ELSEVIER

Gene 230 (1999) 137–144

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

Identification and characterization of a serine hydroxymethyltransferase isoform in *Caenorhabditis briggsae*

G.P. Vatcher [a], W.B. Barbazuk [a,1], N.J. O'Neil [a], M.A. Marra [b], T. Ha [a], D.L. Baillie [a,*]

[a] *Institute of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia, Canada, V5A 1S6*
[b] *Genome Sequencing Center, Washington University School of Medicine, St. Louis, MO 63108, USA*

Received 27 October 1998; received in revised form 8 January 1999; accepted 11 January 1999; Received by A. Dugaiczyk

Abstract

In the nematode *Caenorhabditis elegans*, the maternal effect lethal gene *mel-32* encodes a serine hydroxymethyltransferase isoform. Since interspecies DNA comparison is a valuable tool for identifying sequences that have been conserved because of their functional importance or role in regulating gene activity, *mel-32*(SHMT) genomic DNA from *C. elegans* was used to screen a genomic library from the closely related nematode *Caenorhabditis briggsae*. The *C. briggsae* genomic clone identified fully rescues the Mel-32 phenotype in *C. elegans*, indicating functional and regulatory conservation. Computer analysis reveals that CbMEL-32(SHMT) is 92% identical (97% similar) to CeMEL-32(SHMT) at the amino acid level over the entire length of the protein (484 amino acids), whereas the coding DNA is 82.5% identical (over 1455 nucleotides). Several highly conserved non-coding regions upstream and downstream of the *mel-32*(SHMT) gene reveal potential regulatory sites that may bind *trans*-acting protein factors.

Keywords: Conserved regulatory elements; Evolutionary relationship; Nucleotide sequence analysis

1. Introduction

Serine hydroxymethyltransferase (SHMT) is a ubiquitous, highly conserved pyridoxal 5′-phosphate (PLP) dependent enzyme that catalyses the reversible conversion of glycine and 5,10-methylenetetrahydrofolate to serine and tetrahydrofolate. The by-products of this reaction are the major source of one-carbon units for the biosynthesis of purines, pyrimidines, thymidylate, methionine, and lipids (Schirch, 1982; Garrow et al., 1993). Recent evidence suggests that there is a subcellular partitioning of the reactions catalysed by SHMT, with the serine-to-glycine conversion occurring in the mitochondria, and the glycine-to-serine conversion occurring in the cytoplasm (Narkewicz et al., 1996; Kastanos et al., 1997).

A comparison of the determined and predicted amino acid sequences of many SHMTs reveals a striking degree of conservation in this enzyme from bacteria to man. When SHMT isoforms are compared large stretches of completely conserved amino acids, including the active site, are revealed (Garrow et al., 1993; McNeil et al., 1994; Usha et al., 1994; Vatcher et al., 1998).

Estimates using comparative sequence analysis suggest that *C. elegans* and *C. briggsae* diverged anywhere from 20 to 60 million years ago (Emmons et al., 1979; Prasad and Baillie, 1989; Heschl and Baillie, 1990; Kennedy et al., 1993). The two species are still extremely similar in development, morphology, and behaviour but have been evolutionarily separated long enough for unconstrained sequences to diverge considerably. Most highly conserved sequences are actively transcribed or represent functional regulatory elements (Heschl and Baillie, 1990; Kloek et al., 1996).

Abbreviations: *bcl*, B-cell lymphoma; bp, base pair; Cb, *Caenorhabditis briggsae*; Ce, *Caenorhabditis elegans*; CED, cell death; CEH, *C. elegans* homeobox; COL, collagen; *dpy*, dumpy; EMS, ethylmethanesulfonate; *ges*, gut esterase; H/His, histidine; *gpd*, glyceraldehyde-3-phosphate dehydrogenase; HLH, helix–loop–helix; *hsp*, heat shock protein; MEC, mechanosensory abnormal; *mel*, maternal effect lethal; N, asparagine; PCR, polymerase chain reaction; PLP, pyridoxal 5′-phosphate; *rol*, roller; SDS, sodium dodecyl sulfate; SHMT, serine hydroxymethyltransferase; SSC, standard saline citrate; TRA, transformer; *unc*, uncoordinated.

* Corresponding author. Tel.: +1 604-291-4597; fax: +1 604-291-4597.
E-mail address: dbaillie@darwin.mbb.sfu.ca (D.L. Baillie)
[1] Present address: Genome Sequencing Center, Washington University School of Medicine, St. Louis, MO 63108, USA.

0378-1119/99/$ – see front matter
PII: S0378-1119(99)00076-1

A comparison of the SHMT protein and DNA sequences of two very closely related species, *C. elegans* and *C. briggsae*, could reveal the functionally important protein domains and regulatory sites. In fact, previous interspecies sequence comparisons have established that exons and some regulatory sequences are highly conserved between *C. elegans* and *C. briggsae*, whereas introns and non-regulatory sequences are not (Emmons et al., 1979; Rose and Snutch, 1984; Heine and Blumenthal, 1986; Prasad and Baillie, 1989; Kuwabara and Shah, 1994). Diverging flanking sequences allow easy identification of coding regions as highly conserved open reading frames that differ primarily at the third base position of codons (Kennedy et al., 1993; Lai et al., 1996).

Previously, we described the identification and characterization of a *C. elegans* homolog of SHMT, *mel-32*(SHMT). Seventeen ethylmethanesulfonate (EMS) induced mutant alleles of *mel-32*(SHMT) were reported, defining 13 new residues that are potentially essential to SHMT's catalytic activity or structural integrity (Vatcher et al., 1998). Each of the 17 *mel-32*(SHMT) mutations causes a recessive maternal effect lethal (Mel) phenotype. Animals homozygous for these SHMT mutations have no observable mutant phenotype, but their offspring display an embryonic lethal phenotype. These *C. elegans mel-32*(SHMT) mutations represent the first case where SHMT has been shown to be an essential gene. Here, we identify a SHMT homolog in the closely related nematode *C. briggsae* and show that *C. briggsae* genomic DNA fully complements the *mel-32*(SHMT) mutation in *C. elegans*. Computational analysis reveals that the DNA and protein sequences in *C. elegans* and *C. briggsae* are highly conserved. Several highly conserved elements are also detected in the non-coding DNA upstream and downstream of *mel-32*(SHMT). These elements are probably *cis*-acting control sites to which *trans*-acting protein factors bind.

2. Materials and methods

2.1. Growth and handling of C. elegans

All strains of *C. elegans* were grown at 20°C on Nematode Growth Media plates streaked with *Escherichia coli* strain OP50 as a food source. Standard genetic manipulation followed previously described protocols (Brenner, 1974).

2.2. Computer analysis

DNA and protein sequence comparisons were performed with the CLUSTALW (Thompson et al., 1994) and DOTTER (Sonnhammer and Durbin, 1995) programs. Database searches were performed with the BLAST (provided by the NCBI server) and FASTA programs (Pearson and Lipman, 1988; Altschul et al., 1990). PCR primers were designed with the aid of Oligo (Rychlik and Rhoads, 1989).

2.3. Probe preparation

A 2071-bp *Eco*RI (GibcoBRL) fragment of the *mel-32*(SHMT) rescuing subclone pC05.11 (Vatcher et al., 1998) (representing residues 36 421–38 492 of the C05D11 GenBank Accession No. U00048) was double-gel-purified with a Qiagen Qiaquick gel extraction kit. Fifty nanograms of this DNA were mixed with 1 ng of pBluescript II KS^+ (Stratagene) and randomly labeled with [α-P^{32}] dCTP (Amersham) using a random prime kit (Pharmacia) to a final activity of 1.6×10^6 cpm/ml.

2.4. C. briggsae genomic blot

A *C. briggsae* genomic blot carrying 36 000 individual *C. briggsae* fosmid DNA samples (Genome Systems Inc.) was prehybridized in 6× SSC, 0.5% SDS, and 5× Denhardt's for 1 h at 55°C, and then hybridized with the labeled probe at low stringency (6× SSC, 0.5% SDS and 100 μg/ml of salmon sperm at 55°C) for 16 h. The filter was then washed in 2× SSC and 0.5% SDS, twice at 55°C and once at room temperature, and exposed to Kodak x-Omat film for 22 h.

2.5. Preparation of transgenic strains

The *C. briggsae* fosmid clones detected by the blotting procedure (see Section 3.1), G18D16 and G28M02, were kindly provided by the Washington University Genome Sequencing Center. Single colonies were grown overnight and the fosmid DNA harvested by standard alkaline lysis (Sambrook et al., 1989). These fosmids were then injected into the syncytial gonad of adult wild-type N2 (Bristol) *C. elegans* hermaphrodites together with the dominant marker *rol-6(su1006)* (Kramer et al., 1990; Mello et al., 1991) contained on plasmid pCes1943 (S.J.M. Jones and W.B. Barbazuk, pers. commun.) as previously described (Janke et al., 1997). Stable transgenic strains, i.e. those expressing the roller phenotype in successive generations, were PCR tested for the presence of fosmid vector DNA. Primers specific for the pfos-1 vector (Fos-L: 5′CACTTATTCAGGCGTAGC, and Fos-R: 5′ CAATGAAAGACGGTGAGC) were used with control primers R5 and R11 as previously described (Janke et al., 1997).

2.6. Mutant rescue

Males heterozygous for the canonical allele of *mel-32*, *s2518*, were constructed by crossing *mel-32* homozygotes

to wild-type males. These males [genotype *dpy-17(e164) mel-32(s2518) unc-32(e189)/+++*] were crossed to hermaphrodites bearing the transgenic *C. briggsae* fosmids and scored for rescue as previously described (Janke et al., 1997). Briefly, if the *C. briggsae* genomic DNA rescues the *C. elegans mel-32*(SHMT) mutation, then the homozygous *mel-32* F_2 individuals will survive.

3. Results and discussion

3.1. Identification and characterization of a C. briggsae homolog of SHMT

A *C. briggsae* genomic blot was screened with a labeled piece of *C. elegans* genomic DNA representing residues 36421 to 38492 of C05D11 (GenBank entry U00048). This probe contains 1002 bp (almost two-thirds) of the 3′ end of the *mel-32*(SHMT) gene, 414 bp of intergenic sequence, and 767 bp of the 3′ end of gene 1 of C05D11 (a homolog of the *S. cerevisiae* chromosome XV reading frame ORF YOL098c). One nanogram of Bluescript vector was included in the labeling and hybridization reactions (see Section 2.3) because Bluescript has homology to the fosmid vector (pfos-1). This allowed a weak staining of the grid pattern on the *C. briggsae* blot and aided in the identification of positive clones. A large number of false positives (identified by the manufacturer and the Washington University Genome Sequencing Center: list available at http://genome.wustl.edu/gsc/briggsae/false_positives.html) were detected on the blot (data not shown). In addition, three hybridization signals were detected at positions G18D16, G28M02, and G28G12. The identification of these clones allowed a fosmid contig in the *C. briggsae mel-32* genomic region to be constructed (J.E. Schein, pers. commun.). The overlapping fosmids G18D16 and G28M02 were subsequently sequenced by the Washington University Genome Sequencing Center (sequences available at ftp://genome.wustl.edu/pub/gsc1/sequence/st.louis/briggsae/). A dot-matrix analysis of the *C. briggsae* fosmids with the Ce*mel-32*(SHMT) genomic DNA revealed the Cb*mel-32*(SHMT) homolog at positions 25491–27109 of G18D16 (data not shown). When the genomic and intergenic sequences of the *C. elegans* and *C. briggsae mel-32* genes are compared in a dot-matrix plot (Fig. 1), the exons and several highly conserved non-coding regions are revealed (see below). The syntenic array of genes in the flanking sequence has allowed us to identify this gene unambiguously as the *mel-32* homolog, as opposed to the other as-yet unidentified serine hydroxymethyltransferase isoform in *C. elegans*.

Using the previously determined *C. elegans mel-32*(SHMT) sequence, we predicted the translation start and stop sites and the presumed intron/exon junctions for the *C. briggsae mel-32*(SHMT) gene. This prediction was facilitated by the high degree of coding sequence conservation. The predicted *C. briggsae* gene contains four exons coded in 1618 bp of genomic sequence (Fig. 2). The size and positions of the exons, 1455 nucleotides coding for 484 amino acids, are the same in *C. elegans* and *C. briggsae mel-32*(SHMT) (Fig. 4A). Computer analysis reveals that the *C. elegans* and *C. briggsae* coding DNA is 82.5% identical, and the predicted proteins are 92% identical and 97% similar.

Other *C. elegans*/*C. briggsae* comparisons have revealed proteins that have a very high degree of conservation over their entire lengths. MEC-4, an ion channel subunit, is 97% identical between *C. elegans* and *C. briggsae* (Lai et al., 1996). The cuticle collagen proteins COL-12 and DPY-7 are 95 and 93% identical between *C. elegans* and *C. briggsae* (Gilleard et al., 1997a,b). For COL-12, it is known that intermolecular interactions occur throughout the full length of the protein and that most substitutions have steric effects that interfere with the ability of the monomers to trimerize efficiently or participate in higher-order interactions (Gilleard et al., 1997a). The strong constraint on evolutionary divergence throughout the whole lengths of COL-12, DPY-7, and MEL-32(SHMT) suggests that these proteins have complex higher-order interactions. In fact, other isoforms of SHMT are known to form either dimers or tetramers of identical subunits.

In contrast to the above examples, some *C. elegans*/*C. briggsae*, comparisons have revealed clear regions of greater and lesser homology within a given polypeptide. Areas of high homology may indicate the presence of distinct functional domains, suggesting that these regions play crucial roles in the structure or function of these proteins. CED-9 (a *bcl-2* homolog) is only 70% identical overall, but there are stretches with a very high similarity (Hengartner and Horvitz, 1994). The *C. elegans* and *C. briggsae* HLH-1 proteins (homologs of the MyoD family of helix–loop–helix transcription factors) share a 68% amino acid identity overall, but the basic helix–loop–helix region has 100% identity, whereas the final exon has only 30% identity (Krause et al., 1994). *C. elegans* and *C. briggsae* TRA-1 (a sex-determining gene) share only a 44% amino acid identity, but there are blocks of high identity separated by greatly diverged stretches (de Bono and Hodgkin, 1996). Finally, the homeodomains of *C. elegans* and *C. briggsae* CEH-24 are identical, while the rest of the protein shares only a 79% identity (Harfe and Fire, 1998).

The extremely high conservation of *mel-32* is spread over the entire length of the protein and does not allow for the identification of specific amino acids important for function, but suggests that there is strong evolutionary pressure on the majority of the protein sequence. It may be that substitutions in the SHMT sequence interfere with the ability of the protein to form com-

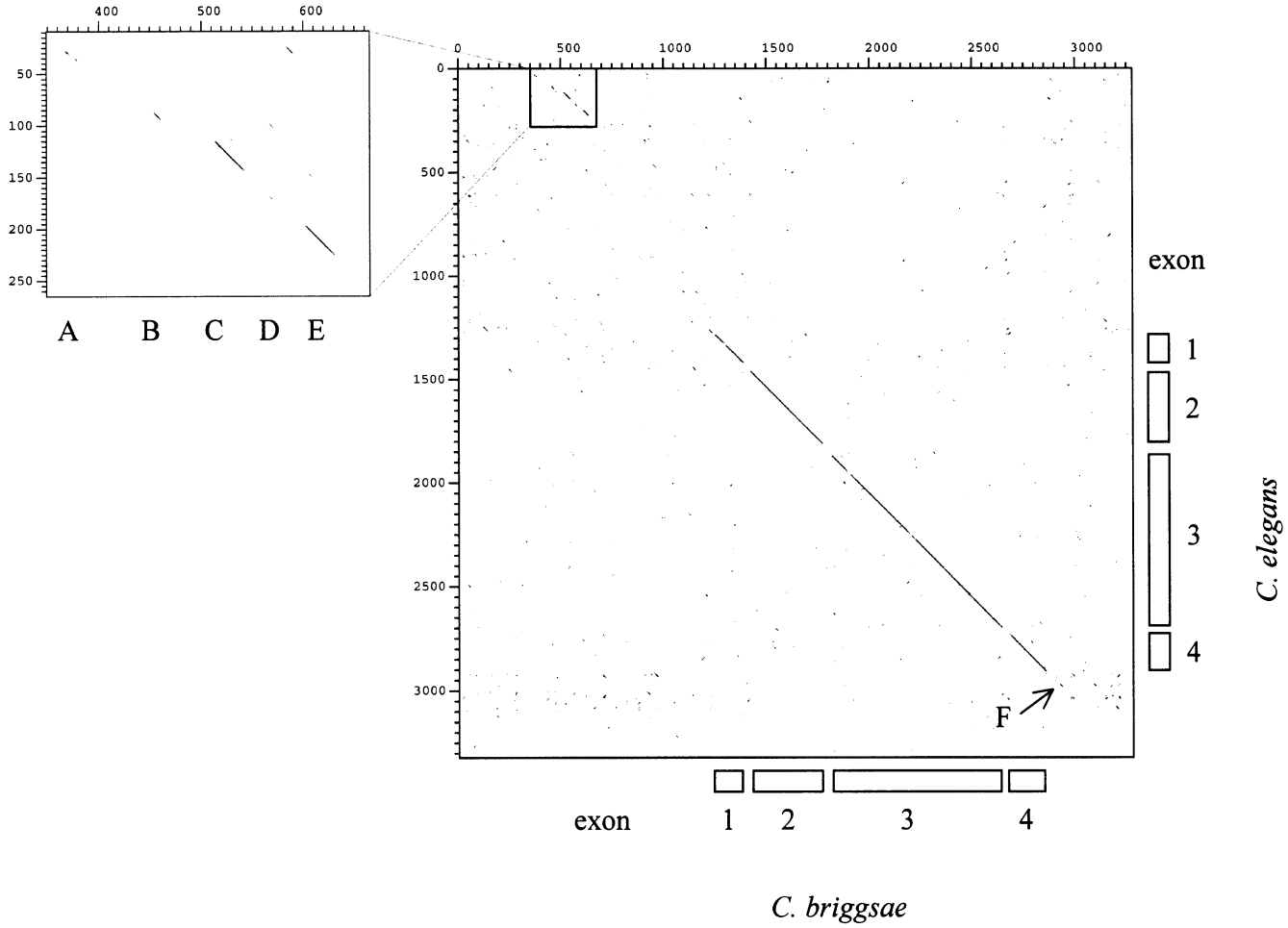

Fig. 1. Dot-matrix plot of the *mel-32* genomic region. The *C. briggsae* sequence is on the horizontal axis, and the *C. elegans* sequence is on the vertical axis. The *mel-32* exons are shown outside the alignment. The inset shows a blow-up of the region that contains the upstream conserved non-coding sequences. A, B, C, D, E, and F correspond to the conserved regions in Fig. 4 and Table 1.

plexes. The sites of amino acid changes probably identify positions in the three-dimensional structure of SHMT which are the most flexible to change while still maintaining the protein's physiological functions. Important structural or enzymatic domains would be less able to tolerate amino acid replacements.

Caenorhabditis introns have few conserved sequences, except for splice sites. Where introns do exist in both species, their positions and splice donor/acceptor sites are generally conserved: the globin genes (Kloek et al., 1996), *mec-4* (Lai et al., 1996), *tra-2* (Kuwabara, 1996), *hsp-3* (Heschl and Baillie, 1990), *gpd-2* and *gpd-3* (Lee et al., 1992), *ges-1* (Kennedy et al., 1993), *dpy-7* (Gilleard et al., 1997b), *unc-119* (Maduro and Pilgrim, 1996), and *ced-9* (Hengartner and Horvitz, 1994) have conserved intron positions and splice sites. The position and splice donor and acceptor site consensus sequences of *mel-32* (SHMT) are also identical between *C. elegans* and *C. briggsae* (Fig. 2). As expected, the intronic sequences and sizes are not conserved. The three intron sizes for *C. elegans* and *C. briggsae* are 45/54, 48/63, and 52/46 nucleotides, respectively (Fig. 4A). The intronic DNA is 57.5% identical over 145 and 163 total nucleotides, with most of the homology residing in the conserved splice sites. *mel-32* does not conform to the general pattern of smaller and fewer introns in *C. briggsae*. The normal disparity in intron size and number between *C. elegans* and *C. briggsae* has been observed in many genes: *mec-3* (Xue et al., 1992), *ges-1* (Kennedy et al., 1993), *tra-1* (de Bono and Hodgkin, 1996), and *unc-119* (Maduro and Pilgrim, 1996).

3.2. *Rescue of the* C. elegans mel-32*(SHMT) mutation with* C. briggsae *genomic DNA*

Interspecies rescue experiments show that *C. elegans* and *C. briggsae* proteins are functionally equivalent, and that transcriptional and translational control mechanisms have been maintained through tens of millions of years of evolution. *C. briggsae* genomic DNA has been used to rescue a *C. elegans hlh-1* nonsense mutation (Krause et al., 1994), the *C. elegans unc-119* phenotype

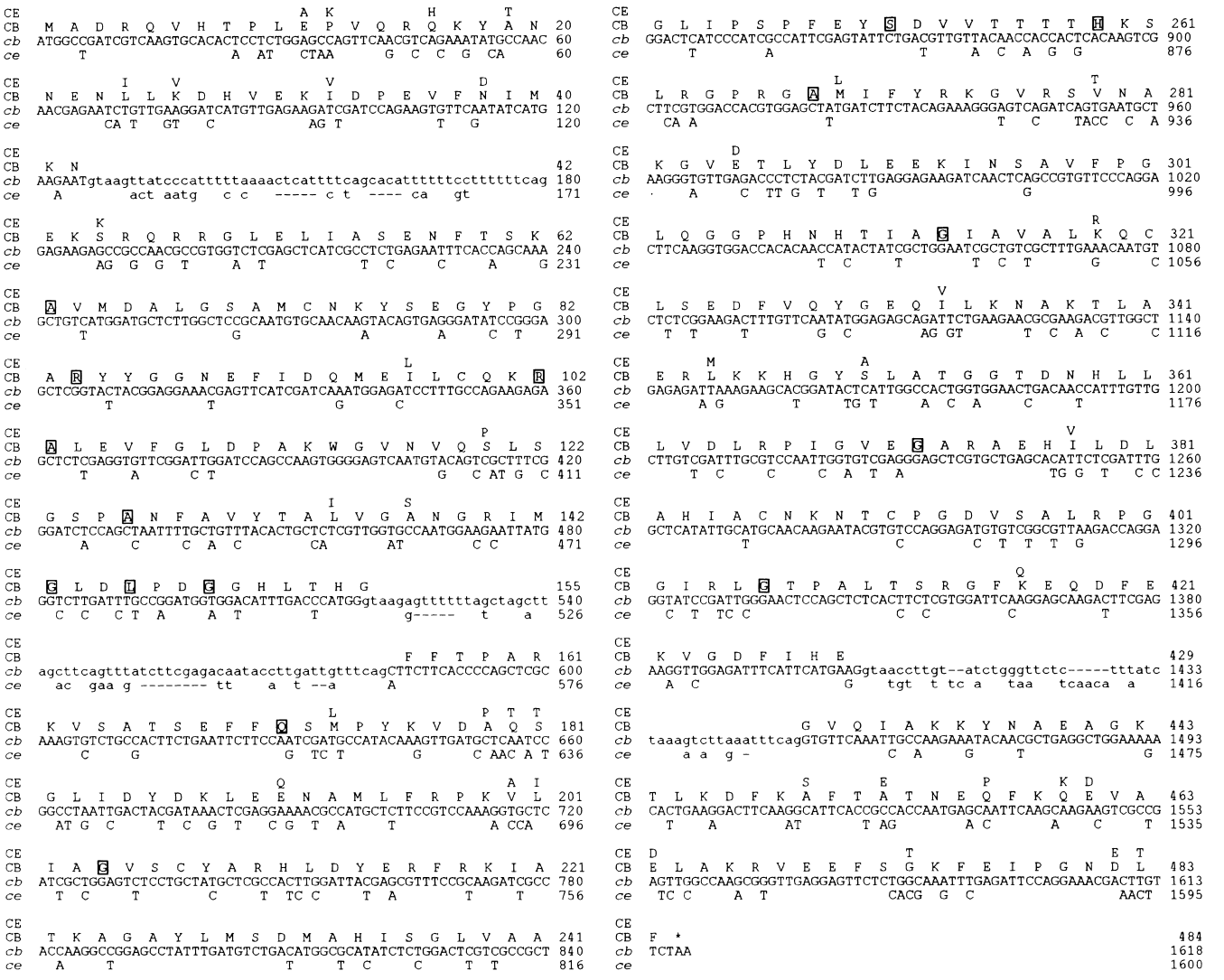

Fig. 2. Nucleotide sequence of the *C. briggsae* SHMT genomic region and derived amino acid sequence. The entire genomic region (from start codon to stop codon) for *C. briggsae* (cb) and the conceptual translation product (CB) are shown. The *C. elegans* nucleotide (ce) and amino acid (CE) sequences are listed only where there is a difference. The amino acids mutated in the *C. elegans mel-32* alleles (Vatcher et al., 1998) are boxed in the corresponding briggsae sequence. Intronic regions are in lower case, and gaps are indicated by a '–'.

(Maduro and Pilgrim, 1996), and *C. elegans tra-1* null mutations (de Bono and Hodgkin, 1996). The rescue of the *C. elegans mel-32*(*s2518*) mutation with *C. briggsae* genomic DNA shows that the CbMEL-32(SHMT) protein is functional in *C. elegans* and that it is properly transcribed and translated. It also suggests that the promoter and enhancer regions share enough homology to regulate the proper temporal and developmental control of *mel-32*(SHMT). The fact that the *C. briggsae* protein is functional in *C. elegans* is not surprising, as they are 97% similar, with only 14 non-conservative changes over 484 amino acids. All of the essential *C. elegans* amino acids previously identified (Vatcher et al., 1998) are conserved (Fig. 2).

When the *C. elegans* and *C. briggsae* MEL-32(SHMT) proteins are compared to other SHMT isoforms, long stretches of high homology are revealed. An important exception occurs at residue 138 of the *C. elegans* and *C. briggsae* proteins (Fig. 3). The properties of three site-directed mutants of sheep liver cytosolic SHMT were previously reported (Jagath et al., 1997). Conserved histidines 134, 147, and 150, which correspond to residues 138, 151, and 154 in MEL-32(SHMT), were mutated into asparagines. The H147N mutant had only 3% bound PLP, whereas the H150N mutant had less than 2% of wild-type activity. These residues are 100% conserved (Fig. 3), and His-147 is probably involved in cofactor binding, whereas His-150 may be the base that abstracts the α-proton of the substrate. Jagath et al. (1997) state that His-134 is conserved in all tetrameric eukaryotic SHMTs but is replaced by a glycine in prokaryotic SHMTs that exist as dimers. The sheep H134N mutant did not form tetramers, had a lowered enzymatic activity and PLP

```
C. briggsae MEL-32 GANGRIMGLDLPDGGHLTHGFFTP 159
C. elegans MEL-32  GSNGRIMGLDLPDGGHLTHGFFTP 159
human C-SHMT       EPHGRIMGLDLPDGGHLTHGFMTD 156
yeast C-SHMT       KPHERLMGLYLPDGGHLSHGYATE 147
human M-SHMT       QPHDRIMGLDLPDGGHLTHGYMSD 158
yeast M-SHMT       NVGERLMGLDLPDGGHLSHGYQLK 163
E. coli SHMT       EPGDTVLGMNLAHGGHLTHGSPVN 134
rabbit C-SHMT      EPHGRIMGLDLPDGGHLTHGFMTD 156
rabbit M-SHMT      QPHDRIMGLDLPDGGHLTHGYMSD 179
sheep C-SHMT       EPHGRIMGLDLPDGGHLTHGFMTD 156
                     ..*. *  ****.**
```

Fig. 3. Comparison of SHMTs. The numbers indicate the positions in the corresponding proteins. C represents the cytosolic isoform, M represents the mitochondrial isoform, a '*' represents amino acids identical in all the proteins, and a '.' represents a conservative change. The residues of interest are highlighted (see Section 3.2 for details).

binding, and had an altered quaternary structure. This suggests that His-134 is involved in dimer–dimer interactions and may have a role in the maintenance of the enzyme's tetrameric structure. The MEL-32(SHMT) proteins have an asparagine at this position (Fig. 3) suggesting that the *Caernorhabditis* enzymes exist as dimers, that they use the asparagine to form a hydrogen bond to stabilize the tetramer, or that they use an alternate residue to form dimer–dimer interaction. It is worth noting that the yeast mitochondrial isoform has a glycine at this position, the same as the prokaryotic isoforms, suggesting that this isoform may also exist as a dimer.

Table 1
Summary of conserved non-coding regions in the *C. elegans* and *C. briggsae mel-32*(SHMT) genomic DNA regions

Region	Location (relative to ATG)		Matches	Percentage homology
	C. elegans	*C. briggsae*		
A	−897	−1267	13/15	87
B	−795	−1192	11/11	100
C	−751	−1180	35/44	80
D	−694	−1122	16/20	80
E	−661	−1096	33/44	75
F	+1660	+1672	21/23	91

It is perhaps surprising that the Cb*mel-32*(SHMT) gene is properly transcribed in *C. elegans*. The entire upstream and downstream regions share only 43% homology overall, similar to the value obtained when the two sequences are randomized before alignment. However, the dot matrix analysis (Fig. 1) reveals that there are stretches of highly conserved sequence. These conserved islands probably represent sequence elements that bind regulatory factors. The six regions with the highest homologies range in size from 11 to 44 nucleotides and share between 75 and 100% identity. The regions are summarized in Table 1 and Fig. 4B. Fig. 4 reveals that the relative order of regions A–F is conserved but that the relative spacing between them is different.

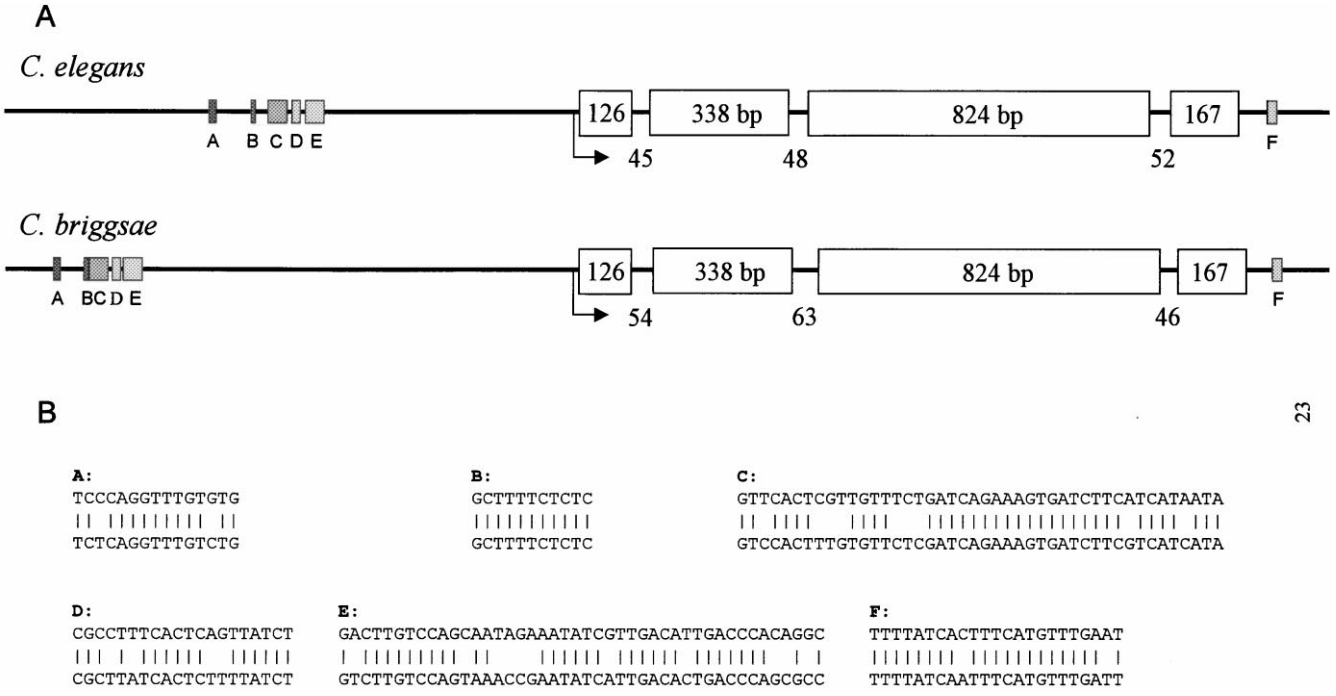

Fig. 4. (**A**) *mel-32* exon, intron, and non-coding conserved region spacing. Exons are represented by open boxes, and non-coding conserved regions (A–F) are represented by shaded boxes. Exon sizes are listed inside the boxes, and intron sizes are indicated below the line. The presumed SL1 splice sites are indicted by arrows. (**B**) Conserved non-coding sequences. The *C. elegans* sequence is on the top, and the *C. briggsae* sequence is on the bottom. Identical residues are indicated by a '|'. A–F indicate the regions in Figs. 1 and 4A, and Table 1.

The upstream cluster (regions A–E) is spread over a slightly larger area in *C. elegans* than is *C. briggsae* (237 nucleotides versus 172 nulceotides). Fig. 4 also reveals that the *C. briggsae* cluster is approximately 400 bp further upstream than the *C. elegans* cluster. When *C. elegans* and *C. briggsae* genes are compared, there are usually short stretches of highly conserved non-coding sequences both upstream and downstream of the exons. This conservation is probably significant since *C. elegans* and *C. briggsae* are sufficiently diverged to remove any non-selected similarity (Prasad and Baillie, 1989). Conserved regions upstream of genes are believed to control developmental and tissue-specific transcription. The *hlh-1* (Krause et al., 1994), *ges-1* (Kennedy et al., 1993), *dpy-7* (Gilleard et al., 1997b), *unc-119* (Maduro and Pilgrim, 1996), *mec-3* (Xue et al., 1992), and *ceh-24* (Harfe and Fire, 1998) genes all have upstream regions conserved in *C. elegans* and *C. briggsae*. These genes have essentially identical activation patterns when transformed into the heterologous species. Conserved regions downstream of genes are believed to regulate polyadenylation and translation, as has been shown for the *tra-2* gene (Jan et al., 1997). Future deletion analysis of the *mel-32* promoter should reveal which, if any, of the conserved regions are important in the transcriptional control of this essential gene.

3.3. Conclusions

In summary, we have identified a *C. briggsae* serine hydroxymethyltransferase isoform that is 92% identical and 97% similar to the *C. elegans* MEL-32(SHMT) protein. All of the previously identified essential residues in CeMEL-32(SHMT) (Vatcher et al., 1998) are conserved in CbMEL-32(SHMT). In addition, there are conserved regions upstream and downstream of the *mel-32*(SHMT) gene. Interspecies phenotypic rescue of the *C. elegans* Mel-32 phenotype with *C. briggsae* genomic DNA implies that these conserved regions regulate the transcription and translation of the *mel-32* gene. The *Caenorhabditis* MEL-32(SHMT) protein differs from other characterized SHMTs at position 138. The *C. elegans* and *C. briggsae* proteins have an asparagine at this position, whereas other isoforms have either a glycine or a histidine. This may indicate that, in nematodes, SHMT forms a dimer.

Acknowledgements

We thank Jacquie Schein for mapping data and for providing the *C. briggsae* fosmid clones. This work was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) operating grant to DLB and a Natural Sciences and Engineering Research Council of Canada (NSERC) postgraduate scholarship (PGSB) to G.P.V.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Brenner, S., 1974. The genetics of *Caenorhabditis elegans*. Genetics 77, 71–94.

de Bono, M., Hodgkin, J., 1996. Evolution of sex determination in *Caenorhabditis*: unusually high divergence of *tra-1* and its functional consequences. Genetics 144, 587–595.

Emmons, S., Klass, M., Hirsh, D., 1979. Analysis of the constancy of DNA sequences during development and the evolution of the nematode *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA 76, 1333–1337.

Garrow, T.A., Brenner, A.A., Whitehead, V.M., Chen, X.N., Duncan, R.G., Korenberg, J.R., Shane, B., 1993. Cloning of human cDNAs encoding mitochondrial and cytosolic serine hydroxymethyltransferases and chromosomal localization. J. Biol. Chem. 268, 11910–11916.

Gilleard, J.S., Henderson, D.K., Ulla, N., 1997a. Conservation of the *Caenorhabditis elegans* cuticle collagen gene *col-12* in *Caenorhabditis briggsae*. Gene 193, 181–186.

Gilleard, J.S., Barry, J.D., Johnstone, I.L., 1997b. *cis* regulatory requirements for hypodermal cell-specific expression of the *Caenorhabditis elegans* cuticle collagen gene *dpy-7*. Mol. Cell. Biol. 17, 2301–2311.

Harfe, B.D., Fire, A., 1998. Muscle and nerve-specific regulation of a novel NK-2 class homeodomain factor in *Caenorhabditis elegans*. Development 125, 421–429.

Heine, U., Blumenthal, T., 1986. Characterization of regions of the *Caenorhabditis elegans* X chromosome containing vitellogenin genes. J. Mol. Biol. 188, 301–312.

Hengartner, M.O., Horvitz, H.R., 1994. *C. elegans* cell survival gene *ced-9* encodes a functional homolog of the mammalian proto-oncogene *bcl-2*. Cell 76, 665–676.

Heschl, M.F., Baillie, D.L., 1990. Functional elements and domains inferred from sequence comparisons of a heat shock gene in two nematodes. J. Mol. Evol. 31, 3–9.

Jagath, J.R., Sharma, B., Rao, N.A., Savithri, H.S., 1997. The role of His-134, -147, and -150 residues in subunit assembly, cofactor binding, and catalysis of sheep liver cytosolic serine hydroxymethyltransferase. J. Biol. Chem. 272, 24355–24362.

Jan, E., Yoon, J.W., Walterhouse, D., Iannaccone, P., Goodwin, E.B., 1997. Conservation of the *C. elegans tra-2* 3′UTR translational control. EMBO J. 16, 6301–6313.

Janke, D.L., Schein, J.E., Ha, T., Franz, N.W., Oneil, N.J., Vatcher, G.P., Stewart, H.I., Kuervers, L.M., Baillie, D.L., Rose, A.M., 1997. Interpreting a sequenced genome: toward a cosmid transgenic library of *Caenorhabditis elegans*. Genome Res. 7, 974–985.

Kastanos, E.K., Woldman, Y.Y., Appling, D.R., 1997. Role of mitochondrial and cytoplasmic serine hydroxymethyltransferase isozymes in de novo purine synthesis in *Saccharomyces cerevisiae*. Biochemistry 36, 14956–14964.

Kennedy, B.P., Aamodt, E.J., Allen, F.L., Chung, M.A., Heschl, M.F.P., McGhee, J.D., 1993. The gut esterase gene (*ges-1*) from the nematodes *Caenorhabditis elegans* and *Caenorhabditis briggsae*. J. Mol. Biol. 229, 890–908.

Kloek, A.P., McCarter, J.P., Setterquist, R.A., Schedl, T., Goldberg, D.E., 1996. *Caenorhabditis* globin genes: rapid intronic divergence contrasts with conservation of silent exonic sites. J. Mol. Evol. 43, 101–108.

Kramer, J.M., French, R.P., Park, E.C., Johnson, J.J., 1990. The

Caenorhabditis elegans rol-6 gene, which interacts with the *sqt-1* collagen gene to determine organismal morphology, encodes a collagen. Mol. Cell Biol. 10, 2081–2089.

Krause, M., Harrison, S.W., Xu, S.Q., Chen, L., Fire, A., 1994. Elements regulating cell- and stage-specific expression of the *C. elegans* MyoD family homolog *hlh-1*. Dev. Biol. 166, 133–148.

Kuwabara, P.E., Shah, S., 1994. Cloning by synteny: identifying *C. briggsae* homologous of *C. elegans* genes. Nucleic Acids Res. 22, 4414–4418.

Kuwabara, P.E., 1996. Interspecies comparison reveals evolution of control regions in the nematode sex-determining gene *tra-2*. Genetics 144, 597–607.

Lai, C.C., Hong, K., Kinnell, M., Chalfie, M., Driscoll, M., 1996. Sequence and transmembrane topology of MEC-4, an ion channel subunit required for mechanotransduction in *Caenorhabditis elegans*. J. Cell Biol. 133, 1071–1081.

Lee, Y.H., Huang, X.Y., Hirsh, D., Fox, G.E., Hecht, R.M., 1992. Conservation of gene organization and *trans*-splicing in the glyceraldehyde-3-phosphate dehydrogenase-encoding genes of *Caenorhabditis briggsae*. Gene 121, 227–235.

McNeil, J.B., McIntosh, E.M., Taylor, B.V., Zhang, F.R., Tang, S., Bognar, A.L., 1994. Cloning and molecular characterization of three genes, including two genes encoding serine hydroxymethyltransferases, whose inactivation is required to render yeast auxotrophic for glycine. J. Biol. Chem. 269, 9155–9165.

Maduro, M., Pilgrim, D., 1996. Conservation of function and expression of *unc-119* from two *Caenorhabditis* species despite divergence of non-coding DNA. Gene 183, 77–85.

Mello, C.C., Kramer, J.M., Stinchcomb, D., Ambros, V., 1991. Efficient gene transfer in *C. elegans* after microinjection of DNA into germline cytoplasm: extrachromosomal maintenance and integration of transforming sequences. EMBO J. 10, 3959–3970.

Narkewicz, M.R., Sauls, S.D., Tjoa, S.S., Teng, C., Fennessey, P.V., 1996. Evidence for intracellular partitioning of serine and glycine metabolism in Chinese hamster ovary cells. Biochem. J. 313, 991–996.

Pearson, W.R., Lipman, D.J., 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85, 2444–2448.

Prasad, S.S., Baillie, D.L., 1989. Evolutionarily conserved coding sequences in the *dpy-20–unc-22* region of *Caenorhabditis elegans*. Genomics 5, 185–198.

Rose, A.M., Snutch, T.P., 1984. Isolation of the closed circular form of the transposable element Tc1 in *Caenorhabditis elegans*. Nature 311, 485–486.

Rychlik, W., Rhoads, R.E., 1989. A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA. Nucleic Acids Res. 17, 8543–8551.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: a Laboratory Manual. 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schirch, L., 1982. Serine hydroxymethyltransferase. Adv. Enzymol. 53, 83–112.

Sonnhammer, E.L., Durbin, R., 1995. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene 167, GC1–10.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Usha, R., Savithri, H.S., Rao, N.A., 1994. The primary structure of sheep liver cytosolic serine hydroxymethyltransferase and an analysis of the evolutionary relationships among serine hydroxymethyltransferases. Biochim. Biophys. Acta 1204, 75–83.

Vatcher, G.P., Thacker, C.M., Kaletta, T., Schnabel, H., Schnabel, R., Baillie, D.L., 1998. Serine hydroxymethyltransferase is maternally essential in *Caenorhabditis elegans*. J. Biol. Chem. 273, 6066–6073.

Xue, D., Finney, M., Ruvkun, G., Chalfie, M., 1992. Regulation of the *mec-3* gene by the *C. elegans* homeoproteins UNC-86 and MEC-3. EMBO J. 11, 4969–4979.