Methylation patterns in the human muscle-specific enolase gene (*ENO3*)

Mina PESHAVARIA and Ian N. M. DAY*

University Department of Clinical Biochemistry, Level D, South Laboratory and Pathology Block, Southampton General Hospital, Tremona Road, Southampton SO9 4XY, U.K.

The methylation status in the human-muscle enolase gene (*ENO3*) was assayed. Previous sequence data and *MspI* cleavage sites indicate the presence of a 5' CpG-rich island of at least 4 kb: none of 22 characterized *MspI* CCGG sites is methylated in any of muscle, sperm or brain DNA. However a complex pattern of complete and partial methylation of *MspI* sites that is different between tissues is observed within the *ENO3* gene: events at one

site may be specific to muscle DNA. The absence of methylation in the promoter region of the *ENO3* gene makes it unlikely that methylation plays a causal role either in transcriptional events or in the divergence of enolase-isogene regulation. The role of tissue-specific methylation events within *ENO3* remains to be determined.

INTRODUCTION

Enolase (EC 4.2.1.11) catalyses the glycolytic step interconverting 2-phosphoglycerate and phosphoenolpyruvate. Although this step appears to be at 'equilibrium' *in vivo* and is not a classical site of metabolic pathway regulation [1], tissue-specific isoforms are found and these display minor kinetic differences [2,3]. Additionally the absolute cellular level of enolase activity is 20-fold higher in muscle (per mg of total soluble protein) than in other tissues [1], possibly reflecting the heavy metabolic demands of skeletal muscle. The possible role of a wide range of isoenzyme systems and their cognate genes and their differential regulation have been extensively investigated. Instances of mechanistic value to the organism have been identified at the substrate, kinetic, macromolecular, translational and transcriptional levels and these roles become clear as a complete description of the system is obtained (for a recent overview, see [4]). Three major tissue-specific isoforms of enolase have been identified in mammals and birds: γ or neuron-specific enolase (NSE), expressed primarily in neurons; β or muscle-specific enolase (MSE), expressed in striated muscle; and α or non-neuronal enolase (NNE) expressed in fetal and other adult cell types. A variety of rat and human NSE, NNE and MSE cDNA and genomic clones have been isolated ([5–13] and references therein), including most recently genomic clones representing human NSE [11] and MSE [12].

It has been established that there is transcriptional switching of the rat NSE gene during neuronal development [10] and earlier *in vitro* translation studies of chicken muscle have suggested a comparable event for the MSE gene during muscle development [14]. Extensive general information has emerged about myogenesis and muscle-gene expression, including the identification of master genes such as myoblast determination gene-1 (*Myo D1*) that can trigger undifferentiated cells to adopt a muscle phenotype [15]. A common pattern of regulation of the genes responsible for the phenotype involves transcriptional regulation by specific protein factors binding to specific motifs near the transcriptional start site [16–21]. In addition to these features of differential gene regulation, patterns of gene methylation can be correlated (either as a cause or as a

consequence) with gene activity both in cell lines ([22] and references therein) and *in vivo* in tissues including muscle ([23] and references therein). This work set out to apply cloned human-genomic *ENO3* to the analysis of methylation patterns in *ENO3* in different tissues, using methylation-sensitive and insensitive restriction enzymes.

MATERIALS AND METHODS

Reagents

Radiochemicals and Hybond-N were from Amersham International (Amersham, Bucks., U.K.). Restriction enzymes and other DNA-modifying enzymes were either from Northumbria Biologicals Ltd., Cramlington, Northumbria, U.K. or from Anglian Biotech Ltd., Colchester, Essex, U.K. Chemicals and other reagents were from Sigma Chemical Co., Poole, Dorset, U.K. Cell culture reagents and plastics were from Gibco-BRL, Paisley, Scotland.

Preparation of probes and plasmid constructs

Restriction fragments representing regions of human *ENO3* were isolated from the clone λ HGME, as described previously [12], using standard laboratory methods [24–27] for isolation, labelling to specific activities of 10^8 – 10^9 d.p.m./ μ g and recloning.

Analysis of *ENO3* methylation patterns

Genomic DNA was isolated [27] from post-mortem (average 24 h) samples of adult human liver, muscle and semen of healthy males [28]. Preliminary digests of genomic DNA with the enzyme(s) representing the *in vivo* ends of the restriction fragment that was to be used as the probe were followed by digestion using restriction enzymes that are differentially sensitive to methylation of their target sequences. Typically the probes were *PstI* fragments and the secondary digests were with the methylation-sensitive enzyme *HpaII* as against its isoschizomer, *MspI*, which is not sensitive to methylation at their common recognition sequence, 5'-CCGG [26]. A sample of 10 μ g of digested genomic DNA was electrophoresed, either in standard agarose gels or in denaturing acrylamide gels, the latter to resolve size fragments

Figure 1 Map of the restriction sites, probe-region nomenclature (regions A, B, C, D, E, G and H) and positions of the first seven exons (I–VII) of the human *ENO3* gene

MspI sites are marked by a vertical bar and circle below the line of the gene. Coding exons are shown as solid boxes (■), non-coding exon (I) as a hatched box (▨). Gaps indicate sequence that has not been determined (ND). The methylation status of the *MspI* sites is coded as follows: open circles (○), unmethylated in all tissues examined; left-half-filled circles (◐), partially methylated in tissues examined; right-half-filled circles (◑), methylated in sperm DNA but not in brain nor in muscle DNA; lower-half-filled circles (◒), fully methylated in sperm and in brain DNA but partially unmethylated in muscle DNA.

Table 1 Summary of *ENO3* regions and probes used showing *MspI* sites and methylation status

See Figure 1 and the Results section for details.

Region/probe	Size (kb)	Region represented	Number of <i>MspI</i> sites	Methylation status of <i>MspI</i> sites
G	2.35	'Distal' upstream region	At least 11	Constitutively unmethylated
D	0.80	'Distal' upstream region	At least 7	Constitutively unmethylated
B	1.62	'Proximal' (0.7 kb) upstream region, TATA box, exon I, intron I	4	Constitutively unmethylated
C	1.15	Exons II, III, IV, introns 2, 3	4	At least partially methylated in tissues examined
E	0.53	Exon V, introns 4, 5	2	Methylated in sperm DNA, but not in brain or in muscle DNA
H/A	1.66	Exons VI, VII, introns 6, 7	1	Fully methylated in sperm and in brain DNA, partially unmethylated in muscle DNA

from 50–500 bp. Electrobloods were prepared, essentially as described by Church and Gilbert [29,30]. Identification of the smallest *MspI* fragments (down to 50 bp) required careful attention to the specific activity of the probe, the blocking and hybridization conditions and the autoradiography procedure, as described previously [29,30]. All blots were prepared using Hybond-N and were probed at 'high' stringency [$0.1 \times$ standard saline citrate (SSC); $1 \times$ SSC is 0.15 M NaCl and 0.015 M sodium citrate)/ 0.1% w/v SDS at 65°C].

RESULTS

The nomenclature of genomic regions used as probes and their relationships to *MspI* sites identified within the gene sequence [12] are illustrated in Figure 1. In outline, region B spans the near upstream region (0.7 kb), exon I and intron 1; regions G and D represent a further upstream region (3.15 kb) and regions C, E, A and H represent the gene interior. Further detail is presented in Table 1.

In Figures 2 and 3, probes B, G and D detect fragments in *PstI*–*MspI* double digests that represent a total of at least 20 *MspI* sites (see Table 1). All of these sites are cleaved as well by *PstI*–*HpaII* in the three tissues examined, indicating a complete absence of methylation at CCGG sites within a 3.5 kb region

upstream of *ENO3*. The *MspI* sites in B are all clustered within a CpG-rich region 5' to the TATA box, none falling within exon I or intron 1, therefore all of the unmethylated sites identified here fall upstream to the gene.

Figure 4 presents a more complex pattern, with differences observed between *MspI* and *HpaII* digests of the same tissue and between *HpaII* digests of different tissues. The differences are as follows. (1) *HpaII* does not cleave the targets for probes C, E or H in sperm DNA, indicating that the respective 4, 2 and 1 *MspI* sites within these targets in sperm DNA are methylated. (2) The single *MspI* site in H is only cleaved by *HpaII* in muscle, and is not cleaved in brain or in sperm DNA. The cleavage was estimated (by eye) to be 30–50% complete. (3) The two *MspI* sites in E are not cleaved at all by *HpaII* in purified sperm DNA, but are fully cleaved by *HpaII* in brain and muscle DNA. (4) Region C is at least partially uncleaved by *HpaII* both in muscle and in brain DNA, indicating that all four *MspI* sites in region C in each of these tissues are at least partially methylated. (5) The signal strength (per mol) is substantially weaker for target C than for H, perhaps reflecting the effect of pre-absorption of probe C (which contains *Alu* sequences) with total genomic DNA. Since the small subfragments of C cannot be identified reliably in brain and muscle *MspI* tracks, the exact methylation status of target C in these tissues remains uncertain.

Figure 2 Analysis of the methylation of *ENO3*: electrophoretograms of digested human DNA, probed using *ENO3* region B

Electrophoretograms of restriction-enzyme digests of human DNA after electrophoresis through 5% acrylamide [7 M urea, 0.5 × TBE (final concentration 0.044 M Tris, 0.044 M borate, 1 mM EDTA)], probed using *ENO3* region B, a *PstI*–*PstI* fragment (see Figure 1). The tissue of origin of the DNA is indicated (brain, muscle or sperm) and all tracks represent a double digest, first with *PstI* and then either with *MspI* (M) or with *HpaII* (H). Bands were identified from the known sequence [12] and hence the *MspI* restriction map of region B, the adjacent electrophoresis of an *MspI* digest of purified DNA representing region B and from end-labelled *MspI*-digested pUC18 size markers (sizes indicated). Both 5-day (a) and overnight (b) autoradiographs are presented so that both larger (e.g., 1282 nt) and smaller (e.g., 65 nt) fragments can be visualized. One faint band, present only in the brain tracks, reflects residual incomplete *PstI* digestion of DNA from this tissue.

Figure 4 Analysis of the methylation of *ENO3*: Southern blots of digested human DNA with probes representing the region exon II–VII

Southern-blot analysis of the interior of the *ENO3* gene (exons II–VII) using a mixture of probes (C, a *PstI*–*PstI* fragment; E, a *PstI*–*PstI* fragment; and A, a *PstI*–*XbaI* fragment; refer to Figure 1). Standard conditions for non-denaturing agarose gel electrophoresis and Southern blotting were used. The tissues of origin of the total human DNA are indicated: P marks tracks that were digested only with *PstI*; M and H, tracks digested with *MspI* or *HpaII* respectively, both after an initial *PstI* digestion. A minor persistent difference in the *PstI* fragments between tissues was noted that probably represents residual incomplete digestion in the brain and muscle tracks. The complex pattern of differences between *MspI* and *HpaII* tracks within and between tissues is discussed in detail in the text. The sizes indicated (right-hand side) represent *HindIII*-digested bacteriophage λ . The sizes and designation (Figure 1) of the *ENO3* *PstI* fragments are marked as well (left-hand side).

Figure 3 Analysis of the methylation of *ENO3*: electrophoretograms of digested human DNA probed using (a) *ENO3* region G and (b) *ENO3* region D

Conditions and size markers are as in Figure 2. All the tracks represent double digests of human DNA with *PstI* followed by either *MspI* (M) or *HpaII* (H). The source tissue is as indicated. The fragments visualized have not been assigned at the sequence level. Note that no fragments differ between the *MspI* and *HpaII* tracks for a given tissue; slight differences between the tissues represent slight residual differences in *PstI* digestion between the tissues.

DISCUSSION

In sperm, brain and muscle genomic DNA, no methylated 5'-CCGG sites were identified at the 5' end of the *ENO3* gene or upstream of it: 22 sites were examined within a 3.5 kb region (probes B, G and D; Figures 2 and 3). Genes in mammals are

typically marked by a 5' CpG island that is relatively spared from the CpG depletion that is typical of the rest of the genome [31–33]. It is thought that the hypermutability of methylated CpG sites accounts for the general depletion of CpG that is found in mammalian DNA [34]. The mechanisms that maintain CpG islands are obscure, but a number of 'housekeeping' genes have been shown to possess constitutively unmethylated CpG sites around their 5' ends [35]. It has been noted as well that CpG/GpC ratios in 5' islands tends to be higher in 'housekeeping' genes than in tissue-specific genes [23]. We have presented summary data for human and rat enolase genes previously that show that the enolase gene family reinforces this observation ([36]; Table 1 in [37]). In a 919-bp region upstream of the transcriptional start site, the CpG/GpC ratios are as follows: for the human NNE gene, 0.95; for the human NSE gene, 0.54; for the human MSE gene, 0.54; for the rat NSE gene, 0.60; and for the rat MSE gene, 0.47 (the average ratio in mammalian DNA is approx. 0.25). Thus NNE, which is expressed in fetal tissues and in most adult tissues, is transcribed from a gene with essentially no CpG depletion in its 5' CpG island, whereas the tissue-specific (neuronal and muscle) isogenes are approximately two-fold depleted in CpG in the corresponding region. In this work we have presented evidence that the human MSE gene is constitutively unmethylated at the CpG sites that were studied (that is, the 5'-CCGG sites) within its 5' CpG island. Unless transient methylation takes place at some stage within the germ line, it is not clear how CpG depletion by the deamination of methyl-CpG could account for the divergence in the CpG/GpC ratios that is noted above for the human MSE gene as against the human NNE gene. However there may be other bulk or specific properties of these regions that influence their nucleotide and sequence content in ways that are different from the rest of the

genome. Distinctive chromatin superstructure [38] and properties of methylases [39] and methyl-CpG-binding proteins [40] are three general factors that may be involved in the process. Specific transcription factors and enhancer-binding proteins might influence methylation as well. Conversely the apparently constitutive unmethylation of the *ENO3* 5' region makes it unlikely that methylation could exert any primary role in regulating *ENO3* transcription *in vivo* (and our preliminary data suggest that transcriptional regulation does occur within this region).

Methylation events within the *ENO3* gene are more complex, displaying a pattern of complete and partial methylation that is different between sites and between tissues (Table 1). One clear result is seen, however, relating to a single *MspI* site in region H, within intron 6. This site (see Results section, point 2) is fully methylated in sperm and in brain DNA, but 30–50% of sites are unmethylated in muscle DNA. Since *ENO3* is expressed in all human muscle fibres, albeit to different levels in type I and type II fibres [41,42], but is not expressed at all in sperm or in brain, this site represents the most direct correlation identified between methylation events and transcriptional events in *ENO3*. Its methylation status in relation to development, fibre type and local chromatin structure would be of interest.

Thus the *ENO3* gene has some methylation properties that are typical of other genes, but has specific features as well that are not explained by current models and that may offer further insight into the evolution, structural disposition and expression of the mammalian enolase genes.

M. P. was the recipient of a Ph.D. Studentship from the Sir Halley Stewart Trust. Work in this laboratory is supported by the British Epilepsy Research Foundation and the Medical Research Council. We thank Mrs D. Brown for typing the manuscript and Professor Thompson for departmental facilities and space.

REFERENCES

- Newsholme, E. A. and Start, C. (1973) in *Regulation in Metabolism*, pp. 88–137, John Wiley & Sons, London, New York, Sydney and Toronto
- Pearce, J. M., Edwards, Y. H. and Harris, H. (1975) *Ann. Hum. Genet.* **39**, 263–280
- Marangos, P. J., Parma, A. M. and Goodwin, F. K. (1978) *J. Neurochem.* **31**, 727–732
- Doonan, S., ed. (1990) *Biochem. Soc. Trans.* **18**, 167–196
- Sakimura, K., Kushiya, E., Obinata, M., Odani, S. and Takahashi, Y. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7453–7457
- Sakimura, K., Kushiya, E., Obinata, M. and Takahashi, Y. (1985) *Nucleic Acids Res.* **13**, 4365–4378
- Forss-Petter, S., Danielson, P. and Sutcliffe, J. G. (1986) *J. Neurosci. Res.* **16**, 141–156
- Giallongo, A., Feo, S., Moore, R., Croce, C. M. and Showe, L. C. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6741–6745
- McAleese, S. M., Dunbar, B., Fothergill, J. E., Hinks, L. J. and Day, I. N. M. (1988) *Eur. J. Biochem.* **178**, 413–417
- Sakimura, K., Kushiya, E., Takahashi, Y. and Suzuki, Y. (1987) *Gene* **60**, 103–113
- Oliva, D., Cali, L., Feo, S. and Giallongo, A. (1991) *Genomics* **10**, 157–165
- Peshavaria, M. and Day, I. N. M. (1991) *Biochem. J.* **275**, 427–433
- Craig, S. P., Day, I. N. M., Thompson, R. J. and Craig, I. W. (1990) *Cytogenet. Cell Genet.* **54**, 71–73
- Tanaka, M., Sugisaki, K. and Nakashima, K. (1985) *Biochem. Biophys. Res. Commun.* **133**, 868–872
- Tapscott, S. J., Davis, R. L., Thayer, M. J., Cheng, P.-F., Weintraub, H. and Lassar, A. B. (1988) *Science* **242**, 405–411
- Wingender, E. (1988) *Nucleic Acids Res.* **16**, 1879–1901
- Gustafson, T. A. and Kedes, L. (1989) *Mol. Cell. Biol.* **9**, 3269–3283
- Ng, S.-Y., Gunning, P., Liu, S.-H., Leavitt, J. and Kedes, L. (1989) *Nucleic Acids Res.* **17**, 601–605
- Mar, J. H. and Ordahl, C. P. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6404–6408
- Devlin, B. H., Wefald, F. C., Kraus, W. E., Bernard, T. S. and Williams, R. S. (1989) *J. Biol. Chem.* **264**, 13896–13901
- Fink, J. S., Verhave, M., Kasper, S., Tsukada, T., Mandel, G. and Goodman, R. H. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6662–6666
- Cedar, H. (1988) *Cell* **53**, 3–4
- Edwards, Y. H. (1990) *Philos. Trans. R. Soc. London B* **326**, 207–215
- M13 Handbook, Amersham International, Amersham, U.K.
- Hattori, M. and Sakaki, Y. (1986) *Anal. Biochem.* **152**, 232–238
- Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Berger, S. L. and Kimmel, A. R., ed. (1987) *Methods Enzymol.* **152**
- Gill, P., Jeffreys, A. J. and Werrett, D. J. (1985) *Nature (London)* **318**, 577–579
- Church, G. M. and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1991–1995
- Edwards, Y. H., Charlton, J. and Brownson, C. (1988) *Gene* **71**, 473–481
- Cooper, D. N., Taggart, M. H. and Bird, A. P. (1983) *Nucleic Acids Res.* **11**, 647–658
- Bird, A. P. (1992) *Cell* **70**, 5–8
- Bird, A. P. (1986) *Nature (London)* **321**, 209–213
- Bird, A. P. (1987) *Trends Genet.* **3**, 342–347
- Gardiner-Garden, M. and Frommer, M. (1987) *J. Mol. Biol.* **196**, 261–282
- Peshavaria, M. (1991) Ph.D. Thesis, University of Southampton
- Day, I. N. M. (1992) *Biochem. Soc. Trans.* **20**, 637–642
- Tazi, J. and Bird, A. (1990) *Cell* **60**, 909–920
- Adams, R. L. P., Bryans, M., Rinaldi, A., Smart, A. and Yesufu, H. M. I. (1990) *Philos. Trans. R. Soc. London B* **326**, 189–198
- Meehan, R. R., Lewis, J. D., McKay, S., Kleiner, E. L. and Bird, A. P. (1989) *Cell* **58**, 499–507
- Matsushita, H., Yamada, S., Satoh, T., Kato, K. and Adachi, M. (1986) *Exp. Neurol.* **93**, 84–91
- Ibi, T., Haimoto, H., Nagura, H., Sahashi, K., Kato, K., Mokuno, K., Sugimura, K. and Matsuoka, Y. (1985) *Acta Neuropathol.* **68**, 74–76