165−090) that had been prefolded in an accordion pleat. The filter paper was wedged into the neck of a 20 mL scintillation vial containing 6 mL of Safety-Solve counting fluor (RPI), which was then capped and allowed to equilibrate at room temperature for 2 h to allow the [^{14}C]methanol produced by methyl ester hydrolysis to diffuse into the fluor. The filter paper was then removed, and the vials were counted. Each sample was assayed in duplicate and a no-enzyme blank, generally representing less than 0.014 pmol of methyl groups transferred per min, was subtracted as the background.

Preparation of C. elegans Genomic DNA and Southern Analysis. Genomic DNA was prepared from washed, sucrose-floated nematodes as described by Sulston and Hodgkin (1988), with the following modifications. Proteinase K digestion of intact nematodes was carried out for 30 min at 65 °C in 0.1 M Tris-HCl, pH 8.5, 50 mM EDTA, pH 8.0, 0.2 M NaCl, and 1% SDS. Three phenol extractions and two chloroform extractions were performed. The DNA was ethanol-precipitated and used without further purification. Genomic DNA (10 μg) was digested with restriction enzymes (Promega) and resolved on a 0.8% agarose gel. DNA fragments were transferred to a charge-modified PVDF membrane (Immobilon N, Millipore) and then prehybridized and hybridized according to the directions of the manufacturer with a random-primed [α-^{32}P]dCTP labeled probe. Signals were detected by autoradiography.

Preparation of C. elegans Total RNA and Northern Analysis. Total cellular RNA was prepared from washed,

sucrose-floated nematodes essentially as described (Gross et al., 1990). The RNA (10−15 μg) was fractionated on a 1% agarose-formaldehyde gel as described (Ausubel et al., 1994) and transferred to a nylon membrane (Genescreen, Amersham) and prehybridized and hybridized with a random-primed [α-^{32}P]dCTP labeled probe as described (Church & Gilbert, 1984).

PCR Amplification of C. elegans pcm-1 Coding Sequence from a cDNA Library. A λ UniZap XR cDNA library (8 × 10^8 pfu/mL; Stratagene, cat. no. 937006) constructed from poly-dT-isolated RNA from mixed-stage Bristol (N2) nematodes was a gift from Dr. Alexander van der Bliek, University of California, Los Angeles. PCR primers were synthesized on a Pharmacia-LKB Gene Assembler Plus DNA synthesizer and included T7 (pBluescript) primer GTAATACGACTCACTATAGGG, MBM15 primer GCNCCNCAYATGCAYGC (preregion I), and MT1F primer CTCGAATTCGTNYTNGANGTNGG (region I, contains an *Eco*RI linker). The primary PCR reaction mixture contained 10 mM Tris-HCl, pH 9.0, 50 mM KCl, 1% Triton X-100, 4 mM $MgCl_2$, 0.2 mM dNTP mix (Pharmacia), 40 pmol of T7 (pBluescript) primer, 75 pmol of MBM15, and 1 μL of phage lysate in a final volume of 45 μL. The reaction mixture was heated to 98 °C for 5 min to release the phage DNA and then cooled to 95 °C. Taq DNA Polymerase (Promega) (2.5 units) in 5 μL of reaction buffer was added, and 20 cycles of touchdown PCR (Don et al., 1991) were performed on a MJ Research PTC-100 programmable thermal cycler starting with an annealing temperature of 62 °C for 1 min that was decreased by 0.5 °C each cycle. Primer extension was carried out for 2 min at 72 °C, and denaturation was carried out at 95 °C for 40 s. The reactions were then cycled an additional 10 times at a constant annealing temperature of 52 °C. Reaction products were fractionated on a 1% low melting point agarose gel (Gibco-BRL) and purified with the Magic PCR Prep purification system (Promega) as described in Results and reamplified with the T7 (pBluescript) primer and the MT1F primer. Reaction conditions were as described above except that 2.5 mM $MgCl_2$ was used and 20 and 50 pmol of the T7 and MT1F primers, respectively, were used. Thirty amplification cycles were carried out using a 95 °C denaturation step for 30 s, 55 °C primer annealing step for 30 s, and a 72 °C primer extension step for 1 min. A 700 bp PCR product was gel-purified, digested with *Eco*RI and *Kpn*I, and cloned into the plasmid pUC19 and sequenced by dideoxy-chain termination using the ΔTaq Cycle Sequencing Kit (USB/Amersham).

cDNA Cloning of Nematode L-*Isoaspartyl Methyltransferase.* The radiolabeled 700 bp PCR product described above was used to screen 700000 phage λ plaques from the cDNA library described above. The library was propagated in *E. coli* XL-1 Blue MRF′ cells (Stratagene) and plated out in NZY top agar (Becton Dickson Microbiology Systems) on 137 mm NZY plates (50000 pfu/plate). Plaques were transferred in duplicate onto nitrocellulose membranes (Millipore HATF) as described (Sambrook et al., 1989). The membranes were washed at 42 °C in 5× SSC, 0.5% SDS, and 1 mM EDTA for 30 min. The membranes were then prehybridized for 3 h at 65 °C in 5× SSPE, 5× Denhardt's Solution, 0.5% SDS, and 100 μg/mL sonicated, denatured salmon sperm DNA (Sigma) and then hybridized overnight with a random-primed [α-^{32}P]dCTP (ICN; 3000 Ci/mmol) labeled probe (>1 × 10^9 cpm/μg). The final wash was at 60 °C in 1× SSPE and 0.2% SDS for 15 min. The membranes were exposed to Kodak XAR-5 film for 24−48 h with an intensifying screen (Cronex Quanta III, Dupont) at −80 °C. Secondary and tertiary screening yielded two positive plaques. The insert-containing pBluescript SK(−) phagemids were rescued using the ExAssist helper phage and SOLR *E. coli* cells as described by the manufacturer (Stratagene Uni-Zap XR library instruction manual). The rescued phagemids both contained 900 bp inserts and were designated clone 21 and clone 27. Both strands of the clone 21 insert were completely sequenced using the ΔTaq Cycle Sequencing Kit described above. One strand of the clone 27 insert DNA was sequenced to show that its sequence was identical to that of clone 21.

Genomic Mapping of the C. elegans pcm-1 Gene Encoding the L-*Isoaspartyl Methyltransferase.* Duplicate ordered arrays of yeast artificial chromosome (YAC) clones encompassing over 90% of the *C. elegans* genome were obtained from Dr. Alan Coulson (MRC Laboratory of Molecular Biology, Cambridge, U.K.; Coulson et al., 1991). The membranes were prehybridized for 3 h at 65 °C in 30 mM sodium phosphate, pH 6.2, 100 mM NaCl, 1 mM EDTA, pH 8.0, 0.75% (w/v) sarkosyl, and 17% (w/v) dextran sulfate. The 700 bp *pcm-1* cDNA probe was radiolabeled as described above, and the membranes were hybridized overnight at 65 °C. The final wash was in 50 mM NaCl, 15 mM sodium phosphate, pH 6.2, and 0.5 mM EDTA, pH 8.0, at 50 °C. The membranes were autoradiographed as described above.

PCR Amplification of the Genomic Sequence of the C. elegans pcm-1 Gene. A 2.4 kb genomic PCR product encompassing 80% of the *pcm-1* gene was obtained in a reaction using the cDNA-derived primer pair 13F (CCGTTGATCGTGGAGATTTTGCT; 100 bp 3′ to the start codon) and 16R (CGCGACTTGTCAATGGCACGTAG; reverse complement of sequence 20 bp 5′ to the stop codon). The $MgCl_2$ concentration used was 2.5 mM, and 20 pmol of each primer was used. Thirty-five cycles were carried out using a 30 s, 95 °C denaturation step, a 30 s, 55 °C annealing step, and a 3 min 72 °C primer extension step. The 2.4 kb product was gel purified, subcloned into a pUC19 plasmid vector (Gibco-BRL), and completely sequenced using ^{32}P-end-labeled primers and the ΔTaq Cycle Sequencing Kit described above.

The genomic sequence flanking the 2.4 kb product was amplified by inverse PCR (Ochman et al., 1990). The presence of an *Xho*II restriction site near the 5′ end of the *pcm-1* gene was exploited to obtain an inverse PCR product spanning the 3′ end of the gene. Two micrograms of genomic DNA was digested with the restriction enzyme *Xho*II (Promega). The reaction was stopped and treated with Strataclean resin (Stratagene) to remove the enzyme and then ethanol precipitated and resuspended in 40 μL of sterile H_2O, and 2 μL was taken for a ligation reaction using 1 unit of T4 DNA ligase (Gibco BRL) in a final volume of 40 μL. The ligation reaction was incubated at 37 °C for 90 min, and then the DNA was ethanol-precipitated and resuspended in 20 μL of sterile H_2O. Five microliters of the ligated DNA was used in a PCR reaction using the primer pair 42F (TGTCGTAAAGTTGGTGTAGTGCC; 580 bp upstream of the 3′ end of the 2.4 kb product) and 27R (ATTCAACAACCGTGCCATTTCGG; reverse complement of a sequence 430 bp downstream of the 5′ end of the 2.5 kb product).

The PCR reaction was carried out with 20 pmol of each primer, 5 units of Taq DNA polymerase (Promega), 5 units of Taq Extender PCR additive (Stratagene), 2 mM $MgCl_2$, 0.4 mM dNTP mix (Pharmacia), and the Taq Extender reaction buffer consisting of 20 mM Tris-HCl, pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 1% Triton X-100, and 1 mg/mL bovine serum albumin. Thirty-five amplification cycles were carried out with a 30 s, 95 °C denaturation step, a 30 s, 56 °C annealing step, and a 3 min 72 °C primer extension step. The presence of an *Xba*I restriction site near the 3′ end of the *pcm-1* gene was exploited to obtain an inverse PCR product spanning the 5′ end of the gene. The genomic DNA was digested with the restriction enzyme *Xba*I (Gibco-BRL), ligated, and PCR-amplified as described above. The primer pair used was 27R (see above) and 22F (ATGTTTGCAAGAGTACCACGTTG; 850 bp upstream of the 3′ end of the 2.4 kb product, and 210 bp upstream of the *Xba*I site).

Overexpression of the C. elegans L-*Isoaspartyl Methyltransferase in E. coli.* The pBluescript/PCM-1 (clone 21) plasmid containing the complete cDNA coding region for the methyltransferase was digested with the restriction enzymes *Xho*I and *Xho*II (Promega). The overhangs were filled in using Klenow DNA Polymerase (Gibco-BRL). The 840 bp *pcm-1* cDNA fragment was excised from a 1% low melting point agarose gel and purified using the Magic PCR Prep purification system (Promega). The overexpression plasmid pT7-7 (Tabor & Richardson, 1990) was digested with the restriction enzyme *Sma*I (Gibco-BRL) and then treated with shrimp alkaline phosphatase (USB/Amersham). After heat inactivation of the alkaline phosphatase, the digested plasmid was purified using Strataclean resin and then ethanol-precipitated. The 840 bp fragment encoding the recombinant methyltransferase (rPCM-1) was ligated into the linearized, dephosphorylated pT7-7 plasmid and transformed into *E. coli* strain DH5α (Gibco-BRL). Transformants were selected on LB-ampicillin plates and minipreparations of plasmid DNA were subjected to restriction enzyme mapping to identify constructs containing the 840 bp insert in the proper orientation. The pT7-7/rPCM-1 plasmid was then transformed into *E. coli* strain C600 which carries the kanamycin resistance plasmid pGP1-2 and the gene for phage T7 RNA polymerase under the control of the λp_L promoter (Tabor & Richardson, 1990). Transformants were selected on LB-ampicillin + kanamycin plates, and minipreparations of plasmid DNA were subjected to restriction enzyme analysis to identify transformants carrying both plasmids. Overnight cultures of control C600 cells and the rPCM-1 containing C600 cells were used to inoculate 250 mL each of LB media (100 μg/mL ampicillin, 100 μg/mL kanamycin), and the cultures were shaken at 225 rpm at 30 °C until they reached an A_{600} of approximately 0.5. The cells were then heat induced at 42 °C for 30 min followed by a further 90 min growth at 37 °C. The cells were pelleted at 3000*g* in 50 mL tubes and washed twice in homogenization buffer [10 mM sodium phosphate, pH 7.0, 5 mM EDTA, pH 8.0, 10% (v/v) glycerol] and then resuspended in 2 mL of buffer per gram cell pellet with added 0.5 mM DTT, 0.5 mM PMSF, 40 μg/mL pepstatin, and 40 μg/mL leupeptin. The cells were lysed by sonication using three 30 s pulses at setting no. 4 on a Branson Sonifier model W-350. The cellular debris was at 4 °C at 16000*g* for 10 min, and the cytosol was either used immediately or stored at −80 °C.

Partial Nematode Methyltransferase Purification and Kinetic Constant Determinations. Two liters of rPCM-1 overexpressing C600 cells was induced and homogenized as described above. The lysate was centrifuged for 90 min at 100000*g* to pellet the membranes, and 8 mL of the supernatant was applied to a 10 mL DE52 (Whatman) DEAE anion exchange column equilibrated to pH 8.0 at 4 °C in 20 mM Tris-HCl, 0.2 mM EDTA, 0.5 mM DTT, 10% glycerol, and 0.5 mM PMSF, pH 8.0. After loading, the column was washed in equilibration buffer at a flow rate of 0.4 mL/min, and 2 mL fractions were collected. The methyltransferase was eluted isocratically from the column, and the peak activity was found in fractions 11−19. Enzyme from fraction 17 (694 pmol min^{-1} mL^{-1}; specific activity 1100 pmol min^{-1} mg^{-1}) was utilized for kinetic assays. The native enzyme from *C. elegans* cytosol was also purified 8-fold by DEAE chromatography under similar conditions. Enzyme from fraction 15 (29.9 pmol min^{-1} mL^{-1}; specific activity 14.7 pmol min^{-1} mg^{-1}) was utilized to compare the affinities of the native and the recombinant enzymes. Synthetic peptides were obtained as described previously (Lowenson & Clarke, 1991a, 1992).

Kinetic Constant Determinations of the E. coli L-*Isoaspartyl Methyltransferase.* Methyltransferase-overproducing *E. coli* HB101 cells containing extrachromosomal copies of the bacterial L-isoaspartyl methyltransferase on the plasmid pMMkatF1 (Fu et al., 1991) were grown overnight in 1 L of Terrific Broth (Sambrook et al., 1989). The cells were pelleted and disrupted with a French press at 16000 psi, and cytosolic extracts were prepared as described (Fu et al., 1991). Methyltransferase assays were carried out for 20 min at 37 °C with 8 μL of cytosol (300 pmol min^{-1} mL^{-1}; specific activity of 3.4 pmol min^{-1} mg^{-1}) as described above. The *E. coli* cytosol was also purified 13.6-fold by DEAE chromatography under conditions similar to those described above but with a 0−500 mM NaCl gradient. Enzyme from fraction 45 (250 pmol min^{-1} mL^{-1}; specific activity 46 pmol min^{-1} mg^{-1}) was utilized in some of the kinetic assays.

RESULTS

Identification of a Protein L-*Isoaspartyl Methyltransferase Activity in Mixed and Dauer Stage Nematodes.* We identified an L-isoaspartyl methyltransferase in cytosolic fractions of 5-day-old mixed-stage nematode cultures incubated with AdoMet. Although we detected very little endogenous activity in the absence of a methyl-accepting substrate, we found a specific activity of 1.74 pmol of methyl groups transferred per min per mg of protein in the presence of the L-isoaspartyl-containing peptide VYP(isoD)HA (Table 1). This activity was comparable to that found in some plant tissues (Mudgett & Clarke, 1993), bacterial species (Li & Clarke, 1992a), and the protozoan *Tetrahymena* (R. Kagan, unpublished data) but was lower than the specific activity found in mammalian tissues such as human erythrocytes (Gilbert et al., 1988) and bovine brain (Johnson et al., 1991) or in certain plant seeds (Mudgett & Clarke, 1993). We also measured this activity in dauer stage larvae. Again, we detected little or no endogenous activity, but in the presence of the peptide methyl acceptor we found a specific activity over 2-fold greater than that of the 5-day mixed adult and larval nematodes (Table 1).

Isolation of cDNA Clones of the C. elegans L-*Isoaspartyl Methyltransferase.* We exploited the high degree of amino

Table 1: L-Isoaspartyl Methyltransferase Activity in Cytosol from 5-Day Mixed Stage and 3-Week Dauer Nematode Cultures[a]

sample	methyltransferase activity (pmol min^{-1} mg^{-1})	sample	methyltransferase activity (pmol min^{-1} mg^{-1})
5-day mixed, endogenous	0.09 ± 0.06	dauer, endogenous	0.10 ± 0.09
5-day mixed + peptide	1.74 ± 0.18	dauer + peptide	4.00 ± 1.21

[a] Bristol (N2) nematodes grown in liquid culture for 5 days to a mixed stage and 3-week-old dauer nematodes were harvested and assayed for endogenous or L-isoaspartyl-peptide-dependent methyltransferase activity as described in Materials and Methods. The isoaspartyl peptide used was VYP(isoD)HA at a concentration of 100 μM. The data represent three independent experiments each performed in duplicate.

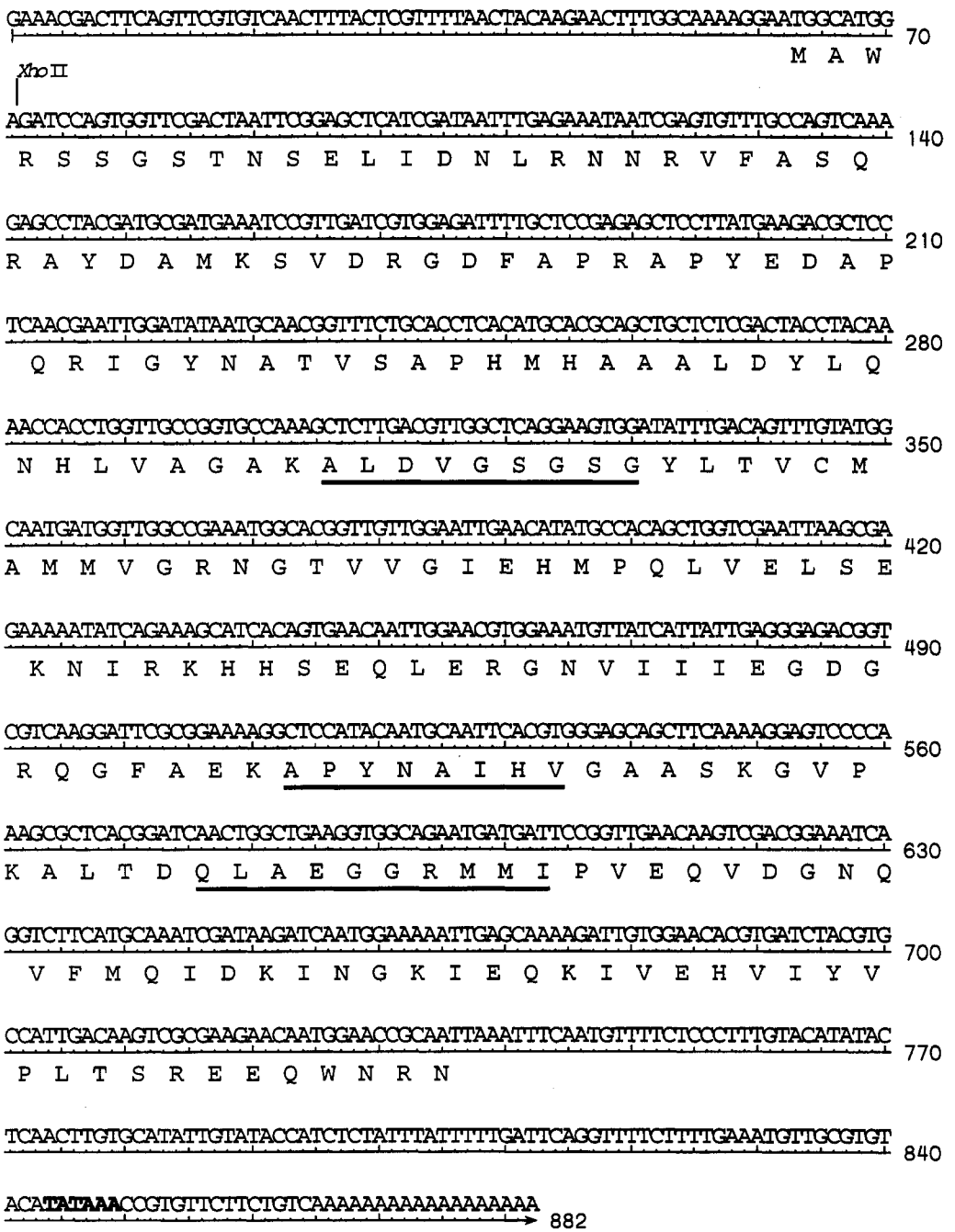

FIGURE 1: Sequence and translation of cDNA clone 21 encoding the *C. elegans* L-isoaspartyl methyltransferase. The pBluescript cDNA clone containing the 882 bp insert encoding the 225 residue *C. elegans* L-isoaspartyl methyltransferase was obtained as described and sequenced. The polyadenylation signal TATAAA at position 844−849 is indicated in boldfaced type. The three methyltransferase AdoMet-binding motifs I, II, and III (Kagan & Clarke, 1994) in the protein sequence are underlined. The sequence has been submitted to Genbank and assigned the accession number U09669.

acid sequence similarity over short regions of the mammalian and plant methyltransferase (Mudgett & Clarke, 1993) to design degenerate oligonucleotides for use in PCR amplification of the *C. elegans* methyltransferase coding sequence from a nematode cDNA library. Initial amplifications with two degenerate primers corresponding to methyltransferase motifs I and II or I and III, or preregion I and motif II or III (Figure 2), yielded only artifactual bands whose translated

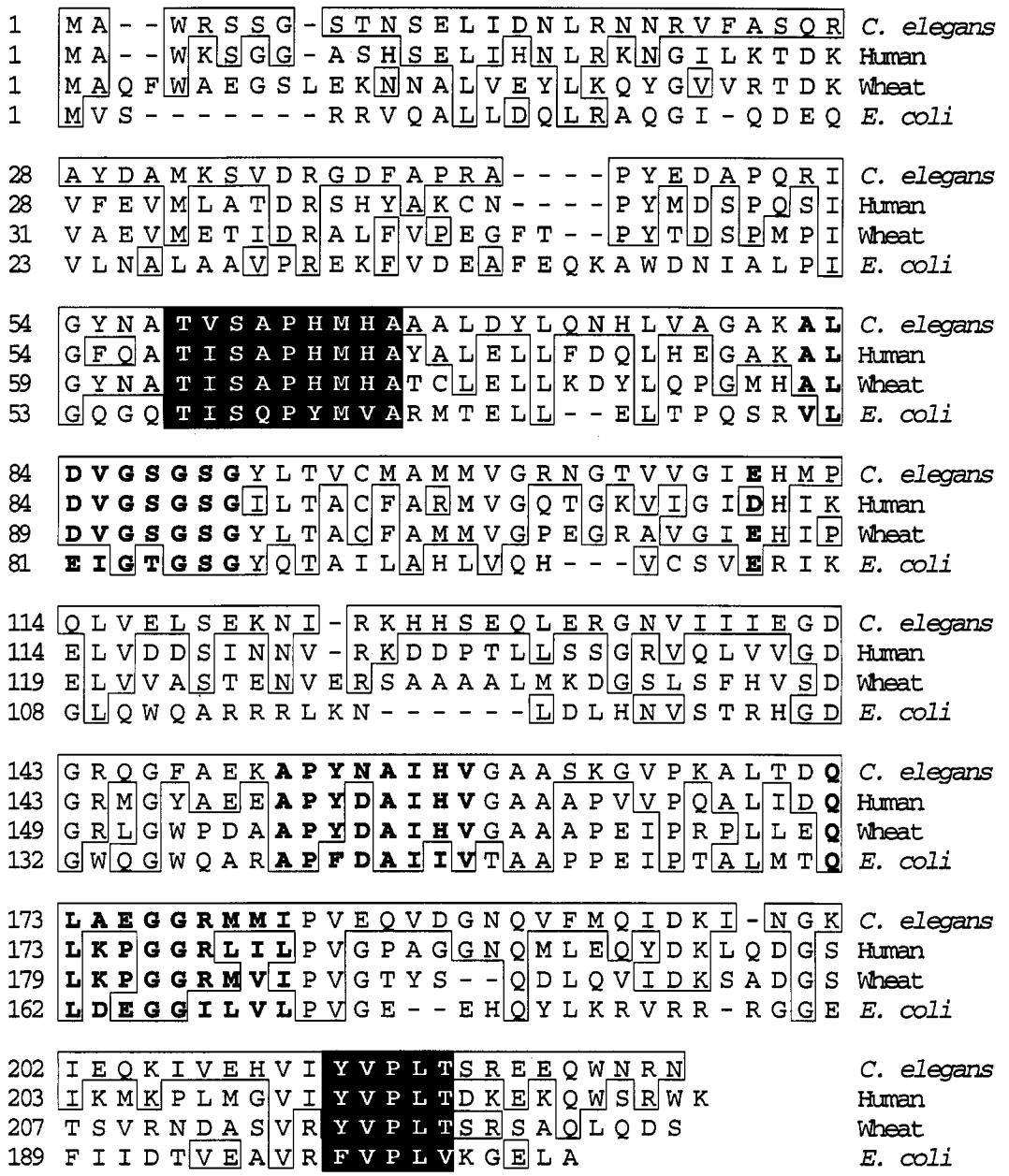

FIGURE 2: Amino acid sequence alignments of *C. elegans* PCM-1 gene product with human, plant, and bacterial L-isoaspartate (D-aspartate) *O*-methyltransferase. The multiple alignment of nematode, human, plant, and bacterial methyltransferases was created with the Megalign program from DNASTAR, Inc. Residues identical to the nematode protein are boxed. The three methyltransferase AdoMet-binding motifs and the conserved acidic residue 19 residues C-terminal to motif I are presented in boldfaced type. Two sequence motifs unique to the L-isoaspartate (D-aspartate) *O*-methyltransferases are presented in reverse type.

sequences showed no resemblance to L-isoaspartyl methyltransferases. Touchdown PCR amplification utilizing the vector-specific T7 primer (adjacent to the expected polyA region of the cDNA) and the L-isoaspartyl methyltransferase-specific 64-fold degenerate primer MBM15, corresponding to preregion I, also yielded a number of products that produced a smear when electrophoresed on a 1% agarose gel. However, Southern blotting of these products utilizing a human L-isoaspartyl methyltransferase cDNA probe detected a clear 780 bp band. The PCR reaction products were then separated on a low melting point agarose gel, and four 5 mm slices from 650−900 bp were excised from the gel. Following purification, the products from each slice were loaded into individual lanes of another agarose gel, separated, Southern blotted, and probed with the mouse methyltransferase cDNA. The amplification product that gave the strongest signal on this blot was then used in a secondary PCR amplification with the T7 primer and the nested 512-fold degenerate methyltransferase primer MT1F, corresponding to methyltransferase motif I. A 700 bp product was obtained that was then digested with *Eco*RI and *Kpn*I and cloned into the corresponding polylinker sites of plasmid pUC19. Sequencing of this product revealed clear deduced amino acid sequence similarities to the mammalian and plant methyltransferases.

The 700 bp fragment, spanning two-thirds of the methyltransferase coding region, was then used to screen the cDNA library for intact cDNA clones. We isolated two positive clones through secondary and tertiary screens. The insert-containing phagemids were rescued by *in vivo* excision,

and the 900 bp insert of one clone (no. 21) was sequenced (Figure 1). The deduced 225 amino acid residue product is 52.7% identical to human L-isoaspartyl methyltransferase, 42.2% identical to wheat methyltransferase, and 29.2% identical to *E. coli* methyltransferase. The deduced methyltransferase sequence contains methyltransferase motifs I, II, and III (Figure 1) that are found in a wide variety of protein, DNA, RNA, and small-molecule methyltransferases (Kagan & Clarke, 1994). The deduced sequence also contains two other motifs found only in other protein L-isoaspartyl methyltransferases (Figure 2). The first precedes region I by 15 residues and has the sequence TISAPHMHA, and the second follows region III by 31 residues and has the sequence YVPLT. Sequencing of one DNA strand of the second cDNA clone revealed no sequence differences with clone 21 (data not shown). Northern analysis of total *C. elegans* RNA with the 700 bp PCR-amplified cDNA probe showed a *pcm-1* major transcript of approximately 900 nt, corresponding to the length of the cloned cDNA (data not shown).

Overexpression of the C. elegans L-*Isoaspartyl Methyltransferase in E. coli.* In order to confirm that the open reading frame that we had cloned coded for an active enzyme, we placed the clone 21 cDNA insert into the expression vector pT7-7 (Tabor & Richardson, 1990). This vector allows transcription of cloned sequences under the control of the T7 RNA polymerase promoter. We exploited the presence of an *Xho*II restriction site at position +10 of the cDNA open reading frame (Figure 1) to clone it in-frame with the initiator methionine in the vector polylinker sequence. As a result, the encoded N-terminus of the construct was altered from MAWRSS in the native protein to MARIRARSS in the recombinant protein. Cytosolic preparations from induced, overexpressing cells yielded a specific L-isoaspartyl methyltransferase activity of 228 ± 10 pmol min^{-1} mg^{-1}, some 130-fold higher than the activity found in crude *C. elegans* cytosol. By comparison, the methyltransferase activity of induced control C600 cells carrying the pT7-7 plasmid with the rPCM-1 construct cloned in the reverse orientation gave a peptide-dependent activity of only 0.021 pmol min^{-1} mg^{-1}.

Substrate Specificity of C. elegans L-*Isoaspartyl Methyltransferase.* Although all of the previously characterized L-isoaspartyl methyltransferases recognize a variety of L-isoaspartyl containing peptides, their affinity for these substrates are variable [Table 3; cf. Lowenson and Clarke (1991a,b)]. Additionally, although clear evidence has been presented for the methylation of D-aspartyl-containing peptides and proteins by the human (McFadden & Clarke, 1982; O'Connor et al., 1984; Lowenson & Clarke, 1992) and bovine (O'Connor et al., 1984) methyltransferase, it does not appear that either the *E. coli* (Fu et al., 1991) or the wheat enzyme (Mudgett & Clarke, 1993) is capable of methylating D-aspartyl residues. We thus decided to compare the ability of the *C. elegans* recombinant enzyme to recognize both L-isoaspartyl, and D-aspartyl-containing peptides. We found that L-isoaspartyl peptides are good substrates for the *C. elegans* methyltransferase. The Michaelis constants of the partially purified, recombinant methyltransferase were determined for a number of L-isoaspartyl-containing peptides (Figure 3 and Table 2). Although the N-terminal of the recombinant enzyme differs at a few positions from the native enzyme (see above), partially purified native *C. elegans*

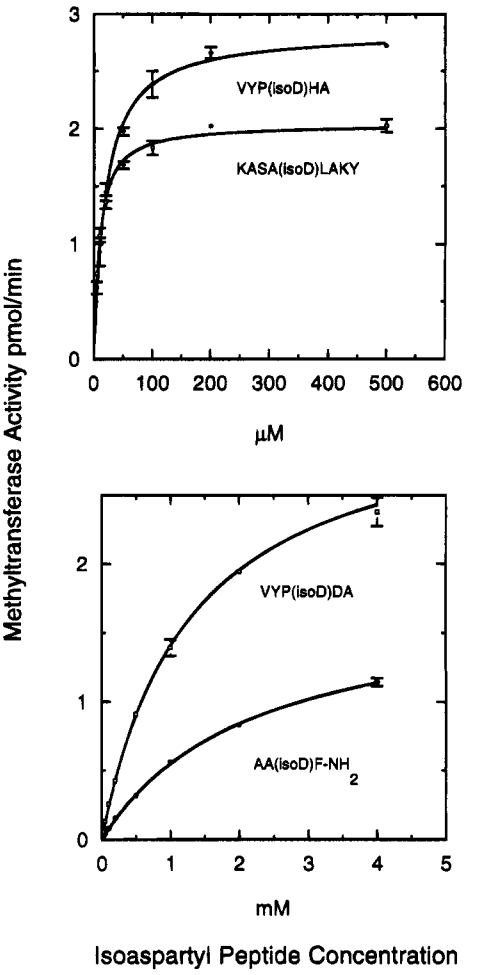

FIGURE 3: Determination of K_m values for the nematode L-isoaspartyl methyltransferase for L-isoaspartyl peptides. Methyltransferase activities were obtained in triplicate determinations as described in Materials and Methods. The error bars indicate one standard deviation. Where no bar is shown, the value of the standard deviation was less than the width of the point. Kinetic constants were determined using the ENZFITTER program as described in Table 2. The solid lines show the theoretical fit to the Michaelis–Menten equation using these values.

enzyme assayed with one of the peptides did not give a significantly different K_m value (Table 2). The K_m values varied over 240-fold between the nonapeptide substrate KASA(isoD)LAKY (9 μM) and the tetrapeptide AA-(isoD)F-NH_2 (2.2 mM), while the maximal velocities showed much less variation (Table 2). The K_m values for the *C. elegans* enzyme for these peptide substrates were found to be 17–67-fold higher than those for the human methyltransferase but were somewhat lower than the values for the wheat methyltransferase (Table 3). The *C. elegans* enzyme also had higher K_m values then the *E. coli* enzyme for three out of the four peptides assayed (Table 3). The human erythrocyte methyltransferase is able to methylate the D-aspartyl-containing peptide KASA(D-Asp)LAKY with a 2.7 mM K_m (Lowenson & Clarke, 1992). However, the same D-aspartate peptide was not recognized by the *C. elegans* recombinant enzyme even at a peptide concentration of 9 mM (Table 4).

Hen egg white ovalbumin is nearly as good a substrate as L-isoaspartyl peptides for the human erythrocyte methyltransferase (K_m = 35 μM; Lowenson & Clarke, 1991a). However, ovalbumin is a poor substrate for the *E. coli* and

Table 2: Kinetic constants of the *C. elegans* L-Isoaspartate *O*-Methyltransferase for AdoMet, L-Isoaspartyl Peptides, and Hen Egg White Ovalbumin[a]

substrate	K_m (μM)	V_{max} (pmol/min)	V_{max} (normalized)	$V_{max\ (norm.)}/K_m$ (M^{-1})
AdoMet	3.05 ± 0.01	3.40 ± 0.01		
KASA(isoD)LAKY	9.12 ± 0.02	2.04 ± 0.01	1	110000
VYP(isoD)HA	19.4 ± 0.5[b]	2.85 ± 0.02	1.40	72000
VYP(isoD)DA	1280 ± 20	3.21 ± 0.02	1.57	1230
AA(isoD)F-NH_2	2220 ± 40	1.77 ± 0.02	0.868	392
ovalbumin	>3000	NA		

[a] Assays were conducted as described in Materials and Methods using the partially purified recombinant enzyme (4 μL; 694 pmol min^{-1} mL^{-1}). The concentration of AdoMet was held at 10 μM for the methyl-acceptor K_m determinations; 200 μM VYP(isoD)HA was used for the AdoMet K_m determination. All amino acid residues in the synthetic peptides were in the L-configuration. Incubations were done for 15 min and were stopped by freezing the reaction tubes in dry ice. The enzymatic activity was determined for 7−8 substrate concentrations, and the kinetic values were calculated from initial velocity vs substrate concentration plots of triplicate determinations for each substrate concentration. The data were fitted to the Michaelis Menten equation by nonlinear regression using the computer program ENZFITTER (Elsevier Biosoft). The K_m and V_{max} values represent mean and the standard error calculated by the program from the nonlinear curve fit. The normalized maximal velocity values were normalized to the V_{max} observed with the KASA(isoD)LAKY peptide. NA, not available. [b] A control assay of native enzyme partially purified on a DEAE cellulose column (15 μL) gave a K_m of 15.8 ± 0.01 μM for this peptide.

Table 3: Comparison of K_m Values of the L-Isoaspartate (D-Aspartate) *O*-Methyltransferase from Bacteria, Plants, Nematodes, and Humans for L-Isoaspartyl Peptides and Hen Egg White Ovalbumin[a]

	KASA(isoD)LAKY (μM)	VYP(isoD)HA (μM)	VYP(isoD)DA (μM)	AA(isoD)F-NH_2 (μM)	ovalbumin (μM)
E. coli	50.6[d]	11.8[d,e]	334[d]	392[b]	727[d]
wheat[c]	12.7	51.7	NA	NA	>1500
C. elegans[d]	9.12	19.4	1280	2220	>3000
human[f]	0.52	0.29	37.7	63.3	35

[a] The kinetic parameters for the recombinant *C. elegans* enzyme are taken from Table 2. The kinetic parameters of the *E. coli* enzyme were determined at 37 °C using a cytosolic extract of an overproducing strain prepared as described in Materials and Methods. The enzymatic activity was determined for 7−8 substrate concentrations, and the kinetic values were calculated from initial velocity vs. substrate concentration plots of triplicate determinations for each substrate concentration by nonlinear regression using the computer program ENZFITTER (Elsevier Biosoft). The K_m and V_{max} values represent mean and the standard error calculated by the program from the non-linear curve fit. [b] Fu et al. (1991) and J. Fu, unpublished data. [c] Values at 25 °C from Mudgett and Clarke (1993). [d] This work. [e] A control assay of native enzyme partially purified on a DEAE cellulose column (10 μL) gave a K_m of 9.2 μM for this peptide. [f] Values at 37 °C from Lowenson and Clarke (1991a).

Table 4: Comparison of the Substrate Specificity of the Human and the *C. elegans* Methyltransferases for a D-Aspartyl-Containing Peptide Substrate[a]

enzyme source	peptide	methyltransferase activity (pmol min^{-1} mg^{-1})	relative activity[b]
C. elegans	KASA(isoD)LAKY[c]	254 ± 39	100%
	KASA(D-Asp)LAKY[d]	0.9 ± 0.2	0.35%
	KASA(D-Asp)LAKY[e]	0.7 ± 0.3	0.28%
	endogenous	0.14 ± 0.06	0.06%
human	KASA(isoD)LAKY[c]	2645 ± 330	100%
	KASA(D-Asp)LAKY[e]	515 ± 25	19%
	endogenous	13 ± 0.4	0.5%

[a] Recombinant *C. elegans* L-isoaspartyl methyltransferase overexpressed in *E. coli* and purified recombinant human L-isoaspartyl (D-aspartyl) methyltransferase (MacLaren & Clarke, 1995) were assayed with either a L-isoaspartyl or a D-aspartyl-containing peptide at 30 °C as described in Materials and Methods. [b] Methyltransferase activity relative to that of the L-isoaspartyl peptide. [c] 140 μM. [d] 750 μM. [e] 9 mM.

the wheat germ methyltransferases (Fu et al., 1991; Mudgett & Clarke, 1993). The *C. elegans* enzyme proved to be similar to the bacterial and wheat germ enzymes in this respect as well. At an ovalbumin concentration of 3 mM, the activity of the recombinant nematode methyltransferase was only 32% of the maximal velocity observed for the L-isoaspartyl peptide VYP(isoD)HA, and saturation could not be reached (Table 2). The Michaelis constants of the *C. elegans* enzyme for the methyl donor AdoMet was also determined. The K_m value of 3.05 μM for AdoMet was close to the 1.9 μM value measured for the bovine brain methyltransferase (Johnson & Aswad, 1993).

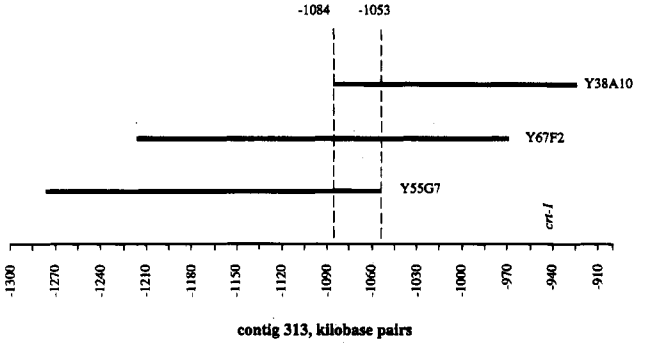

FIGURE 4: Mapping of the *pcm-1* gene to chromosome V. A 700 bp ^{32}P-labeled cDNA probe for the C. elegans *pcm-1* gene was hybridized to an ordered array of YACs encompassing 95% of the *C. elegans genome* (Coulson et al., 1991). The three YACs that gave positive signals mapped to contig 313 and overlapped over a 31 kb region between −1084 and −1053, placing the *pcm-1* gene in the central region of chromosome V. The calreticulin-encoding *crt-1* gene has been previously mapped to positions −942 to −932 (Durbin & Thierry-Mieg, 1993).

Genomic Mapping of the C. elegans Methyltransferase. Southern blotting of *C. elegans* genomic DNA with the cloned 700 bp PCR product detected single restriction fragments in *Bam*HI- and *Hin*dIII-digested genomic DNA (8.7 and 5.6 kb, respectively) suggesting that the methyltransferase was the product of a single gene, designated *pcm-1* (data not shown). Two restriction fragments were detected in *Cla*I- and *Eco*RI-digested genomic DNA, due to the presence of *Cla*I and *Eco*RI sites in the *pcm-1* genomic sequence (see below). In order to ascertain the chromosomal location of the *pcm-1* gene, we obtained an ordered array of yeast artificial chromosome (YAC) clones, encompassing

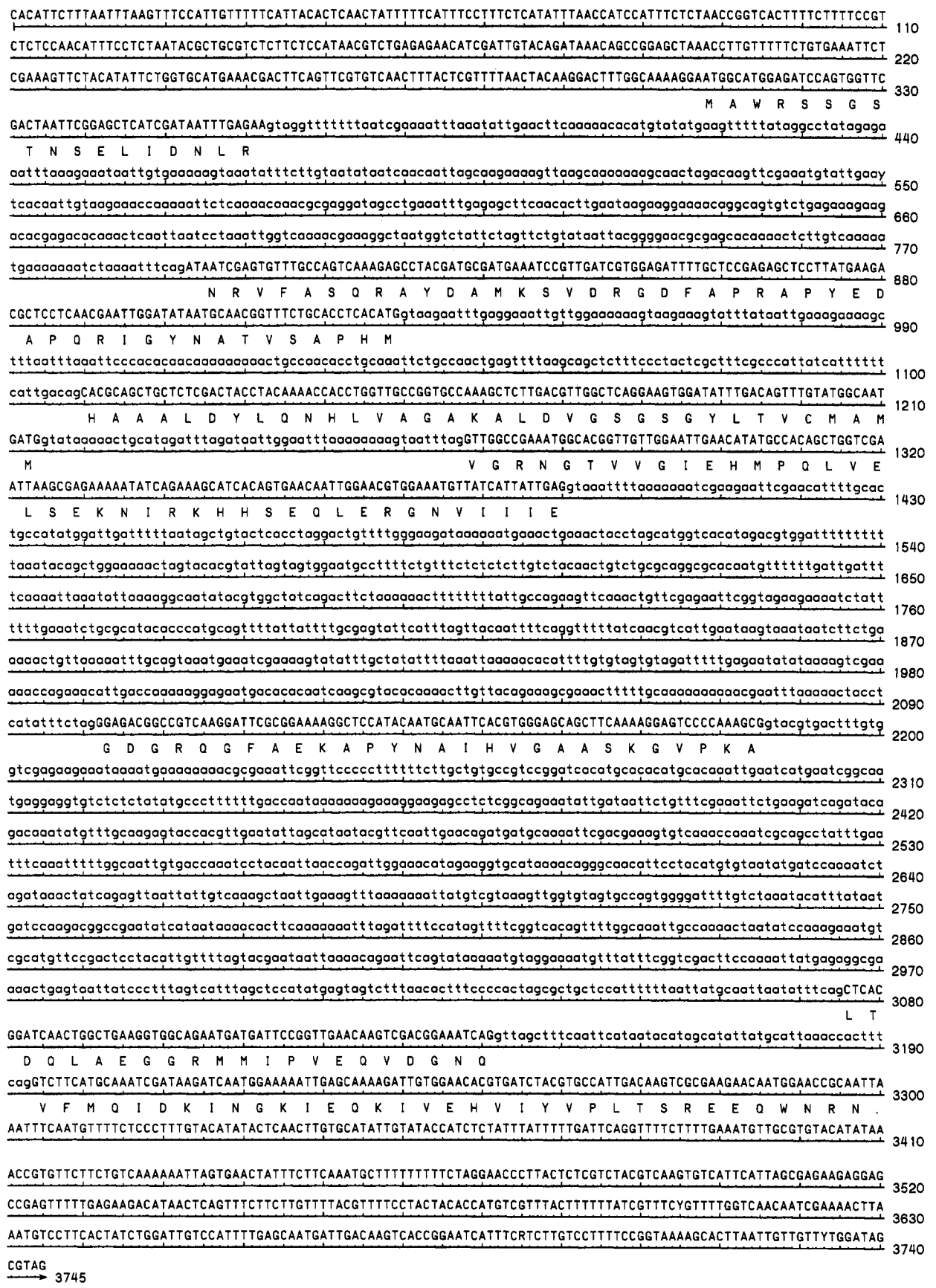

FIGURE 5: Nucleotide sequence encoding the *C. elegans pcm-1* gene on chromosome V. *pcm-1* specific primers were used to obtain one direct genomic PCR product and two inverse genomic PCR products spanning the *pcm-1* gene. The products were cloned and sequenced as described in Materials and Methods. The translation of the seven *pcm-1* exons is shown below the nucleotide sequence. Nucleotide positions 550, 3605, 3695, and 3733 are polymorphic: Y= C or T and R = G or A. The sequence has been submitted to Genbank and assigned the accession number U15129.

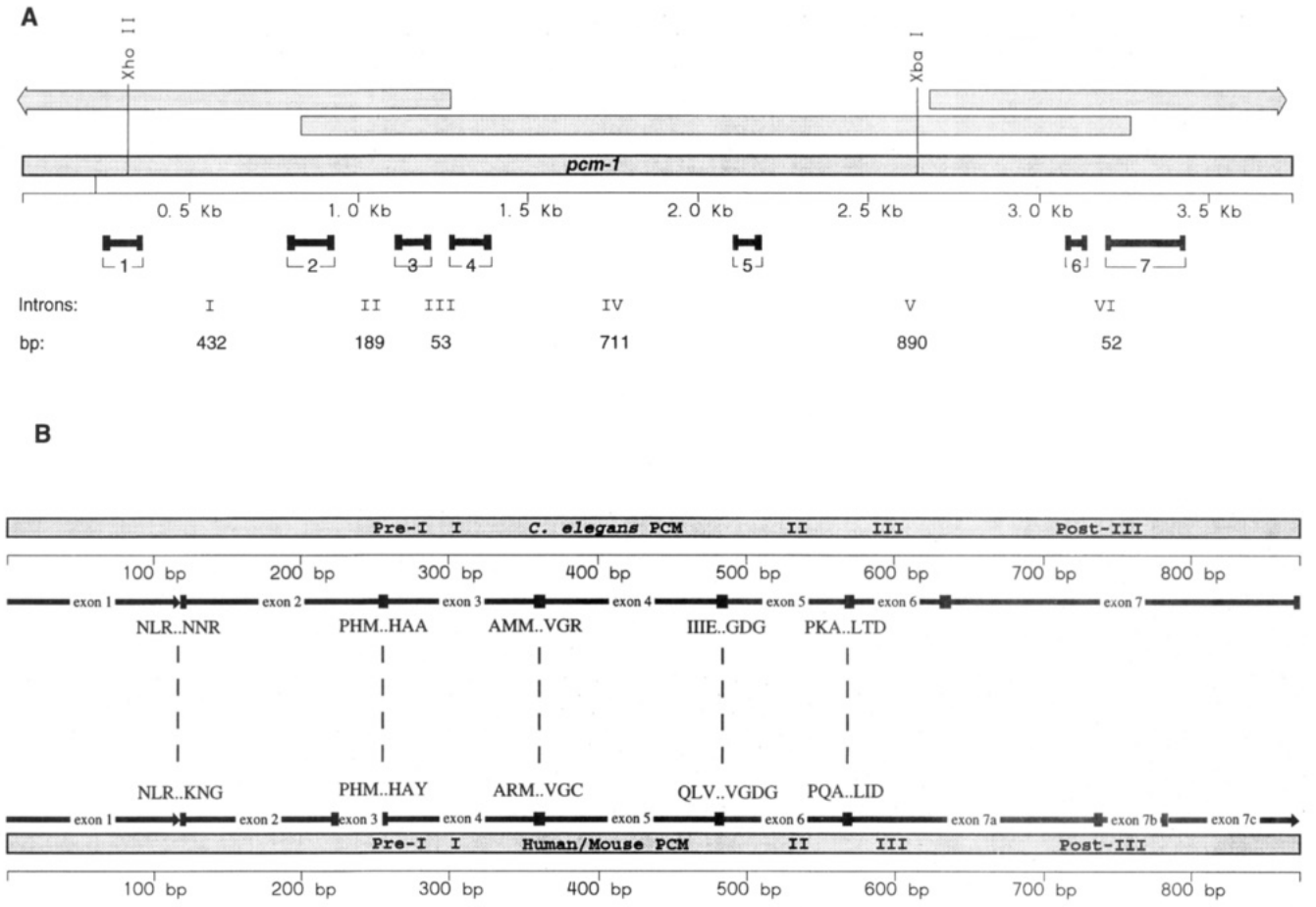

FIGURE 6: Comparison of the genomic organization of the *C. elegans* and mouse genes encoding the L-isoaspartyl methyltransferase. (A) *pcm-1* exon positions are indicated by the solid bars below the scale. Shaded bars A, B, and C represent the direct (A) and inverse (B and C) genomic PCR products used to obtain the genomic nucleotide sequence as described in the text. (B) Comparison of the locations of exon splice junctions for the mouse (Romanik et al., 1992) and the *C. elegans* methyltransferases. The amino acid sequence for the five splice sites conserved between nematode and mouse is given below (nematode) or above (human/mouse) the splice junctions. I, II, III designate the methyltransferase AdoMet-binding motifs; pre-I and post-III designate unique L-isoaspartyl methyltransferase motifs (see Figure 2).

over 90% of the *C. elegans* genome (Coulson et al., 1991). Duplicate arrays were hybridized with the 700 bp cDNA probe. We obtained three positive signals for the YAC clones Y55G7, Y67F2, and Y38A10 (data not shown). These clones map to ctg 313, an 11 megabase contig on chromosome V, and they delineate a 31 kb overlap in the central region of chromosome V (Figure 4; Durbin & Thierry-Mieg, 1993).

Genomic Organization of the C. elegans Methyltransferase Gene. In order to determine the genomic organization of the *pcm-1* gene, we PCR-amplified a 2.4 kb genomic fragment encompassing 80% of the *pcm-1* gene, using the cDNA-derived primer pair 13F and 16R. This product was completely sequenced, and the information was used to design primers for inverse PCR to obtain the remaining 5′ and 3′ sequence of the gene. After *Xho*II digestion and ligation of genomic DNA, the primers 27R and 42F were used in an inverse PCR reaction to amplify a 2.1 kb product. DNA sequencing yielded an additional 470 bp of new sequence information encompassing the 3′ end of the coding region and the 3′ UTR through the polyadenylation site. Digestion of genomic DNA with *Xba*I and ligation followed by an inverse PCR reaction with the primer pair 27R and 22F yielded a 3.3 kb PCR product. DNA sequencing yielded an additional 862 bp of new genomic sequence information, encompassing the 5′ end of the coding region, the 5′ UTR, and the upstream putative promoter region. No TATA box element, or other identifiable promoter motif, was evident in the upstream region of the *C. elegans* methyltransferase gene. In this respect, the nematode gene resembles the TATA-less murine L-isoaspartyl methyltransferase (Romanik et al., 1992).

The nucleotide sequence of the *pcm-1* gene is presented in Figure 5. The gene encompasses seven exons and six introns, spanning 3185 bp from the start of the 5′ UTR to the polyadenylation site. The arrangement of the introns and exons is presented schematically in Figure 6A. The consensi for the *pcm-1* splice donor sites and the splice acceptor sites are AG/GTA and TTTCA/G, respectively, matching the consensi for other surveyed *C. elegans* introns (Felsenstein & Emmons, 1988; Fields, 1990). The nucleotide composition of the introns is 68% AT overall. The largest intron, number V, has a 160 bp region that is 90% identical to the consensus for repetitive element family 357 found in a number of intronic intergenic regions on chromosome III (Agarwal & States, 1994). Interestingly, the positions of the intron−exon splice sites within the coding region of the *C. elegans* gene and the mouse methyltransferase gene (Romanik et al., 1992) are remarkably conserved. Figure 6B shows that the positions of four splice sites are identical in mice and nematodes. The insertion phase of the introns in the coding sequence is also identical in these cases (phase 1 for intron I and phase 0 for introns 2, 3, and 5). A fifth splice site between *pcm-1* exons 4 and 5 (mouse exons 5

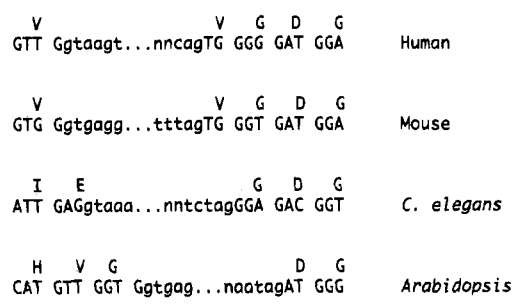

FIGURE 7: Potential intron slippage at a single splice site in human, mouse, nematode, and plant PCM genes. The splice site for intron IV in *C. elegans*, intron V in human (B. Tsai, unpublished results), and mouse (Romanik et al., 1992) and intron III in *Arabidopsis* (M. B. Mudgett, unpublished results) methyltransferase genes is conserved. A slippage of two nucleotides has apparently occurred from nematodes to mammals, and an additional two nucleotide slippage has occurred from nematodes to the *Arabidopsis*.

and 6) is shifted by a single codon. The second intron in the mouse gene is absent in the *C. elegans* gene, resulting in a single exon for the sequence that spans two exons in the mouse gene. In contrast, the splice site between *pcm-1* exons 6 and 7 is absent in the mouse gene. The human methyltransferase gene can undergo alternative splicing at the 3′ end to generate two distinct isozymes (MacLaren et al., 1992) and has additional splice sites for exons 7b, 7c, and 8 that are absent in the *C. elegans* gene. A single intron position is conserved in the human, mouse, *C. elegans*, and *Arabidopsis thaliana* methyltransferase genes (Figure 7). The divergence of the intron position between human and mouse exons 5 and 6, *C. elegans* exons 4 and 5, and *A. thaliana* exons 3 and 4 most likely originated by intron sliding resulting from insertion and deletion of nucleotides (Muller & Schmitt, 1988; Gilbert & Glynias, 1993). The splice acceptor and donor sites of the *C. elegans* gene are shifted by four nucleotides toward the 5′ end of the gene relative to the *Arabidopsis* sequence, and an additional shift of two nucleotides is evident in the human gene (Figure 7).

DISCUSSION

We have identified a protein L-isoaspartyl methyltransferase in *C. elegans*. From sequence analysis of a cDNA encoding this enzyme, we have found greater than 29% identities in deduced amino acid sequence with the *E. coli*, wheat, and mammalian enzymes (Figure 2). This result is consistent with the conserved nature of the L-isoaspartyl methyltransferase sequences seen previously. These enzymes thus meet the criteria of Doolittle's "first edition" ancient proteins that have retained their original function over two billion years of evolution from prokaryotes to eukaryotes (Doolittle et al., 1986). Although the nematode line diverged from the line leading to higher animals some 1000 million years ago (Vanfleteren et al., 1994), the nematode methyltransferase has retained 52.7% amino acid identity to the human enzyme.

We found that the nematode methyltransferase recognizes the same L-isoaspartyl peptide substrates as the human, *E. coli*, and wheat enzymes. However, the affinities of the human methyltransferase for these substrates are 17−67-fold higher than those of the nematode enzyme. The wheat and *E. coli* methyltransferases also exhibit higher K_m values for L-isoaspartyl-containing sequences than the human enzyme (Table 3), suggesting that higher animals may have continued to evolve a higher affinity for L-isoaspartyl-containing sequences. Interestingly, while the human methyltransferase is able to methylate ovalbumin with micromolar K_m values, neither the wheat enzyme nor the *C. elegans* enzyme were able to reach saturation velocity at millimolar concentrations of this protein substrate. Furthermore, the results presented here suggest that the ability to recognize D-aspartyl substrates, a characteristic of the mammalian enzyme, is not present in the nematode, *E. coli*, or wheat enzyme. No methylation of the D-aspartyl-containing peptide KASA(D-Asp)LAKY by the recombinant *C. elegans* enzyme was detected at peptide concentrations of up to 9 mM, while the human enzyme recognizes this peptide with a K_m of 2.7 mM (Table 4). These data may reflect simply a much poorer affinity of the nematode enzyme for all methyl-accepting substrates or the absence of its ability to recognize D-aspartyl sites at all.

Why does the mammalian enzyme appear to recognize substrates with much higher affinities than the nematode enzyme? Computer models have predicted that a 100-fold increase in the K_m of human erythrocyte methyltransferase for an L-isoaspartyl substrate (from 0.5 to 50 μM) would result in an increase from 0.22 to 22 L-isoaspartyl residues per 1000 100 kDa proteins after 120 days (Lowenson, 1991b). In long-lived protein molecules in tissues such as brain and eye lens, this could result in a significant accumulation of these aberrant residues at an earlier age with possible detrimental effects. Thus, the longer lifespans of higher organisms may have provided the selective pressure for the evolution of a mammalian methyltransferase with lower Michaelis constants for its damaged protein substrates. Since the mean lifespan of *C. elegans* is only approximately 20 days, or up to several months in the case of dauer larvae, the higher K_m values of its methyltransferase may not result in a significant accumulation of aberrant aspartate residues. Additionally, the selective pressure that led to the appearance of D-aspartyl methylation in the mammalian enzymes may also be attributed to an increase in lifespan, as D-aspartyl residues accumulate to a much lesser extent than L-isoaspartyl residues over a short period of time (Lowenson, 1991b) and thus may not be significant in short-lived organisms. It is unclear why the affinity of the wheat enzyme for its substrates is also much lower than the mammalian enzyme (Mudgett & Clarke, 1993; cf. Table 3) especially since seeds can remain viable for years. As the water content of viable seeds can be quite low (15−28%; Villiers, 1975), the effective concentration of substrates for the methyltransferase would be expected to be quite high, and thus a higher Michaelis constant may not result in a reduced rate of L-isoaspartyl methylation.

A large number of prokaryotic and eukaryotic methyltransferases and other AdoMet-utilizing enzymes are distinguished by the presence of three sequence motifs, designated I−III, that may be related to the binding of their common cofactor (Kagan & Clarke, 1994). Two recent crystal structures of an adenine-specific *Taq*I DNA methyltransferase (Labahn et al., 1994) and a catechol *O*-methyltransferase (Vidgren et al., 1994) have shown direct contacts of both motif I residues and a downstream glutamate residue with AdoMet and a supporting role of motifs II and III in an adjacent β-pleated sheet structure. These motifs are highly conserved between the L-isoaspartyl methyltransferases of bacteria, plants, and mammals, and now nematodes (Figure 2). Motif I is completely conserved in all four

sequences. The glutamate residue at position 110 in the nematode enzyme, 19 residues C-terminal to motif I, is also conserved in a large number of methyltransferases (Kagan & Clarke, 1994). In the methyltransferase motif II, the near invariant aspartate at position 4 is replaced with an asparagine in the nematode sequence—the only such case in 47 methyltransferase sequences, representing 26 distinct enzymatic activities (Kagan & Clarke, 1994). The recent partial sequencing of the putative L-isoaspartyl methyltransferase gene of *Pseudomonas aeruginosa* also has an asparagine at this position (Genbank accession number D26134). Motif III is also well-conserved. The amino acid sequences of L-isoaspartyl methyltransferases are distinguished from those of other methyltransferases by the presence of two unique motifs. The first is a nine-residue sequence (designated preregion I) about 60 residues from the N-terminus with the consensus sequence TISAPHMHA, and the second (designated postregion III) is the sequence YVPLT near the C-terminus (Figure 2). We hypothesize that these motifs may play an essential role in L-isoaspartate substrate recognition.

We have also identified and characterized the gene for this enzyme, designated *pcm-1*, and found that it is present as a single copy on chromosome V and consists of seven exons and six introns, with an organization remarkably similar to that of the mouse L-isoaspartyl methyltransferase gene (Romanik et al., 1992). However, while the mouse gene is at least 25 kb in length, the nematode gene spans only 3.2 kb. Five intron splice sites are conserved between the nematode and mouse genes. The nematode *pcm-1* splice site between exons 6 and 7 is absent in the mouse methyltransferase gene. The mouse gene has an additional splice site between exons 2 and 3, absent in the nematode *pcm-1* as are the splice sites between exons 7a, 7b, 7c, and 8 (Figure 6). These latter three sites are involved in alternate splicing of the 3′ end of human methyltransferase (MacLaren et al., 1992). No alternately spliced forms of *C. elegans* methyltransferase were isolated in the cDNA library screen, and the genomic sequence shows only a single polyadenylation signal in the 340 bp 3′ to the termination codon. The six introns of the *pcm-1* gene are AT rich (68% overall), and the splice donor and acceptor sites match the established consensus for *C. elegans* genes. However, only two of the six introns (III and VI) are shorter than 75 bp, whereas 64% of surveyed introns in *C. elegans* are shorter than 75 bp [Figure 6; cf. Fields (1990)]. The conserved genomic organization of nematode and mammalian methyltransferase is consistent with an ancestral animal methyltransferase gene with seven introns (five conserved introns, *C. elegans* intron VI and human intron II). Throughout 1000 million years of evolutionary change separating nematodes from mammals, the nematode gene may have lost human intron II and the mammalian gene may have lost *C. elegans* intron VI. It seems unlikely that this genomic arrangement arose through intron gain which should have required the independent insertion of five introns at identical sites in the mammalian and nematode lineages (Moens et al., 1992; Blaxter et al., 1994). The alternate splicing sites at the 3′ end of the human methyltransferase gene may have been a recent modification. A single splice site is conserved between mammalian, nematode, and plant methyltransferase genes (Figure 7). The divergence in position of this splice site may represent intron sliding, resulting from insertion and deletion of nucleotides. Other cases of intron movement in well-conserved genes have been detected in the *Volvox* histone H3 genes (Muller & Schmitt, 1988) and the triose phosphate isomerase genes of *Aspergillus*, maize, and chicken (Gilbert & Glynias, 1993).

Conservation of multiple splice sites between distantly related organisms, such as nematodes and mammals, has been observed in a number of other highly conserved genes. The maize and chicken triose phosphate isomerase genes have five conserved splice sites (Marchionni & Gilbert, 1986) and the vertebrate, plant, and nematode globin genes have a single splice site conserved between *C. elegans* and plants and two other splice sites that are conserved between the plant and the vertebrate genes (Kloek et al., 1993). The *C. elegans ace-1* gene coding for acetylcholinesterase also has a single splice site that is conserved with *Drosophila* and vertebrate ACHE genes (Arpagaus et al., 1994). The high degree of evolutionary conservation in intron positioning indicates that it may have functional significance, and it has been suggested that intron positioning in mRNA may play a role in differential RNA expression (Berman et al., 1990).

What is the role of L-isoaspartyl methyltransferase in *C. elegans*? Under conditions of diminished food supplies or overcrowding, nematode larvae in the L1 stage enter a specialized larval stage known as the dauer stage. Dauer larvae may survive for up to several months without feeding and with a significantly diminished metabolic rate and then resume normal development with no change in the subsequent adult lifespan (Riddle, 1988; Reape & Burnell, 1990). Superoxide dismutase, an enzyme that protects cells from oxidative damage, is detected at 4.5−5-fold higher levels in dauer larvae (Anderson, 1982; Larsen, 1993). In this work, we have observed a 2-fold increase in the specific activity of the L-isoaspartyl methyltransferase in the dauer stage (Table 1). We propose that the increased levels of methyltransferase may help minimize the accumulation of isomerized aspartate residues during this extended period. In *E. coli*, deletion of the *pcm* gene dramatically reduces the survival of cells in stationary phase and increases their sensitivity to heat shock and oxidative stress (Li & Clarke, 1992b), presumably due to the accumulation of aberrant aspartate residues that cannot be repaired. To test this hypothesis in a eukaryotic system, experiments are in progress to obtain Tc1 transposon insertion mutants in which the *pcm-1* gene is inactivated in *C. elegans* (Zwaal et al., 1993).

ACKNOWLEDGMENT

We thank Dr. Jorge Mancillas (University of California, Los Angeles) for providing us with nematode strains, Dr. Alexander van der Bliek (University of California, Los Angeles) for providing us with a cDNA library, and Dr. Alan Coulson (MRC Laboratory of Molecular Biology, Cambridge, U.K.) for providing the YAC array and the YAC clones. We also thank Dr. Mary Beth Mudgett (University of California, Los Angeles) for suggesting the use of the PCR primer MBM15 and Dr. Mudgett and William Tsai (University of California, Los Angeles) for making available the genomic sequences of the *Arabidopsis* and human methyltransferase genes. Finally, we thank Duncan MacLaren (University of California, Los Angeles) for providing us with purified human recombinant methyltransferase.

REFERENCES

Anderson, G. A. (1982) *Can. J. Zool. 60*, 288−291.

Agarwal, P., & States, D. J., Eds. (1994) *The Repeat Pattern Toolkit (RPT): Analyzing the Structure and Evolution of the C. elegans Genome*, 1-1-9 AAAI Press.

Arpagaus, M., Fedon, Y., Cousin, X., Chatonnet, A., Berge, J. B., Fournier, D., & Toutant, J. P. (1994) *J. Biol. Chem. 269*, 9957−9965.

Artigues, A., Birkett, A., & Schirch, V. (1990) *J. Biol. Chem. 265*, 4853−4858.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D., Seidman, J. G., Smith, J. A., & Struhl, K. (1994) *Current Protocols in Molecular Biology*, John Wiley and Sons, New York.

Bailey, J. L. (1967) in *Techniques in Protein Chemistry*, pp 340−341, Elsevier Science Publishing Co., New York.

Berman, S. A., Bursztajn, S., Bowen, B., & Gilbert, W. (1990) *Science 247*, 212−214.

Blaxter, M. L., Ingram, L., & Tweedie, S. (1994) *Mol. Biochem. Parisitol. 68*, 1−14.

Brennan, T. V., Anderson, J. W., Jia, Z., Waygood, E. B., & Clarke, S. (1994) *J. Biol. Chem. 269*, 24586−24595.

Cassada, R. C., & Russell, R. L. (1975) *Dev. Biol. 46*, 326−342.

Chazin W. J., Kordel J., Thulin E., Hofmann T., Drakenberg T., & Forsen S. (1989) *Biochemistry 28*, 8646−8653.

Church, G. M., & Gilbert, W. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1991−1995.

Clarke, S. (1985) *Annu. Rev. Biochem. 54*, 479−506.

Clarke, S., Stephenson, R. C., & Lowenson, J. L. (1992) in *Stability of Protein Phamaceuticals, Part A, Chemical and Physical Pathways of Protein Degradation* (Ahern, T. J., & Manning, M. C., Eds.) pp 1−29, Plenum Press, New York.

Coulson, A., Kozono, Y., Lutterbach, B., Shownkeen, R., Sulston, J., & Waterston, R. (1991) *Bioessays 13*, 413−417.

Dalley, B. K., & Golomb, M. (1992) *Dev. Biol. 151*, 80−90.

Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., & Mattick, J. S. (1991) *Nucleic Acids Res. 19*, 4008.

Doolittle, R. F., Feng, D. F., Johnson, M. S., & McClure, M. A. (1986) *Cold Spring Harbor Symp. Quant. Biol. 51*, 447−455.

Durbin, R., & Thierry-Meig, J. (1993) *ACeDB−A C. elegans Database*, MRC Laboratory of Molecular Biology, Cambridge, U.K.

Felsenstein, K. M., & Emmons, S. W. (1988) *Mol. Cell. Biol. 8*, 875−883.

Fields, C. (1990) *Nucleic Acids Res. 18*, 1509−1512.

Fisher, G. H., Garcia, N. M., Payan, I. L., Cadilla, P. R., Sheremata, W. A., & Man, E. H. (1986) *Biochem. Biophys. Res. Commun. 135*, 683−687.

Fu, J. C., Ding, L., & Clarke, S. (1991) *J. Biol. Chem. 266*, 14562−14572.

Fujii, N., Ishibashi Y., Satoh K., Fujino M., & Harada K. (1994) *Biochim. Biophys. Acta 1204*, 157−163.

George-Nasciemento, C., Lowenson, J., Borissenko, M., Calderon, M., Medina, S. A., Kuo, J., Clarke, S., & Randolph, A. (1990) *Biochemistry 29*, 9584−9591.

Gilbert, J. M., Fowler, A., Bleibaum, J., & Clarke, S. (1988) *Biochemistry 27*, 5227−5233.

Gilbert, W., & Glynias, M. (1993) *Gene 135*, 137−144.

Gross, R. E., Bagchi, S., Lu, X., & Rubin, C. S. (1990) *J. Biol. Chem. 265*, 6896−6907.

Johnson, B. A., & Aswad, D. W. (1993) *Neurochem. Res. 18*, 87−94.

Johnson, B. A., Murray, E. J., Clarke, S., Glass, D. B., & Aswad, D. W. (1987a) *J. Biol. Chem. 262*, 5622−5629.

Johnson, B. A., Langmack, E. L., & Aswad, D. W. (1987b) *J. Biol. Chem. 262*, 12283−12287.

Johnson, B. A., Ngo, S. Q., & Aswad, D. W. (1991) *Biochem. Int. 24*, 841−847.

Johnson, T. E., Tedesco, P. M., & Lithgow, G. J. (1993) *Genetica 91*, 65−77.

Kagan, R. M., & Clarke, S. (1994) *Arch. Biochem. Biophys. 310*, 417−427.

Kloek, A. P., Sherman, D. R., & Goldberg, D. E. (1993) *Gene 129*, 215−221.

Labahn, J., Granzin, J., Schluckebier, G., Robinson, D. P., Jack, W. E., Schildkraut, I., & Saenger, W. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 10957−10961.

Larsen, P. L. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8905−8909.

Li, C., & Clarke, S. (1992a) *J. Bacteriol. 174*, 355−361.

Li, C., & Clarke, S. (1992b) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9885−9889.

Lowenson, J. D., & Clarke, S. (1991a) *J. Biol. Chem. 266*, 19396−19406.

Lowenson, J. D., & Clarke, S. (1991b) *Gerontology 37*, 128−151.

Lowenson, J. D., & Clarke, S. (1992) *J. Biol. Chem. 267*, 5985−5995.

MacLaren, D. C., & Clarke, S. (1995) *Protein Express. Purif. 6*, 99−108.

MacLaren, D. C., Kagan, R. M., & Clarke, S. (1992) *Biochem. Biophys. Res. Commun. 185*, 277−283.

Marchionni, M., & Gilbert, W. (1986) *Cell 46*, 133−141.

Masters, P. M., Bada, J. L., & Zigler, J. J. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 1204−1208.

McFadden, P. N., & Clarke, S. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 2460−2464.

McFadden, P. N., & Clarke, S. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 2595−2599.

Moens, L., Vanfleteren, J., De Baere, I., Jellie, A. M., Tate, W., & Trotman, C. N. A. (1992) *FEBS Lett. 312*, 105−109.

Mudgett, M. B., & Clarke, S. (1993) *Biochemistry 32*, 11100−11111.

Muller, K., & Schmitt, R. (1988) *Nucleic Acids Res. 16*, 4121−4136.

O'Connor, C. M., & Clarke, S. (1985) *Biochem. Biophys. Res. Commun. 132*, 1144−1150.

O'Connor, C. M., Aswad, D. W., & Clarke, S. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7757−7761.

Ochman, H., Medhora, M. M., Garza, D., & Hartl, D. L. (1990) in *PCR Protocols* (Innis, M. A., Gelfand, D. H., Sninsky, J. J., & White, T. J., Eds.) pp 219−227, Academic Press, San Diego, CA.

Ota, I. M., & Clarke, S. (1990) *Arch. Biochem. Biophys. 279*, 320−327.

Reape, T. J., & Burnell, A. M. (1990) *Biochem. Biophys. Res. Commun. 172*, 1013−1021.

Riddle, D. L. (1988) in *The Nematode Caenorhabditis Elegans* (Wood, W. B., Ed.) pp 393−412, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Roher, A. E., Lowenson, J. D., Clarke, S., Wolkow, C., Wang, R., Cotter, R. J., Reardon, I. M., Zurcher, N. H., Heinrikson, R. L., Ball, M. J., & Greenberg, B. D. (1993) *J. Biol. Chem. 268*, 3072−3083.

Romanik, E. A., Ladino, C. A., Killoy, L. C., D'Ardenne, S. C., & O'Connor, C. M. (1992) *Gene 118*, 217−222.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning, a Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Shapira, R., Wilkinson, K. D., & Shapira, G. (1988) *J. Neurochem. 50*, 649−654.

Stadtman, E. R. (1988) *J. Gerontol. 43*, B112−B120.

Sulston, J., & Hodgkin, J. (1988) in *The Nematode Caenorhabditis Elegans* (Wood, W. B., Ed.) pp 587−606 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Tabor, S., & Richardson, C. C. (1990) *J. Biol. Chem. 265*, 8322−8328.

Vanfleteren, J. R., Van De Peer, Y., Blaxter, M. L., Tweedie, S. A. R., Trotman, C., Lu, L., Van Hauwaert, M. L., & Moens, L. (1994) *Mol. Phyogenet. Evol. 3*, 92−101.

Vidgren, J., Svensson, L. A., & Liljas, A. (1994) *Nature 368*, 354−358.

Villiers, T. A. (1975) *Crop Genetic Resources for Today and Tomorrows*, Cambridge University Press, London.

Zwaal, R. R., Broeks, A., van, M. J., Groenen, J. T., & Plasterk, R. H. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 7431−7435.

BI9510035