Open Access Bioinformatics

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Bioinformatic studies of vertebrate enolases: multifunctional genes and proteins

This article was published in the following Dove Press journal:
Open Access Bioinformatics
8 February 2011
Number of times this article has been viewed

Roger S Holmes

School of Biomolecular and Physical Sciences, Griffith University, Nathan, QLD, Australia

Abstract: Enolase (*ENO*) genes and proteins (ENO; EC 4.2.1.11) serve multiple functions in the body, including catalyzing 2-phospho-D-glycerate hydro-lyase activity in glycolysis, assisting hypoxia tolerance, tumor suppression, plasminogen and DNA binding, and acting as a lens crystallin. Comparative ENO amino acid sequences and structures and *ENO* gene locations were examined using data from several vertebrate genome projects. Vertebrate *ENO1*, *ENO2*, and *ENO3* genes usually contained 11 coding exons, while *ENO4* (encoding an ENO-like protein, ENOLL) usually contained 14 coding exons. Vertebrate *ENOF1* (or *ENO5*) genes encode an antisense RNA, which may regulate mitochondrial thymidylate synthase activity that contained 12–15 coding exons. Vertebrate ENO1, ENO2, and ENO3 sequences shared 78%–98% identities but only 19%–24% with ENO4 and >10% predicted sequence identities with vertebrate ENOF1. Sequence alignments, key amino acid residues, and conserved predicted secondary and tertiary structures were examined, including active site residues (absent in ENO4 and ENOF1) and sites for Mg^{2+} and plasminogen binding and for acetylation and phosphorylation. The predicted ENO4 structure contained three N-terminal α-helices, two β-sheets, a poly-proline segment, and an extended C-terminal sequence in addition to the typical α/β barrel structure reported for ENO1–3 sequences. Potential transcription factor binding sites (TFBS) and CpG islands for regulating *ENO* gene expression were identified. Human *ENO1*, *ENO2*, *ENO3*, and *ENOF1* genes each contained CpG islands in the gene promoter regions consistent with higher-than-average levels of expression. Human *ENO3* and *ENO1* gene promoters also contained a diverse range of TFBS. The *ENO4* gene promoter comprised a CpG island and several TFBS, including AHR1 in the 5′-UTR region, which may suggest a role for ENO4 in aryl hydrocarbon ligand binding or metabolism. Phylogeny studies of vertebrate *ENO1*, *ENO2*, and *ENO3* genes and enzymes suggested that they originated in a vertebrate ancestor from gene duplication events of an ancestral *ENO1*-like gene >500 million years ago.

Keywords: vertebrate, amino acid sequence, enolase, evolution, bioinformatics

Introduction

Enolase (ENO; EC 4.2.1.11) genes and proteins serve multiple functions in the body, including catalyzing 2-phospho-D-glycerate hydro-lyase activity in glycolysis[1] or playing roles in hypoxia tolerance,[2] tumor suppression,[3] and cell surface plasminogen binding[4] or acting as a lens *tau*-crystallin,[5] a DNA-binding protein[6] or a tubulin/microtubule binding protein during myogenesis.[7] Three major *ENO*-like genes have been described on the human genome, *ENO1*, *ENO2*, and *ENO3*, which encode the α-, β-, and γ-subunits, respectively.[8–11] Two other human *ENO*-like genes have also been reported, *ENO4* (or *ENOLL*)[12] and *ENO5* (also called *ENOF1* or *ENOSF1*),

Correspondence: Roger S Holmes
School of Biomolecular and Physical Sciences, Griffith University, Nathan, 4111 QLD, Australia
Tel +61 7 3735 5077
Fax +61 7 3735 7773
Email r.holmes@griffith.edu.au

originally identified as encoding an antisense transcript to the thymidylate synthase (*TS*) gene[13] which may play a role in regulating the *TS* locus.[14]

Biochemical studies of vertebrate enolases have characterized several dimeric isozymes containing α-, β-, and γ-subunits, which are differentially but widely distributed in the tissues of the body.[1,15] *ENO3* encodes the β-subunit and is predominantly expressed in muscle, whereas *ENO2* is more restricted to neural tissues (also called neuron-specific enolase or γγ), while *ENO1* is expressed in virtually all the tissues of the body, including embryonic tissues, and encodes the α-subunit.[16,17] During vertebrate development, major changes occur in the expression of these genes with a switch from *ENO1* → *ENO3* and a change from αα to ββ in skeletal muscle and a similar switch from *ENO1* → *ENO2* in nervous tissues with an associated change from αα to an αγ and γγ enolase isozymes.[1,18,19] Evolutionary studies have shown that DNA sequences encoding the enolase gene family are highly conserved from yeast to mammalian organisms and that the gene duplication events generating the *ENO1*, *ENO2*, and *ENO3* genes may have predated the appearance of vertebrates.[15,20]

Structural and molecular modeling studies of lobster,[21] yeast,[22,23] rabbit,[24] and human ENO2[25] enolases have shown that each polypeptide subunit contains at least two major domains with distinct roles. The C-terminal domain folds into an α/β barrel with a typical sequence of $\beta_2\alpha_2(\beta\alpha)6$ in secondary structure which forms the active site,[24] while the N-terminal domain contains a long, flexible loop that folds back onto the active site.[26,27] ENO catalytic activity has an absolute requirement for two divalent cations (Mg^{2+}) binding at distinct sites: a substrate-binding site which induces a conformational change and a chelation site which positions the N-terminal 'flap' over the active site entrance.[21–28] ENO1–3 catalyze the reversible elimination of water from 2-phosphoglycerate (2-PGA) to form phosphoenolpyruvate, and the two active site magnesium ions apparently facilitate the reaction by activating the C2 proton of 2-PGA and stabilizing the charged intermediate.[28]

Structures of three human enolase (*ENO*) genes have been reported, including human *ENO1*,[29] *ENO2*,[9] and *ENO3*.[30] These genes contained 12 exons and showed a high degree of sequence conservation and consistency in the positioning of the introns which suggested a common evolutionary origin for these genes. The 5′-flanking putative gene promoter regions for these genes were also highly conserved and contained a CpG island in each case. The *ENO1* and *ENO2* genes lacked canonical TATA and CAAT boxes in the 5′-promoters, which is consistent with these being housekeeping genes, whereas the human *ENO3* putative 5′-promoter contained an upstream TATA box. Each of these genes undergoes exon shuffling, generating several isoproteins in each case,[31] which may perform functions that are distinct from catalyzing the glycolytic reaction. ENO1 isoforms, for example, serve different roles as a hypoxic stress protein, lens crystallin, autoimmune antigen, cell surface plasminogen receptor, and a transcriptional repressor of the Myc proto-oncogene (called myc-binding protein or MBP1).[2–4,32–34]

This article reports predicted gene structures and amino acid sequences for several vertebrate enolase genes (*ENO*) and enzymes (ENO) previously not reported, including three closely related enolase family members (ENO1, ENO2, and ENO3) and two other enolase-like genes and proteins (ENO4 and ENO5) that have not been extensively investigated. Predicted secondary and tertiary structures for vertebrate enolases and conserved regulatory regions for mammalian *ENO* promoters are also described as well as the structural and evolutionary relationships of these genes and enzymes.

Methods

Vertebrate enolase gene and protein identification

Basic Local Alignment Search Tool (BLAST) studies were undertaken using Web tools from the National Center for Biotechnology Information (see http://blast.ncbi.nlm.nih.gov/Blast.cgi).[35] Protein BLAST analyses used human ENO1, ENO2, ENO3, ENO4, and ENO5 amino acid sequences which are previously described (Table 1). Nonredundant protein sequence databases for several vertebrate genomes were examined using the BLASTP algorithm, including human (*Homo sapiens*),[11] orangutan (*Pongo abelii*) (see http://genome.wustl.edu), rhesus monkey (*Macaca mulatta*),[36] marmoset (*Callithrix jacchus*) (see http://genome.ucsc.edu/), cow (*Bos taurus*),[37] horse (*Equus caballus*),[38] mouse (*Mus musculus*),[39] rat (*Rattus norvegicus*),[40] opossum (*Monodelphis domestica*),[41] chicken (*Gallus gallus*),[42] frog (*Xenopus tropicalis*),[43] zebrafish (*Danio rerio*),[44] and nematode (*Caenorhabditis elegans*) (see http://genome.ucsc.edu/). This procedure produced multiple BLAST 'hits' for each of the protein databases, which were individually examined and retained in FASTA format, and a record was maintained of the sequences for predicted mRNAs and encoded ENO-like proteins. These records were derived from annotated genomic sequences using the gene prediction method: GNOMON and predicted sequences with high similarity scores for human ENO1, ENO2, ENO3, ENO4, and ENO5 (or ENOF1) (see Tables 1 and 2). Predicted ENO-like protein sequences

Table 1 Vertebrate and nematode enolase genes and proteins

Species	Species	*Enolase* gene	Alternate name(s)	RefSeq ID [1]predicted	GenBank ID	UNIPROT ID	Amino acids	Chromosome location	Exons[2] (strand)	Gene size kbps	Expression[3] (tissues)	Subunit MW
Human	*Homo sapiens*	*ENO1*	ENOA; C-myc;	NP_001419	BC001810	P06733	434	1:8,844,009-8,857,554	11 (–ve)	13.5	5.3 (many)	47,169
		ENO2	ENOG	NP_001966	BC002745	P09104	434	12:6,895,258-6,902,221	11 (+ve)	7.0	5.9 (brain)	47,269
		ENO3	ENOB	NP_001967	BC017249	P12939	434	17:4,795,870-4,801,063	11 (+ve)	5.2	3.4 (muscle)	46,932
		ENO4	ENOLL	[1]chr10_1192.1	BX647301	A6NNW6	628	10:118,599,068-118,631,167	14 (+ve)	32.1	0.3 (testis)	68,821
		ENOF1	rTS	NP_974487	BC001285	Q7L5Y1	443	18:664,308-702,587	15 (–ve)	38.3	3.0 (liver)	49,786
Orangutan	*Pongo abelii*	*ENO1*	ENOA	NP_001126461	CR860345	Q5R6Y1	434	1:221,571,852-221,585,217	11 (+ve)	13.4	na	47,169
		ENO2	ENOG	NP_001125817	na	na	434	[4]12:7,145,546-7,151,771	[4]8 (+ve)	[4]6.2	na	47,243
		ENO3	ENOB	[1]XP_002826931	na	na	434	17:4,885,981-4,890,791	11 (+ve)	4.8	na	47,045
		ENO4	ENOLL	[1]XP_002821225	na	na	628	10:116,308,307-116,340,495	14 (+ve)	32.2	na	68,787
		ENOF1	rTS	NP_974487	CR858528	Q7L5Y1	443	18:32,014,225-32,055,160	14 (+ve)	40.9	na	49,786
Rat	*Rattus norvegicus*	*ENO1*	RGD:2553	NP_036686	BC063174	P04764	434	5:167,396,702-167,405,208	11 (+ve)	8.5	2.8 (wide)	47,128
		ENO2	RGD:2554	NP_647541	BC060310	P07323	434	4:160,891,006-160,897,723	11 (–ve)	6.7	1.0 (brain)	47,141
		ENO3	RGD:2555	NP_037081	BC083566	P15429	434	10:57,537,632-57,542,235	11 (+ve)	4.6	0.5 (muscle)	47,014
		ENO4	RGD:1308333	[1]chr1.2629	na	D3ZFY3	574	1:265,466,516-265,487,761	14 (+ve)	21.2	na	63,939
		ENOF1	RGD:3829	[1]chr8.717	na	na	443	8:76,322,813-76,388,182	13 (–ve)	65.4	na	47,563
Opossum	*Monodelphis domestica*	*ENO1*	ENOA	[1]XP_001362200	na	na	434	4:428,960,397-428,973,722	11 (+ve)	13.3	na	47,091
		ENO2	ENOG	[1,4]chr8.625.1	na	na	[4]282	[4]8:108,411,282-108,429,544	[4]9 (+ve)	[4]18.3	na	[4]31,095
		ENO3	ENOB	[1]XP_001366144	na	na	434	2:279,495,723-279,500,514	11 (–ve)	4.8	na	47,108
		ENO4	ENOLL	[1]chr1.10.018	na	na	619	1:93,086,111-93,118,373	14 (–ve)	32.3	na	67,605
Chicken	*Gallus gallus*	*ENO1*	ENOA	NP_990451	D37900	P51913	434	21:3,197,543-3,206,401	11 (–ve)	8.9	na	47,348
		ENO2	ENOG	NP_990207	AB004291	O57391	434	1:80,447,772-80,453,183	11 (+ve)	5.4	(brain)	47,308
		ENO3	ENOB	na	na	P07322	434	na	na	na	na	47,196
		ENO4	ENOLL	[1]chr6.491.1	na	na	567	6:30,441,484-30,456,523	12 (+ve)	15.0	na	62,970
Zebrafish	*Danio rerio*	*ENO1A*	ENOA.1	NP_997887	BC071359	Q6IPQ5	432	[5]23:22,152,060-22,170,789	11 (+ve)	18.7	na	47,060
		ENO1B	ENOA.2	NP_956989	BC059434	na	433	[5]6:44,237,409-44,245,891	11 (+ve)	8.5	na	47,303
		ENO2	ENOB	NP_001003848	BC072713	Q6GQM9	434	[5]19:4,701,771-4,714,958	11 (+ve)	13.2	na	46,841
		ENO3A	ENOC.1	AAH92869	BC092869	Q568G3	433	[5]23:42,253,515-42,263,149	11 (+ve)	9.6	na	47,432
		ENO3B	ENOC.2	AAH92869	BC092869	Q568G3	433	[5]23:42,204,754-42,214,388	11 (+ve)	9.6	na	47,432
		ENO4	ENOLL	NP_001071015	BC124782	Q08BC6	576	[5]17:18,823,353-18,839,296	13 (–ve)	15.9	na	62,758
		ENOF1	rTS; ENOSF1	NP_001070210	BC124261	na	441	[5]7:59,631,111-59,651,359	13 (+ve)	20.2	na	49,474
Nematode	*Caenorhabditis elegans*	*ENO1*	enol-1	NM_110127178	T21B10.2	Q27527	434	II:8932027-8933529	3 (+ve)	1.5	na	46,617

Notes: GenBank IDs are derived from the NCBI database http://www.ncbi.nlm.nih.gov/genbank/; Ensembl ID was derived from Ensembl genome database http://www.ensembl.org; UNIPROT refers to UniprotKB/Swiss-Prot IDs for individual enolases (see http://kr.expasy.org). [1]Predicted Ensembl amino acid sequence; [2]Coding exons; [3]Comparison with rate of average gene expression; [4]Partial sequence observed; [5]Contigs are identified for zebrafish genome.
Abbreviations: RefSeq, reference amino acid sequence; na, not available; bps, base pairs of nucleotide sequences.

Table 2 Vertebrate enolase genes and proteins

Species	Species	*Enolase* gene	Alternate name(s)	RefSeq ID [1]predicted	GenBank ID	UNIPROT ID	Amino acids	Chromosome location	Exons[2] (strand)	Gene size kbps	Expression[3] level (tissues)	Subunit MW
Rhesus	*Macaca mulatta*	*ENO1*	ENOA	NP_001182540	na	na	434	1:11,900,376-11,914,067	11 (–ve)	13.7	na	47,155
		ENO2	ENOB	[1]XP_001110839	na	na	434	11:7,102,131-7,109,383	11 (+ve)	7.3	na	47,243
		ENO4	ENOLL	[1]XP_001095040	na	na	607	9:116,480,199-116,513,823	13 (+ve)	33.6	na	66,187
		ENOF1	rTS	[1]XP_001088491	na	na	403	18:13,384,307-13,430,039	14 (+ve)	45.7	na	44,613
Marmoset	*Callithrix jacchus*	*ENO1*	ENOA	[1]XP_002750292	na	na	434	7:43,334,114-43,347,694	10 (–ve)	13.6	na	47,122
		ENO2	ENOG	[1]XP_002752315	na	na	434	9:18,767,629-18,774,222	11 (+ve)	6.6	na	47,239
		ENO3	ENOB	[1]XP_002763325	na	na	434	5:97,098,884-97,106,607	8 (+ve)	7.7	na	47,053
		ENO4	ENOLL	[1]XP_002756669	na	na	627	12:104,592,284-104,623,808	14 (+ve)	31.5	na	68,882
		ENOF1	rTS	[1]XP_002757153	na	na	443	13:57,475,648-57,517,943	14 (+ve)	42.3	na	49,700
Mouse	*Mus musculus*	*Eno1*	MGI:95393; EnoA	NM_023119	BC089539	P17182	434	4:149,613,591-149,622,661	11 (+ve)	9.1	8.0 (many)	47,141
		Eno2	MGI:95394; EnoG	NM_013509	BC043708	P17183	434	6:124,711,314-124,718,340	11 (–ve)	7.0	5.7 (brain)	47,297
		Eno3	MGI:95395; EnoB	NM_007963	BC013460	P21550	434	11:70,471,376-70,475,931	11 (+ve)	4.6	2.8 (muscle)	47,025
		Eno4	MGI:2141717; EnoII	NM_178689	BC023285	Q8C042	619	19:59,017,947-59,045,230	14 (+ve)	27.3	0.2 (testis)	67,883
Cow	*Bos taurus*	*ENO1*	ENOA	NP_776474	BC103354	Q9XSJ4	434	16:41,639,384-41,650,358	11 (+ve)	11.0	na	47,326
		ENO2	ENOG	NP_001094595	BC150078	A6QR19	434	5:10,523,267-10,529,383	11 (–ve)	6.1	na	47,269
		ENO3	ENOB	NP_001029874	BC102988	Q3ZC09	434	19:26,813,742-26,818,158	11 (–ve)	4.4	na	47,096
		ENO4	ENOLL	[1]chr26.130.1	BC109948	na	592	26:37,637,207-37,663,970	14 (+ve)	26.8	na	65,032
		ENOF1	rTS	NP_001040015	BC112706	na	443	24:36,770,668-36,792,545	13 (–ve)	21.9	na	49,621
Horse	*Equus caballus*	*ENO1*	ENOA	[1]XP_001494912	CX594614	na	434	2:42,139,011-42,150,728	11 (+ve)	11.7	na	47,140
		ENO2	ENOG	[1]XP_001497628	na	na	434	6:34,377,376-34,383,120	11 (+ve)	5.7	na	47,227
		ENO3	ENOB	[1]XP_001504796	na	na	434	11:49,568,859-49,573,066	11 (–ve)	4.2	na	47,053
		ENO4	ENOLL	[1]XP_001497604	na	na	624	1:15,270,783-15,295,862	14 (–ve)	25.1	na	68,408
		ENOF1	rTS	[1]XP_001915737	na	na	380	8:41,192,562-41,219,081	12 (–ve)	26.5	na	42,271
Dog	*Canis familiaris*	*ENO1*	ENOA	[1]XP_536735	na	na	415	5:65,299,748-65,310,742	12 (–ve)	11.0	na	45,100
		ENO2	ENOG	[1]XP_534902	na	na	434	27:41,149,648-41,155,612	11 (–ve)	6.0	na	47,186
		ENO3	ENOB	[1]XP_536606	na	na	434	5:34,659,440-34,663,207	11 (–ve)	3.8	na	47,055
		ENO4	ENOLL	[1]chr28.31.004	na	na	538	28:30,393,882-30,417,107	13 (+ve)	23.2	na	59,170
		ENOF1	rTS	[1]XP_848625	na	na	443	7:70,539,571-70,568,299	14 (–ve)	28.7	na	49,753
Frog	*Xenopus tropicalis*	*ENO1*	ENOA	NP_989144	BC061287	Q6P8E1	434	[5]207:168,511-177,564	11 (+ve)	9.1	na	47,584
		ENO3	ENOB	NP_001080346	BC096516	Q4VA70	434	[5]5154:14-6,373	na	6.4	na	47,306
		ENO4	ENOLL	NM_001126520	BC158253	na	574	[5]859:194695-219824	13 (+ve)	25.1	na	62,651
		ENOF1	rTS	NP_001119992	BC155430	Q6INX4	436	[5]84:1,744,319-1,768,607	13 (+ve)	24.3	na	50,162

Notes: GenBank IDs are derived from the NCBI database http://www.ncbi.nlm.nih.gov/genbank/; Ensembl ID was derived from Ensembl genome database http://www.ensembl.org; UNIPROT refers to UniprotKB/Swiss-Prot IDs for individual enolases (see http://kr.expasy.org). [1]Predicted Ensembl amino acid sequence; [2]Coding exons; [3]Comparison with rate of average gene expression; [4]Contigs are identified for frog genome. **Abbreviations:** RefSeq, reference amino acid sequence; na, not available; bps, base pairs of nucleotide sequences.

were then subjected to analyses of predicted protein and gene structures.

BLAT analyses were subsequently undertaken for each of the predicted ENO amino acid sequences using the UC Santa Cruz genome browser (see http://genome.ucsc.edu/cgi-bin/hgBlat)[45] with the default settings to obtain the predicted locations for each of the vertebrate ENO-like genes, including predicted exon boundary locations and gene sizes for coding exons. Structures for human *ENO1*, *ENO2*, *ENO3*, *ENO4*, and *ENO5* isoforms (splicing variants) were obtained using the AceView Web site to examine predicted gene and protein structures (see http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/index.html?human).[31] The UC Santa Cruz genome browser (http://genome.ucsc.edu)[45] was used to examine comparative structures for vertebrate and *C. elegans* enolase genes and for identifying predicted CpG islands and transcription factor binding sites (TFBS) for human *ENO1*, *ENO2*, *ENO3*, *ENO4*, and *ENO5* genes.

Predicted structures, properties, and alignments of vertebrate enolase-like sequences

Predicted secondary and tertiary structures for human and other vertebrate ENO-like proteins were obtained using the PSIPRED v2.5 Web site tools (see http://bioinf.cs.ucl.ac.uk/psipred/psiform.html)[46] and SWISS-MODEL Web tools (see http://swissmodel.expasy.org/), respectively.[47,48] The reported tertiary structure for human ENO1[49] was used as the reference for the ENO1 tertiary structure, with a modeling range of residues 3–432; the reported structure for human ENO2 served as a reference for human ENO2 and ENO3 (modeling range of 2–431),[25,50] and the structure for *Escherichia coli* enolase served as a reference for human ENO4 (modeling range of 67–577).[51] Alignments of human ENO1–5 sequences were assembled using the ClustalW2 multiple sequence alignment program (see http://www.ebi.ac.uk/Tools/clustalw2/index.html).[52]

Results and discussion

Alignments of human enolase amino acid sequences

Amino acid sequence alignments for previously reported human (*H. sapiens*) ENO1 (α),[49,53] ENO2 (γ),[8,9,54] ENO3 (β),[24,55] ENO4 (also called ENOLL),[12] and ENO5 (or ENOF1)[13] sequences are shown in Figure 1 (also see Table 1). The amino acid sequences for the human ENO1, ENO2, and ENO3 subunits contained 434 residues, whereas the predicted human ENO4 and ENO5 sequences (deduced from respective

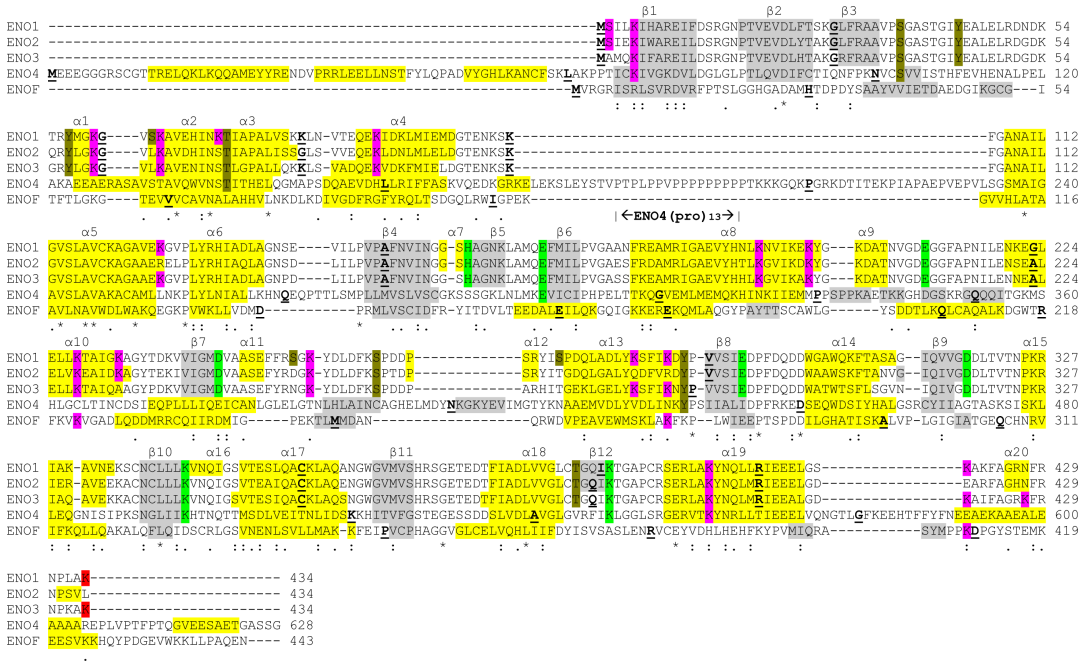

Figure 1 Amino acid sequence alignments for human ENO1, ENO2, ENO3, ENO4 (ENOLL), and ENO5 (ENOF1) sequences. See Table 1 for sources of human ENO1, ENO2, ENO3, ENO4, and ENO5 sequences (the latter two are predicted sequences); Symbol * shows identical residues for proteins; :, similar alternate residues; ., dissimilar alternate residues; key enolase active site, Mg^{2+}-binding and substrate-binding residues are in shaded green; predicted or known sites for acetylation are in shaded purple; predicted or known sites for phosphorylation are in shaded khaki; conserved C-terminal lysine residues for human ENO1 and ENO3 are in shaded red; β-sheets (β1–β12) are numbered according to human ENO1 sequences[49] and are in shaded grey; α-helices are also numbered according to the ENO1structure[47] and are in shaded yellow; bold underlined font shows residues corresponding to known or predicted exon start sites; exon numbers refer to human *ENO1* gene coding exons.

Table 3 Comparative amino acid sequences at key sites for vertebrate enolases (ENO1, ENO2, and ENO3)

Enolase gene (number)	Ser2 Acet	Lys5 Acet	Ser37 Phos	Tyr44 Phos	Tyr57 Phos	Lys60 Acet	Ser63 Phos	Lys64 Acet	Lys71 Acet	Thr72 Phos	Lys89 Acet	Lys126 Acet	His158 S binds	Glu167 S binds	Lys193 Acet	Lys199 Acet	Glu210 Act site	Lys228 Acet
***ENO1* (14)**	13 [1]el Pro	13 [1]mo Arg	**14**	13 [1]el His	11 [1]mo Phe [1]ra Phe [1]el His	**14**	12 [1]fr Gly [1]el Leu	12 [1]mo Gln [1]fr Arg	11 [1]fr Glu [1]zf1B Gln [1]el Glu	11 [1]fr Phe [1]zf1B His [1]el Lys	13 [1]el Asp	**14**	**14**	**14**	**14**	13 [1]el Arg	**14**	12 [1]op Thr [1]el Asn
***ENO2* (12)**	11 [1]ch Ala	9 [1]op Arg [1]ch Arg [1]zf Ser	**12**	**12**	11 [1]ch Phe	**12**	0 (Leu)	11 [1]ch Gln	0 (Ser/Thr)	11 [1]mo Arg	**12**	3 Arg/Ser	**12**	**12**	**12**	**12**	**12**	**12**
***ENO3* (12)**	3 9 (Ala)	**12**	**12**	10 [1]ch His [1]zf His	11 [1]ch Phe	**12**	0 (Leu)	**12**	11 [1]hu Ser	11 [1]zf Asp	**12**	**12**	**12**	**12**	**12**	**12**	**12**	**12**

Enolase gene (number)	Lys233 Acet	Asp245 Mg^{2+}	Ser254 Phos	Lys256 Acet	Ser263 Phos	Ser272 Phos	Lys281 Acet	Lys285 Acet	Tyr287 Phos	Glu293 [1]Mg^{2+}	Asp318 [1]Mg^{2+}	Lys343 Act Site	Thr390 Phos	Lys394 S binds	Lys406 Acet	Lys420 Acet	Lys434 [2]Micro
***ENO1* (14)**	12 [1]op Thr [1]el Asn	**14**	8 [1]ra Ala ch Asp fr Asp [1]zf1A Gly [1]zf1B Asp [1]el Asp	13 [1]zf1B Gln	13 [1]el Asn	10 [1]mo Thr [1]ra Thr [1]ho Thr [1]co Thr	12 [1]fr Met [1]el Gln	11 [1]mo Gln [1]co Arg [1]zf1B Glu	**14**	**14**	**14**	**14**	**14**	**14**	**14**	13 [1]el Asp	8 [1]op Gln [1]ch Asn [1]zf1A Ile [1]fr Asn [1]zf1B Asn [1]el Val
***ENO2* (12)**	**12**	**12**	0 (Asp/Glu)	**12**	**12**	3 (Thr)	0 (Gln)	0 (Arg)	11 [1]zf Phe	**12**	**12**	**12**	**12**	**12**	**12**	0 (Glu)	0 (L)
***ENO3* (12)**	2 (Ala)	**12**	1 (Asn)	11 [1]ch Arg	**12**	5 (Thr)	11 [1]ch Arg	11 [1]mo Gln	**12**	**12**	**12**	**12**	**12**	**12**	**12**	**12**	**11** [1]zf leu

Notes: Numbers of vertebrate ENO1, ENO2, and ENO3 amino acid sequences examined are shown: ENO1 (14), ENO2 (12), and ENO3 (12); an amino acid is identified at each key site; any observed substitution at the site is also identified; invariant vertebrate sites are shown in **bold**, invariant phosphorylation sites are in shaded blue, invariant acetylation sites are in shaded pink, active site or substrate binding invariant sites are in shaded yellow, and isozyme 'specific' variations in amino acids are shown in green.

Abbreviations: El, nematode (*C. elegans*); hu, human; mo, mouse; ra, rat; ho, horse; co, cow; fr, frog; zf, zebrafish; ch, chicken; op, opossum.

nucleotide sequences in each case) contained 628 and 443 amino acids, respectively (Figure 1).

Previous biochemical and genetic analyses of human and mouse ENO1,[49,56,57] ENO2,[8,9,58] and ENO3[54,59] have enabled predictions of key residues for human ENO1, ENO2, and ENO3. These included active site residues: Glu210 (proton donor), Lys343 (proton acceptor), His158, Glu167, Glu293, Asp318, and Lys394 (substrate-binding sites); Mg^{2+} chelating sites: Asp245, Glu293, and Asp318 (required for catalysis and stabilizing the dimer); C-terminal Lys434 for ENO1 and ENO3 (required for the binding to neuronal and skeletal muscle plasma membranes, respectively); acetylated residues: Ser2 and Lys residues **60**, **64**, 71 (ENO1), **89**, 126 (ENO1 and ENO3), **193**, **199**, **228**, 233 (ENO1 and ENO2), **256**, 281, and 285 (ENO1 and ENO3), **406** and 420 (ENO1 and ENO3); and phosphorylated residues Ser**37**, 63, and 263 (ENO1) and **272**; **Tyr44**, **57,** and **287**; and **Thr72** and **390** (residues in **bold** are shared among human ENO1, ENO2, and ENO3 sequences). Given the potential roles of acetylated and phosphorylated ENO sites in regulating cellular metabolism,[56,57] the existence of conserved and isozyme-specific sites are of particular significance for these enzymes.

Table 3 compares the amino acid residues localized in each of these key sites for ENO1, ENO2, and ENO3 from 13 vertebrate species and for ENO1 obtained for the nematode, *C. elegans*. In addition to the active site and substrate-binding and Mg^{2+}-chelating sites, several previously described acetylation and phosphorylation sites were also strictly conserved among the ENO sequences examined. These included phosphorylation sites Ser37, Ser263, and Thr390; and lysine acetylation sites 60, 89, 193, 199, and 406. In contrast, there are a number of predicted 'isozyme'-specific translationally modified sites, including two ENO1 phosphorylated residues: Ser63 (ENO1)/Leu63 (ENO2 and ENO3); Ser254 (ENO1)/Asp or Glu (ENO2) and Asn (ENO3); and four ENO1 acetylated residues: Lys71 (ENO1 and ENO3)/Ser or Thr71 (ENO2); Lys281 (ENO1 and ENO3)/Gln281 (ENO2); Lys285 (ENO1 and ENO3)/Arg285 (ENO2); and Lys420 (ENO1 and ENO3)/Glu420 (ENO2). In addition, the C-terminal 434Lys, previously shown to play a microlocalization role for ENO1 in neuronal cells[17,18] and for ENO3 in binding the ββ-isozyme to muscle filaments,[7] is predominantly conserved among the vertebrate sequences examined, but has been substituted with Leu434 for the 12 vertebrate ENO2 sequences. With the exception of the differences for the C-terminal residue, the significance of these isozyme-specific changes in sequences remains to be determined.

Alignments of vertebrate ENO1, ENO2, and ENO3 and nematode ENO1 amino acid sequences examined showed between 69% and 97% identities, suggesting that these are products of one gene family, which is highly conserved during vertebrate and invertebrate evolution (Table 4). In addition, sequences of multiple zebrafish (*D. rerio*) ENO1 (designated as ENO1A and ENO1B) and ENO3 (ENO3A and ENO3B) also showed similar or identical sequences (88% and 100%, respectively). Comparisons of the vertebrate ENO1–3 sequences with the predicted ENO4 (also called ENOLL) and ENO5 (also called ENOF1), however, showed large differences in sequence identity, with only ~20% identical sequences for the ENO1–3 proteins with ENO4 and <10% with ENO5, suggesting that the latter are members of distinct *ENO* gene families.

Secondary and tertiary structures for vertebrate enolases

Predicted secondary structures for human ENO4 and ENO5 sequences were compared with those previously reported for human ENO1,[49,57] ENO2,[23,25] and ENO3[55] (Figure 1). α-Helix and β-sheet structures for these sequences were numbered as for those described for human ENO1.[49] The predicted human ENO4 secondary structure[51] was similar to those for human ENO1–3 sequences, although a number of additional structures were observed, including three N-terminal α-helices; a poly-proline ENO4 sequence (residues 177–233); two β-sheets (designated as ENO4 β5 and β6) (residues 394–415); and an extended C-terminal sequence (residues 605–628) containing an α-helix. The predicted ENO5 secondary structure also exhibited similarities with the human ENO1–3 secondary structures which were previously reported[25,49,55] although major differences were apparent, particularly in the extended C-terminal region. Given that the *ENO5* (also called *ENOF1*) gene has a proposed role in encoding an antisense transcript for *TS*[13] and in regulating the *TS* locus,[14] it is likely that ENOF1 does not function in catalyzing the glycolytic enolase reaction. This is supported by examining the predicted human ENO5 amino acid sequence (Figure 1) for which several key ENO residues have been substituted, including the active site residues and conserved acetylated and phosphorylated residues for ENO1–3.

Figure 2 compares previously reported structures for human ENO1,[49] ENO2,[23,25] and ENO3[55] protein sequences with a predicted tertiary structure for the human ENO4 (or ENOLL) subunit (based on the reported tertiary structure for E. coli enolase[51]). Two major differences were observed

Table 4 Percentage identities for vertebrate and nematode enolase amino acid sequences

	Hu ENO1	Co ENO1	Zf ENO1A	Zf ENO1B	Hu ENO2	Co ENO2	Zf ENO2	Hu ENO3	Co ENO3	Zf ENO3A	Zf ENO3B	Hu ENO4	Co ENO4	Zf ENO4	Hu ENOF1	Co ENOF1	Zf ENOF1	El ENO1
Hu ENO1	100	**95**	**90**	**86**	**83**	**82**	**79**	**83**	**83**	**84**	**84**	24	20	23	7	7	7	**75**
Co ENO1	**95**	100	**89**	**86**	**83**	**82**	**78**	**83**	**84**	**84**	**84**	23	20	23	7	7	7	**72**
Zf ENO1A	**90**	**89**	100	88	**83**	**83**	**78**	**82**	**83**	**84**	**84**	24	19	23	6	8	7	**72**
Zf ENO1B	**86**	**86**	**88**	100	**81**	**80**	**78**	**81**	**81**	**82**	**82**	23	19	23	4	7	4	**72**
Hu ENO2	**83**	**83**	**83**	**81**	100	**98**	**83**	**83**	**83**	**82**	**82**	24	20	22	3	6	4	**73**
Co ENO2	**82**	**82**	**83**	**80**	**98**	100	**82**	**82**	**82**	**81**	**81**	22	20	22	3	5	4	**73**
Zf ENO2	**79**	**78**	**78**	**78**	**83**	**82**	100	**80**	**81**	**76**	**76**	22	19	23	7	8	7	**69**
Hu ENO3	**83**	**83**	**82**	**81**	**83**	**82**	**80**	100	**97**	**85**	**85**	23	20	22	5	6	4	**73**
Co ENO3	**83**	**84**	**83**	**81**	**83**	**82**	**81**	**97**	100	**85**	**85**	23	20	22	5	6	4	**73**
Zf ENO3A	**84**	**84**	**84**	**82**	**82**	**81**	**76**	**85**	**85**	100	100	21	21	22	6	8	6	**75**
Zf ENO3B	**84**	**84**	**84**	**82**	**82**	**81**	**76**	**85**	**85**	100	100	21	21	22	6	8	6	**75**
Hu ENO4	24	23	24	23	24	22	22	23	22	21	21	100	**83**	36	7	7	4	20
Co ENO4	20	20	19	19	20	20	19	20	20	21	21	**83**	100	34	3	5	3	21
Zf ENO4	23	23	23	23	22	22	23	22	22	22	22	36	34	100	1	1	1	22
Hu ENOF1	7	7	6	4	3	3	7	5	5	6	6	7	3	1	100	**88**	**73**	5
Co ENOF1	7	7	8	7	6	5	8	6	6	8	8	7	5	1	**88**	100	**73**	4
Zf ENOF1	7	7	7	4	4	4	7	4	4	6		4	3	1	**73**	**73**	100	5
El ENO1	**75**	**72**	**72**	**72**	**73**	**73**	**69**	**73**	**73**	**75**	**75**	20	21	22	5	4	5	100

Notes: Numbers show the percentage of amino acid sequence identities. Numbers in **bold** show higher sequence identities for lipases from the same gene family: ENO1, ENO2, ENO3, ENO4, and ENOF1 (or ENO5).
Abbreviations: Hu, human; co, cow; zf, zebrafish; el, nematode (*C. elegans*).

for the ENO4 tertiary structure: an extended chain region corresponding to a poly-proline amino acid segment (ENO4 residues 177–232) not present in the human ENO1–3 structures and two additional β-sheet segments (human ENO4 residues 334–354 and 394–415) (also see Figure 1). The rest of the predicted ENO4 structure was similar to those previously described for human ENO1–3 subunits, although the homology model for ENO4 did not include the extended N-terminal and C-terminal regions. Although ENO4 displays a similar structure to the human ENO1–3 active site zone (see Figure 2), the amino acid alignments results (Figure 1) show that ENO4 lacks active site and other key residues and would not be expected to function in catalyzing the glycolytic reaction and an alternate role should be considered.

Gene locations and exonic structures for vertebrate *ENO* genes

Tables 1 and 2 summarize the predicted locations for vertebrate and nematode (*C. elegans*) enolase-like (*ENO*) genes based upon BLAT interrogations of genomes using the reported sequences for human ENO1 (α),[49,53] ENO2 (γ),[8,9] ENO3 (β),[55] ENO4 (also called ENOLL),[12] and ENO5 (or ENOF1)[13] and the predicted sequences for other vertebrate ENO-like proteins using the UC Santa Cruz genome browser.[45] The mammalian *ENO*-like genes were transcribed on either strand, depending on the *ENO* gene or genome examined. Figure 1 summarizes the coding exon start sites for human *ENO1*, *ENO2*, and *ENO3* genes, showing 11 coding exons in identical or similar positions which are consistent with previous reports.[9,53,54] In contrast, the human *ENO4* gene contained 13 coding exons, including two additional exons encoding an extended N-terminal sequence and a poly-proline segment not observed in the human ENO1–3 and ENOF1 sequences. In addition, the ENO4 coding exon start sites were in distinct positions to those reported for human *ENO1*, *ENO2*, and *ENO3* genes. Comparisons of predicted *ENO4* vertebrate gene structures showed that each contained predominantly 14 coding exons, whereas dog, frog, and zebrafish *ENO4* genes contained 13 coding exons. Genomic analyses of the human *ENO5* gene showed that 15 predicted coding exons were present, which corresponded to different locations to those previously reported for human *ENO1*, *ENO2*, and *ENO3*[9,53,54] and described here for human *ENO4* (Figure 1). Comparative analyses of vertebrate *ENO5* genes showed that the number of coding exons varied with the species examined, from 12 for the horse *ENO5* (or *ENOF1*) gene to 15 for the human *ENO5* gene.

Figure 3 shows the predicted structures of mRNAs for human *ENO1*, *ENO2*, *ENO3*, *ENO4*, and *ENO5* transcripts for the major isoform in each case.[31] These human mRNA transcripts varied in length from 5.2 to 13.5 kb for the *ENO1–3* genes and up to 32 and 38 kb for human *ENO4* and *ENO5* genes, respectively. The human *ENO1–3* gene transcripts contained extended 5′-untranslated (UTR) and 3′-UTR regions, with the latter also observed for the *ENO4* gene and the former for the human *ENO5* gene. Human *ENO1* and *ENO5* transcripts were encoded on the negative strand, whereas human *ENO2*, *ENO3*, and *ENO4* transcripts were transcribed on the positive strand (Table 1). In each case for these human *ENO* genes, a CpG island was observed within the 5′-regulatory promoter regions for these genes. CpG islands are typically observed within the promoters of housekeeping genes and may enhance a high level of expression for these genes.[60] The levels of expression for these human *ENO* genes have been compared with the average level of gene expression observed in the human genome[31] (see Table 1). Four of these genes exhibited higher levels of expression (*ENO1* (x5.3), *ENO2* (x5.9), *ENO3* (x3.4), and *ENO5* (x3.0)), whereas human *ENO4* exhibited a lower-than-average level of expression (x0.3). The higher levels of *ENO1–3* expression were also observed for the mouse and rat genes (see Tables 1 and 2).

In addition to the CpG islands observed for the human *ENO* gene promoters, these sequences also contained 78 predicted TFBS including within the 5′-upstream promoter region (20 sites), 5′-UTR promoter region (7 sites), and intron 1 (seven sites), which are usually associated with regulating gene transcription (Table 5). Several of these sites have been shown to play significant roles in regulating *ENO* gene expression. For example, the HIF1 site (hypoxia-inducible factor located in the 5′-UTR region of *ENO1*) activates genes encoding proteins that mediate responses to hypoxia, in association with coactivator proteins, such as the CREB-binding protein,[2,61] which also has a binding site in this region. An alternate translated form of ENO1 (named MBP1), which is predominantly located in the nucleus, has been characterized as a c-Myc promoter-binding protein that negatively controls transcription of this proto-oncogene[32] which identifies *ENO1* as a potential tumor suppressor. Sousa and coworkers[62] have also shown that interferons induce *ENO1* expression in target cells by activating mitogen-activated protein kinases and the transcription factor (CREB). The 5′-noncoding exon for *ENO3* contains many predicted motifs for transcriptional regulation, including Sp1, activator protein 1 and 2, CCAAT box transcription factor/nuclear factor 1, and

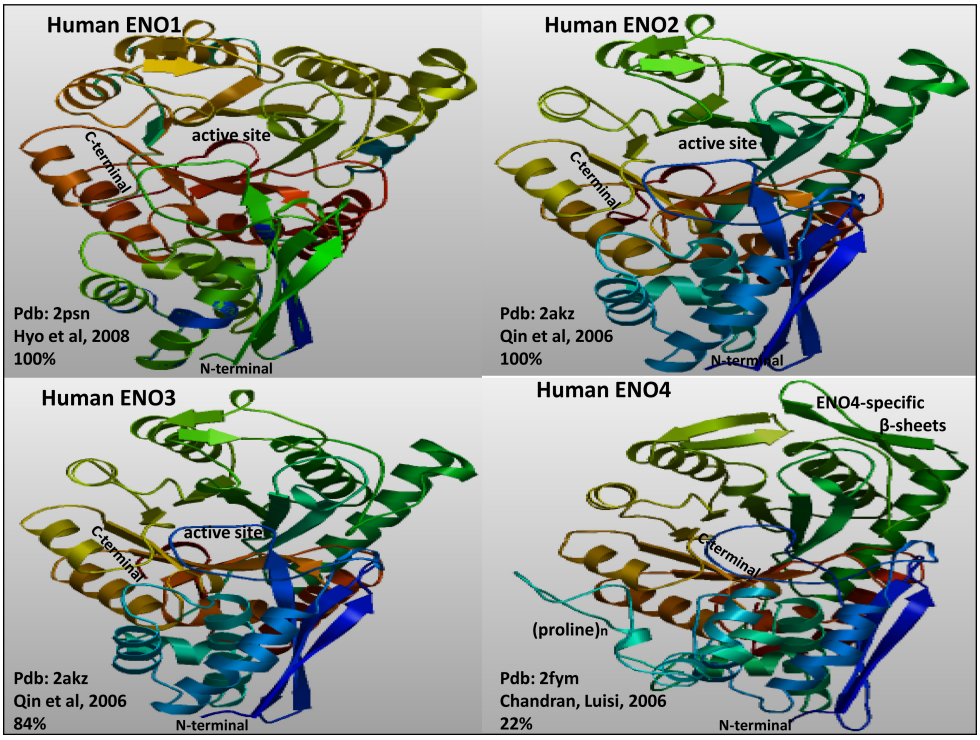

Figure 2 Known or predicted tertiary structures for human ENO1, ENO2, ENO3, and ENO4 (ENOLL). Tertiary structures were obtained using SWISS-MODEL methods; the rainbow color code describes the known tertiary structures from the N- (blue) to C-termini (red color) for human ENO1,[49] ENO2,[25] and ENO3;[55] and the predicted structure for human ENO4 (ENO4 (ENOLL) structure based on E. coli enolase[51]); arrows indicate directions for β-sheets; known or predicted active site, N-terminal and C-terminal regions are shown, as are predicted structures and locations for ENO4 β5 and β6 sheets and a poly-proline ((proline)n) sequence.

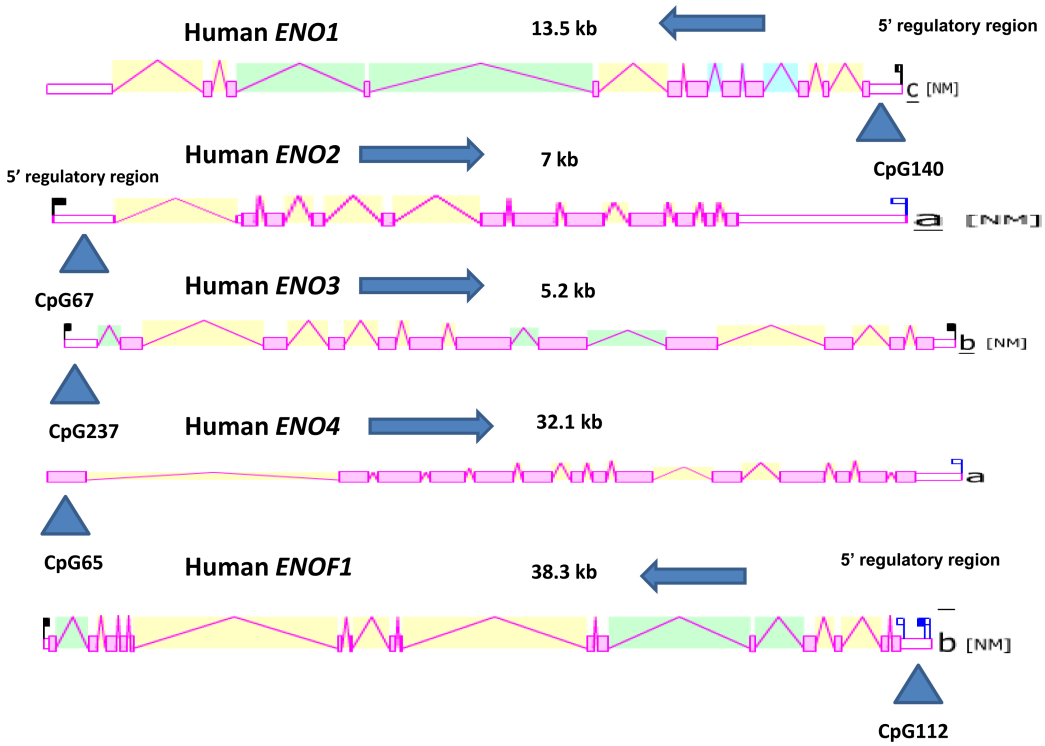

Figure 3 Gene structures and major splicing variant for human *ENO1*, *ENO2*, *ENO3*, *ENO4*, and *ENO5* (*ENOF1*) gene transcripts. Derived from AceView,[31] mature isoform variants (a) are shown with capped 5′- and 3′-ends for the predicted mRNA sequences, NM refers to the NCBI reference sequence, exons are in shaded pink, untranslated 5′- and 3′-sequences are in open pink, introns are represented as pink lines joining exons, the directions for transcription are shown as 5′ → 3′, sizes of mRNA sequences are shown in kilobases (kb), CpG islands are identified and numbered. (a), (b) and (c) refer to the major isoforms of enolase genes.

cyclic AMP, and for muscle-specific *ENO3* gene regulation, a CC(A + T-rich) 6GG box, M-CAT-box CAATCCT, and two myocyte-specific enhancer-binding factor 1 boxes.[30,63] Muscle-specific *ENO3* gene enhancers are also located within the first intron that bind myocyte-specific enhancer factor 2 proteins and G-rich-box binding factors.[64] A TFBS (AHR1) and a CpG island for regulating *ENO4* gene expression were also identified in the 5′-UTR region for *ENO4*, which may suggest a role for ENO4 (or ENOLL) in aryl hydrocarbon ligand binding or metabolism.

Comparative tissue expression of mouse enolase genes and differential functions for vertebrate enolases

Figure 4 presents 'heat maps' showing comparative gene expression for various mouse tissues obtained from GNF Expression Atlas Data using GNF1M (mouse) chips[65] (see http://genome.ucsc.edu; http://biogps.gnf.org). These data supported differential tissue expression for mouse enolase genes: *Eno1* showing highest levels in embryonic tissues, kidney, and brown adipose tissue; *Eno2* with highest expression levels in neural tissues; *Eno3* with highest levels of expression in skeletal muscle, heart, brown fat, bone, and prostate; and *Eno5* (or *Enof1*) showing a broad tissue expression profile. This is consistent with previous reports for these genes.[59,66,67] There were no 'heat map' results available for the mouse *Eno4* gene. Overall, however, mouse and human *ENO1*, *ENO2*, *ENO3*, and *ENO5* (or *ENOF1*) tissue gene expression levels were >3 times higher than the average level of gene expression which supports the key role played by these proteins and enzymes in glycolysis and in various multifunctional roles in the body (for ENO1–3)[23,49] and in regulating TS activity (ENOF1).[13,14] In contrast, the average human and mouse *ENO4* gene expression was below the average (×0.2–0.3) with highest levels observed in testis in each case (Tables 1 and 2).

Enolase genes and proteins are multifunctional with the three major genes (*ENO1*, *ENO2*, and *ENO3*) encoding α-, γ-, and β-subunits which form dimeric isozymes, performing a primary role in glycolysis, catalyzing 2-phospho-D-glycerate hydro-lyase activity.[1] These isozymes are also subject to differential localization within the cell. The α-subunit contains a C-terminal lysine residue, which facilitates binding with plasminogen in the neuronal plasma membrane and promotes its activation,[8,49] while the β-subunit C-terminal lysine facilitates binding to troponin in the Z-line of striated muscle fibers.[68,69] A reported association of a genetic deficiency for the enolase β-subunit with muscle weakness supports a role for this isozyme in localized adenosine triphosphate production within muscle fibers.[70] The physiological importance of enolase subcellular localization has also been demonstrated in studies of flagellar motility and energy production in *Chlamydomonas reinhardtii*.[71] ENO1 (αα) also plays a role in hypoxia tolerance,[2] tumor suppression,[3] cell surface plasminogen binding,[4] or acting as a lens *tau*-crystallin.[5] A differentially translated isoform of ENOαα (called MBP-1) also binds to the c-Myc promoter and acts as a transcriptional repressor and DNA-binding protein and is a potential candidate as a tumor suppressor.[6,32]

In contrast to the multifunctional roles in carbohydrate metabolism and other processes for α-, β-, and γ-subunits containing enolases, *ENO5* (*ENOF1* or *ENOSF1*) was originally identified as encoding an antisense transcript to the *TS* gene[13] with a proposed role in regulating the *TS* locus by the synthesis of signaling molecules involved in the downregulation of TS.[14] *ENO4* (or *ENOLL*) has also been reported as an enolase-like gene[12] but has only been described at the transcript level as yet. The lack of active site residues for this 'predicted protein' is suggestive of another function similar to that of catalyzing enolase activity, either as a protein (ENO4 or ENOLL) or as a transcript. The predicted 3-D structure for ENO4[51] shows significant similarities with the α-, β-, and γ-subunits, although with two additional β-sheets which may overlay the active site and an extended poly-proline chain, which may suggest another function to that of catalyzing enolase activity (Figure 2). It is relevant to note that the homology modeling method used to derive a predicted 3-D structure for human ENO4 was based on E. coli enolase[51] rather than on a mammalian α-, β-, and γ-containing subunit structure. E. coli enolase also plays a role within a multienzyme complex called the RNA degradosome[72] which may suggest a similar role for human ENO4. The location of a TFBS (AHR1) within the ENO4 promoter region also suggests a role for ENO4 (or ENOLL) in aryl hydrocarbon ligand binding or metabolism.

Phylogeny and divergence of enolase sequences

A phylogenetic tree (Figure 5) was calculated by the progressive alignment of 39 vertebrate ENO1, ENO2, and ENO3 amino acid sequences with the nematode (*C. elegans*) enolase sequence serving as the 'root' for the tree (see Tables 1 and 2). The phylogram showed clustering of the ENO sequences into three groups which were consistent with their evolutionary relatedness, as well as groups for ENO1, ENO2, and ENO3,

Table 5 Predicted transcription factor binding sites identified for human *Enolase* genes

Human *ENO* gene	5′-region	Exon1/5′UTR	Predicted transcription factor binding sites/CpG islands													
			Intron 1	Exon 2	Intron 2	Exon 3	Exon 4	Exon 5	Exon 6	Exon 7	Intron 7	Exon 8	Intron 8	Exon 9	Exon 11	Intron 12	Exon 15
ENO1		CpG140															
	CP2, c-Fos	[1]HIF1	NF-IL6	IK1		GATA	PBX1	TCF11	CUTL1	ZEB1		ZID					
	NF-E2	[2]CREB							STAT			c-Myb					
ENO2		CpG67															
		NF-E2	EVI1	GATA1	HMX3	NF-IL6	GATA	PAX4		HEN1	PBX1	NF-E2		NF-E2			
			LYF1	IK1		PPAR-g	IK3										
ENO3		CpG237															
	STAT, cREL	[3]Sp1	RORA1	PPAR-g			Hox-A9		PAX6		SRF1	PPAR-g	SREBP1		HLF		
	NF-κ-B, NRF2		Hox-A9						CUTL1		ARNT						
	LCR-F1, NF-E2		[4]MEF2A								SOX9						
	ARNT, c-Myc		ZID														
	c-Fos, n-Myc																
	USF, ARP1																
	LUN1																
ENO4		CpG65															
	PAX4, HOX-A9	ARNT, RFX1					GATA			SRF1		HNF1	NKX	NKX		PAX4	
	OCT1, SEF1	MIF1											EVI1			OCT	
ENO5		CpG112															
(ENOF1)								OCT				CDC5					LYF1

Notes: The human genome browser[45] was used to examine the predicted transcription factor binding sites for human *ENO1*, *ENO2*, *ENO3*, *ENO4*, and *ENO5* (or *ENOF1*) genes; UNIPROT IDs are shown for each protein. The CpG islands are individually numbered and are in shaded yellow.

Abbreviations: ARNT, aryl hydrocarbon receptor (P27540); ARP1, COUP transcription factor (P24468); CDC5, cell division cycle 5-like protein (Q99459); CP2, α-globin transcription factor (Q9ERA0); c-Fos, proto-oncogene factor (P01100); c-Myb, Myb transcriptional activator (P06876); c-Myc, proto-oncogene (P01106); n-Myc, proto-oncogene (P03966); CREB, cyclic-AMP responsive-element binding protein 1 (P16220); cREL, proto-oncogene (Q04864); CUTL1, homeobox protein cut-like 1 (P39880); EVI1, transcriptional regulator (P14404); GATA, transcription factor (P15976); HEN1, helix-loop-helix protein 1 (Q02575); HIF1, hypoxia-inducible factor 1 (Q16665); HLF, hepatic leukemia factor (Q16534); HMX3, homeobox regulator (P42581); HNF1, hepatocyte nuclear factor 1-alpha (P15257); Hox-A9, homeobox protein (P31269); IK1/IK3, ikaros transcription regulator proteins (Q13422); LCR-F1, transcription activator (Q14494); LUN1, E3-ubiquitin protein ligase (Q9 NS56); LYF1, transcription regulator (Q03267); MEF2A, myocyte specific enhancer 2A (Q02078); MIF1, homocysteine-responsive endoplasmic reticulum-resident ubiquitin-like domain member 1 protein (Q15011); NF-E2, transcription regulator protein (O14867); NF-IL6, CCAAT/enhancer-binding protein (P17676); NF-κ-B, transcription factor (P19838); NKX, homeobox protein (P78426); NRF2, nuclear factor 2 (Q16236); OCT1, POU transcription factor (P14859); PAX4, paired box protein 4 (P32115); PAX6, paired box 6 (P26367); PBX1, leukemia transcription factor (P40424); PPAR-g, peroxisome proliferator-activated receptor γ (P37238); RORA1, nuclear receptor ROR-α (P35397); SEF1, SL3–3 enhancement factor 1 (P30561); RFX1, MHC class II regulatory factor (P48377); Sox9, transcription factor (P48436); Sp1, transcription factor (P08047); SREBP1, sterol regulatory protein 1 (P36956); SRF1, serum transcription factor (P11831); STAT, transcription activator (P42224); TCF11, erythroid transcription factor (Q14494); USF, upstream stimulatory factor 1 (P22415); ZEB1, zinc-finger E-homeobox 1 (Q13088); ZID, zinc-finger/BTB protein 6 (Q15916); 5′ region, the upstream 5′ promoter; 5′-UTR, the 5′ untranslated region of the gene.

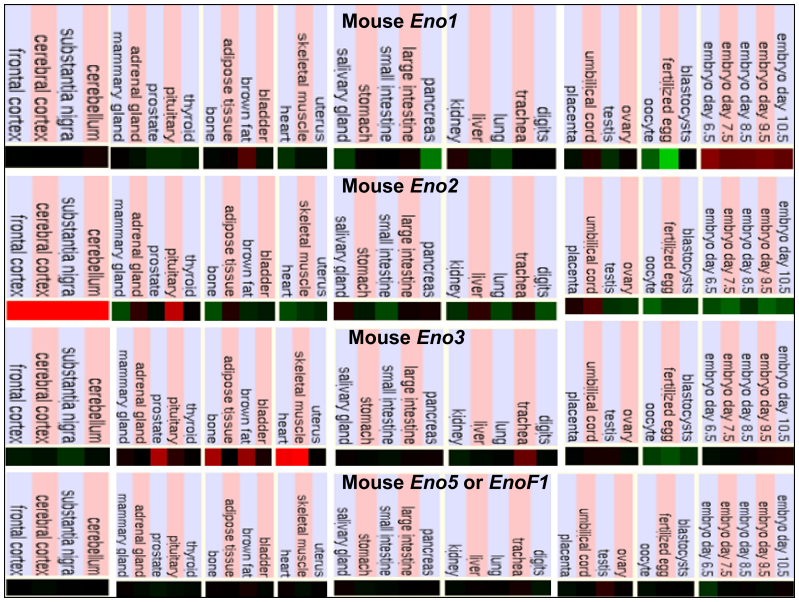

Figure 4 Comparative tissue expression for mouse enolase genes (*Eno1*, *Eno2*, *Eno3*, and *Enof1*). Expression 'heat maps' (GNF Expression Atlas 2 data)[65] were examined for comparative gene expression levels among mouse tissues for *Eno1*, *Eno2*, *Eno3*, and *Eno5* (or *Enof1*) genes showing high (red), intermediate (black), and low (green) expression levels, derived from mouse genome browser.[45]

and which were significantly different from each other (with bootstrap values of 91–100). It is apparent from this study of vertebrate *ENO1*, *ENO2*, and *ENO3* genes and proteins that these are ancient proteins for which a proposed common ancestor for these genes may have predated the appearance of fish >500 million years ago.[73] Tracy and Hedges[20] have examined this timing event further and have concluded that the *ENO1* and *ENO3* genes appeared first in actinopterygian, sarcopterygian, and chondrichthyan fishes and that the third gene duplication event generating *ENO2* occurred subsequently to the divergence of living agnathans (jawless fish, eg, lamprey) (~550 million years ago). Liang and coworkers[14] have also conducted a comprehensive phylogenetic analysis of ENO5 (also called ENOF1 and rTS) protein and showed that it has an extended distribution profile and exists in some groups of eubacteria, two fungal lineages, and most animal species from insects to mammals, demonstrating that *ENOF1* (*ENO5*) is a very ancient gene in biological evolution. There are no reports available on the phylogeny of the vertebrate ENO4 (or ENOLL) protein, although this present study shows that the gene is present among all vertebrate genomes examined.

Conclusion

These results demonstrate that vertebrate enolase (*ENO*) genes and encoded enzymes (ENO) comprise at least three distinct forms of enolases: 1) ENO1, ENO2, and ENO3; 2) ENO4; and 3) ENO5 (or ENOF1). The first group is further subdivided into three families corresponding to *ENO1*, *ENO2*, and *ENO3* genes, observed for each of the vertebrate genomes examined and previously reported for many mammals and chicken.[8–11] ENO1, ENO2, and ENO3 enzymes not only have a primary enzymatic role in catalyzing the 2-phospho-D-glycerate hydro-lyase activity in glycolysis[1] but also perform a number of other functions. For ENO1, these include a role in hypoxia tolerance,[2] tumor suppression,[3] cell surface plasminogen binding,[4] and acting as a lens *tau*-crystallin,[5] and for an isoform of ENO1 (called MBP-1), the functions include binding to the c-Myc promoter and serving as a potential tumor suppressor.[6,30] In contrast, vertebrate *ENO4* genes have only been described at the transcript level,[12] and little is known concerning the potential role(s) of the vertebrate *ENO4* gene, its transcript, or encoded protein, although the gene is present throughout vertebrate evolution. *ENO5* (*ENOF1* or *ENOSF1*) was originally identified as an antisense transcript to the *TS* gene, and a mitochondrial protein (ENOF1) has been reported which may play a role in regulating the TS locus.[13,14]

ENO1, ENO2, and ENO3 were each encoded by single genes among most vertebrate genomes examined, with the exception of the zebrafish genome, which contained two *ENO1*-like and *ENO3*-like genes. These genes are highly but differentially expressed in human and mouse tissues,

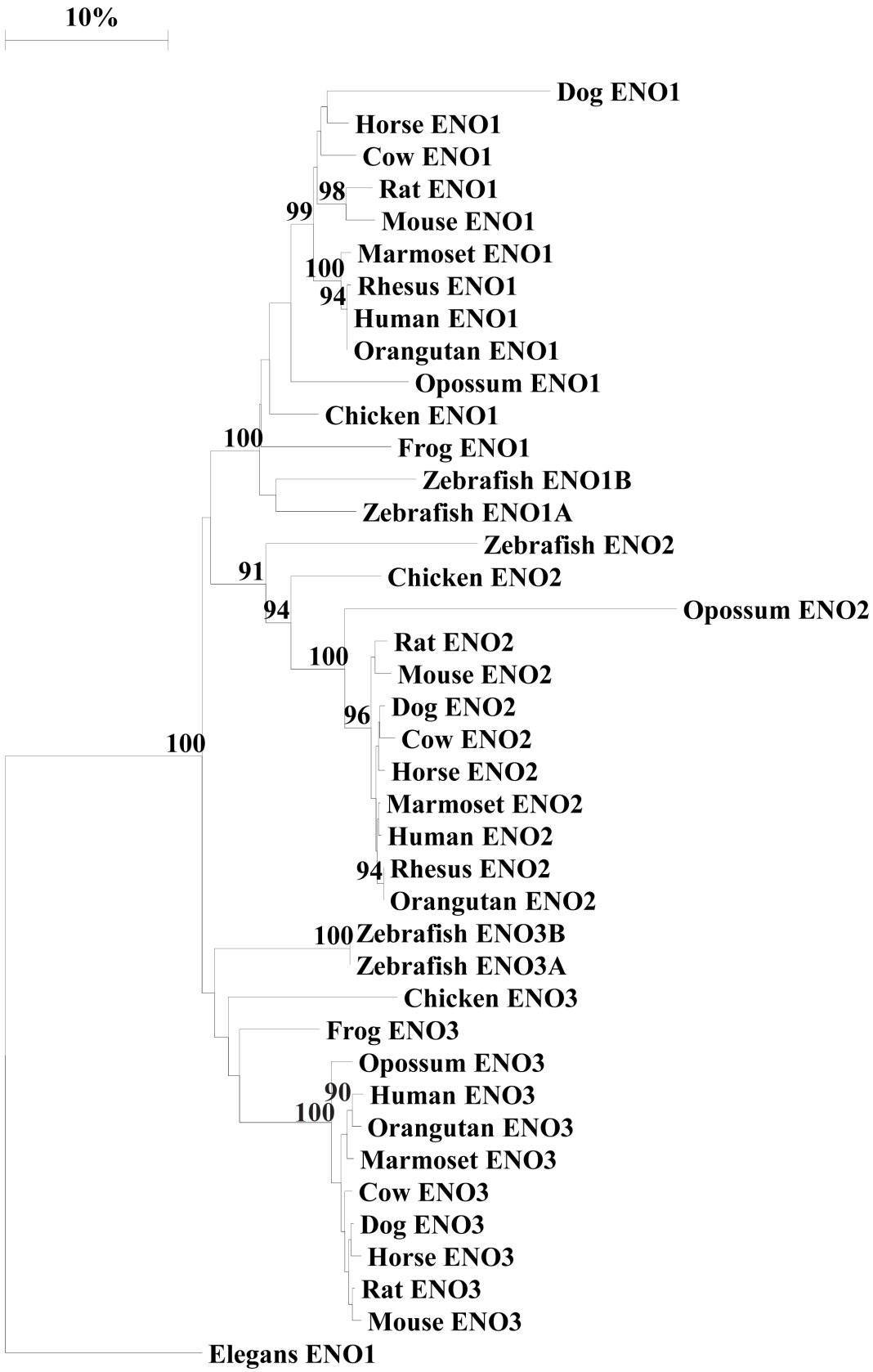

Figure 5 Phylogenetic tree of vertebrate enolase (ENO1, ENO2, and ENO3) with nematode enolase amino acid sequences. The tree is labeled with the enolase name and the name of the animal and is 'rooted' with the nematode enolase sequence (*C. elegans*). Note the three major clusters corresponding to the *ENO1*, *ENO2*, and *ENO3* gene families. A genetic distance scale is shown (% amino acid substitutions). The number of times a clade (sequences common to a node or branch) occurred in the bootstrap replicates are shown. Only replicate values of 90 or more which are highly significant are shown with 100 bootstrap replicates performed in each case.

with *ENO1* expression predominating in embryonic tissues; *ENO2* in neural tissues; and *ENO3* in skeletal and heart muscles,[59,64–66] and usually contained 11 coding exons. Predicted structures for human ENO4[51] and human ENO5 (Figure 1) proteins showed similarities with human ENO1,[49] ENO2,[23,25] and ENO3.[55] Human ENO4, however, exhibited at least four distinct structures, including an extended N-terminal sequence containing three predicted α-helices, two additional β-sheets, a poly-proline chain, and an extended C-terminal region with an additional predicted α-helix. Comparisons of ENO1, ENO2, and ENO3 amino acid sequences from vertebrates representative of mammals, birds, amphibians, and bony fish demonstrated that these are highly conserved proteins during evolution, not only for active site residues, but also for those involved in posttranslational changes, such as acetylation and phosphorylation, and for the C-terminal lysine residue for ENO1 and ENO3 sequences, which participate in localizing the enzyme within cell macromolecular structures.[7,68,69] Vertebrate ENO1, ENO2, and ENO3 sequences shared 78%–98% identities, but only 19%–24% with ENO4 and >10% predicted sequence identities with vertebrate ENOF1. Sequence alignments, key amino acid residues, and conserved predicted secondary and tertiary structures were examined, including active site residues (absent in ENO4 and ENOF1) and sites for Mg^{2+} and plasminogen binding and for acetylation and phosphorylation. Mutation studies of yeast enolase have served as useful models for examining specific roles of individual residues and likely impacts of genetic deficiencies for human enolases.[74]

Potential TFBS and CpG islands for regulating *ENO* gene expression were identified using bioinformatic techniques. Human *ENO1*, *ENO2*, *ENO3*, and *ENOF1* genes each contained CpG islands in the gene promoter regions consistent with the higher-than-average levels of *ENO* gene expression observed. Human *ENO3* and *ENO1* gene promoters also contained a diverse range of TFBS. The *ENO4* gene promoter contained a CpG island and several TFBS, including AHR1 in the 5′-UTR region, which may suggest a role for ENO4 in aryl hydrocarbon ligand binding or metabolism. Phylogeny studies of vertebrate *ENO1*, *ENO2*, and *ENO3* genes and enzymes suggested that they originated in a vertebrate ancestor from gene duplication events of an ancestral *ENO1*-like gene >500 million years ago, which is consistent with an earlier study[20] but is in contrast to the *ENOF1* gene which has a much wider biological distribution, including some eubacteria, fungi, as well as invertebrate and vertebrate animals.[14]

Acknowledgments

The expert assistance and advice of Dr Laura Cox of the Southwest Foundation for Biomedical Research, San Antonio, TX, USA and Dr Bharet Patel of Griffith University, Australia, are gratefully acknowledged.

Disclosure

The author reports no conflicts of interest in this work.

References

1. Rider CC, Taylor CB. Enolase isoenzymes in rat tissues. Electrophoretic, chromatographic, immunological and kinetic properties. *Biochim Biophys Acta*. 1974;365(1):285–300.
2. Semenza GL, Jiang BH, Leung SW, et al. Hypoxia response elements in the aldolase A, enolase 1, and lactate dehydrogenase A gene promoters contain essential binding sites for hypoxia-inducible factor 1. *J Biol Chem*. 1996;271(51):32529–32537.
3. He P, Naka T, Serada S, et al. Proteomics-based identification of alpha-enolase as a tumor antigen in non-small lung cancer. *Cancer Sci*. 2007;98(8):1234–1240.
4. Wygrecka M, Marsh LM, Morty RE, et al. Enolase-1 promotes plasminogen-mediated recruitment of monocytes to the acutely inflamed lung. *Blood*. 2009;113(22):5588–5598.
5. Kim RY, Lietman T, Piatigorsky J, Wistow GJ. Structure and expression of the duck alpha-enolase/tau-crystallin-encoding gene. *Gene*. 1991;103(2):193–200.
6. Wang W, Wang L, Endoh A, Hummelke G, Hawks CL, Hornsby PJ. Identification of alpha-enolase as a nuclear DNA-binding protein in the zona fasciculata but not the zona reticularis of the human adrenal cortex. *J Endocrinol*. 2005;184(1):85–94.
7. Keller A, Peltzer J, Carpentier G, et al. Interactions of enolase isoforms with tubulin and microtubules during myogenesis. *Biochim Biophys Acta*. 2007;1770(6):919–926.
8. McAleese SM, Dunbar B, Fothergill JE, Hinks LJ, Day IN. Complete amino acid sequence of the neurone-specific gamma isozyme of enolase (NSE) from human brain and comparison with the non-neuronal alpha form (NNE). *Eur J Biochem*. 1988;178(2):413–417.
9. Oliva D, Barba G, Barbieri G, Giallongo A, Feo S. Cloning, expression and sequence homologies of cDNA for human gamma enolase. *Gene*. 1989;79(2):355–360.
10. Feo S, Oliva D, Barbieri G, Xu WM, Fried M, Giallongo A. The gene for the muscle-specific enolase is on the short arm of human chromosome 17. *Genomics*. 1990;6(1):192–194.
11. Lander ES, Linton LM, Birren B, et al; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860–921.
12. Deloukas P, Earthrowl ME, Grafham DV, et al. The DNA sequence and comparative analysis of human chromosome 10. *Nature*. 2004;429(6990):375–381.
13. Dolnick BJ. Cloning and characterization of a naturally occurring antisense RNA to human thymidylate synthase mRNA. *Nucleic Acids Res*. 1993;21(8):1747–1752.
14. Liang P, Nair JR, Song L, McGuire JJ, Dolnick BJ. Comparative genomic analysis reveals a novel mitochondrial isoform of human rTS protein and unusual phylogenetic distribution of the *rTS* gene. *BMC Genomics*. 2005;6:125.
15. Verma M, Dutta SK. DNA sequences encoding enolase are remarkably conserved from yeast to mammals. *Life Sci*. 1994;55(12):893–899.
16. Sakimura K, Kushiya E, Ohshima-Ichimura Y, Mitsui H, Takahashi Y. Structure and expression of rat muscle-specific enolase gene. *FEBS Lett*. 1990;277(1–2):78–82.

17. Keller A, Bérod A, Dussaillant M, Lamandé N, Gros F, Lucas M. Coexpression of alpha and gamma enolase genes in neurons of adult rat brain. *J Neurosci Res*. 1994;38(5):493–504.
18. Schmechel DE, Brightman MW, Marangos PJ. Neurons switch from non-neuronal enolase to neuron-specific enolase during differentiation. *Brain Res*. 1980;190(1):195–214.
19. Tanaka M, Sugisaki K, Nakashima K. Purification, characterization, and distribution of enolase isozymes in chicken. *J Biochem*. 1985;98(6):1527–1534.
20. Tracy MR, Hedges SB. Evolutionary history of the enolase gene family. *Gene*. 2000;259(1–2):129–138.
21. Duquerroy S, Camus C, Janin J. X-ray structure and catalytic mechanism of lobster enolase. *Biochemistry*. 1995;34(39):12513–12523.
22. Lebioda L, Stec B, Brewer JM. The structure of yeast enolase at 2.25-Å resolution. An 8-fold beta + alpha-barrel with a novel beta beta alpha alpha (beta alpha)6 topology. *J Biol Chem*. 1989;264(7): 3685–3693.
23. Reed GH, Poyner RR, Larsen TM, Wedekind JE, Rayment I. Structural and mechanistic studies of enolase. *Curr Opin Struct Biol*. 1996;6(6):736–743.
24. Kornblatt MJ, Zheng SX, Lamandé N, Lazar M. Cloning, expression and mutagenesis of a subunit contact of rabbit muscle-specific (betabeta) enolase. *Biochim Biophys Acta*. 2002;1597(2):311–319.
25. Chai G, Brewer JM, Lovelace LL, Aoki T, Minor W, Lebioda L. Expression, purification and the 1.8 angstroms resolution crystal structure of human neuron specific enolase. *J Mol Biol*. 2004;341(4):1015–1021.
26. Wedekind JE, Poyner RR, Reed GH, Rayment I. Chelation of serine 39 to Mg^{2+} latches a gate at the active site of enolase: structure of the bis(Mg^{2+}) complex of yeast enolase and the intermediate analog phosphonoacetohydroxamate at 2.1-Å resolution. *Biochemistry*. 1994;33(31):9333–9342.
27. Poyner RR, Larsen TM, Wong SW, Reed GH. Functional and structural changes due to a serine to alanine mutation in the active-site flap of enolase. *Arch Biochem Biophys*. 2002;401(2):155–163.
28. Schreier B, Höcker B. Engineering the enolase magnesium II binding site: implications for its evolution. *Biochemistry*. 2010;49(35):7582–7589.
29. Giallongo A, Oliva D, Calì L, Barba G, Barbieri G, Feo S. Structure of the human gene for alpha-enolase. *Eur J Biochem*. 1990;190(3):567–573.
30. Peshavaria M, Day IN. Molecular structure of the human muscle-specific enolase gene (ENO3). *Biochem J*. 1991;275(Pt 2):427–433.
31. Thierry-Mieg D, Thierry-Mieg J. AceView: a comprehensive cDNA-supported gene and transcripts annotation. *Genome Biol*. 2006;7 Suppl 1: S12.1–S12.14.
32. Feo S, Arcuri D, Piddini E, Passantino R, Giallongo A. ENO1 gene product binds to the c-myc promoter and acts as a transcriptional repressor: relationship with Myc promoter-binding protein 1 (MBP-1). *FEBS Lett*. 2000;473(1):47–52.
33. Lopez-Alemany R, Suelves M, Diaz-Ramos A, Vidal B, Munoz-Canoves P. Alpha-enolase plasminogen receptor in myogenesis. *Front Biosci*. 2005;10:30–36.
34. Perconti G, Ferro A, Amato F, et al. The Kelch protein NS1-BP interacts with alpha-enolase/MBP-1 and is involved in c-Myc gene transcriptional control. *Biochim Biophys Acta*. 2007;1773(12):1774–1785.
35. Altschul SF, Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403–410.
36. Gibbs RA, Rogers J, Katze MG, et al; Rhesus Macaque Genome Sequencing and Analysis Consortium. Evolutionary and biomedical insights from the rhesus macaque genome. *Science*. 2007;316(5822):222–234.
37. Bovine Genome Project. 2008. Available from: http://www.hgsc.bcm.tmc.edu/projects/bovine. Accessed January 6, 2011.
38. Horse Genome Project. 2008. Available from: http://www.uky.edu/Ag/Horsemap/. Accessed January 6, 2011.
39. Waterston RH, Lindblad-Toh K, Birney E, et al; Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature*. 2002;420(6915):520–562.
40. Gibbs RA, Weinstock GM, Metzker ML, et al; Rat Genome Sequencing Project Consortium. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature*. 2004;428(6982): 493–521.
41. Mikkelsen TS, Wakefield MJ, Aken B, et al. Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. *Nature*. 2007;447(7141):167–177.
42. International Chicken Genome Sequencing Consortium. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*. 2004;432(7018):695–716.
43. Hellsten U, Harland RM, Gilchrist MJ, et al. The genome of the western clawed frog Xenopus tropicalis. *Science*. 2010;328(5978):633–636.
44. Sprague J, Bayraktaroglu L, Clements D, et al. The Zebrafish Information Network: the zebrafish model organism database. *Nucleic Acids Res*. 2005;34(Database issue):D581–D585.
45. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res*. 2002;12(6):996–1006.
46. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics*. 2000;16(4):404–405.
47. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*. 1997; 18(15):2714–2723.
48. Kopp J, Schwede T. The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models. *Nucleic Acids Res*. 2004;32(Database issue):D230–D234.
49. Kang HJ, Jung SK, Kim SJ, Chung SJ. Structure of human alpha-enolase (hENO1), a multifunctional glycolytic enzyme. *Acta Crystallogr D Biol Crystallogr*. 2008;64(Pt 6):651–657.
50. Qin J, Chai G, Brewer JM, Lovelace LL, Lebioda L. Fluoride inhibition of enolase: crystal structure and thermodynamics. *Biochemistry*. 2006;45(3):793–800.
51. Chandran V, Luisi BF. Recognition of enolase in the *Escherichia coli* RNA degradosome. *J Mol Biol*. 2006;358(1):8–15.
52. Larkin MA, Blackshields G, Brown NP, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007;23(21):2947–2948.
53. Giallongo A, Feo S, Moore R, Croce CM, Showe LC. Molecular cloning and nucleotide sequence of a full-length cDNA for human alpha enolase. *Proc Natl Acad Sci U S A*. 1986;83(18):6741–6745.
54. Calì L, Feo S, Oliva D, Giallongo A. Nucleotide sequence of a cDNA encoding the human muscle-specific enolase (MSE). *Nucleic Acids Res*. 1990;18(7):1893.
55. Vollmar M, Krysztofinska E, Chaikuad A, et al. Crystal structure of human beta enolase ENOB. Protein Data Bank Entry 2XSX. 2010; unpublished.
56. Kim SC, Sprung R, Chen Y, et al. Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol Cell*. 2006; 23(4):607–618.
57. Zhou W, Capello M, Fredolini C, et al. Mass spectrometry analysis of the post-translational modifications of alpha-enolase from pancreatic ductal adenocarcinoma cells. *J Proteome Res*. 2010;9(6):2929–2936.
58. Ballif BA, Carey GR, Sunyaev SR, Gygi SP. Large-scale identification and evolution indexing of tyrosine phosphorylation sites from murine brain. *J Proteome Res*. 2008;7(1):311–318.
59. Lamandé N, Mazo AM, Lucas M, et al. Murine muscle-specific enolase: cDNA cloning, sequence, and developmental expression. *Proc Natl Acad Sci U S A*. 1989;86(12):4445–4449.
60. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci U S A*. 2006;103(5):1412–1417.
61. Safran M, Kaelin WG Jr. HIF hydroxylation and the mammalian oxygen-sensing pathway. *J Clin Invest*. 2003;111(6):779–783.
62. Sousa LP, Brasil BS, Silva Bde M, et al. Characterization of alpha-enolase as an interferon-alpha 2 alpha 1 regulated gene. *Front Biosci*. 2005;10:2534–2547.
63. Taylor JM, Davies JD, Peterson CA. Regulation of the myoblast-specific expression of the human beta-enolase gene. *J Biol Chem*. 1995;270(6): 2535–2540.
64. Feo S, Antona V, Barbieri G, Passantino R, Calì L, Giallongo A. Transcription of the human beta enolase gene (ENO-3) is regulated by an intronic muscle-specific enhancer that binds myocyte-specific enhancer factor 2 proteins and ubiquitous G-rich-box binding factors. *Mol Cell Biol*. 1995;15(11):5991–6002.

65. Su AI, Wiltshire T, Batalov S, et al. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci U S A*. 2004;101(16):6062–6067.
66. Keller A, Ott MO, Lamandé N, et al. Activation of the gene encoding the glycolytic enzyme beta-enolase during early myogenesis precedes an increased expression during fetal muscle development. *Mech Dev*. 1992;38(1):41–54.
67. Gerhard DS, Wagner L, Feingold EA, et al; MGC Project Team. The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome Res*. 2004;14(10B):2121–2127.
68. Merkulova T, Lucas M, Jabet C, et al. Biochemical characterization of the mouse muscle-specific enolase: developmental changes in electrophoretic variants and selective binding to other proteins. *Biochem J*. 1997;323(Pt 3):791–800.
69. Keller A, Demeurie J, Merkulova T, et al. Fibre-type distribution and subcellular localisation of alpha and beta enolase in mouse striated muscle. *Biol Cell*. 2000;92(7):527–535.
70. Comi GP, Fortunato F, Lucchiari S, et al. Beta-enolase deficiency, a new metabolic myopathy of distal glycolysis. *Ann Neurol*. 2001;50(2): 202–207.
71. Mitchell BF, Pedersen LB, Feely M, Rosenbaum JL, Mitchell DR. ATP production in *Chlamydomonas reinhardtii* flagella by glycolytic enzymes. *Mol Biol Cell*. 2005;16(10):4509–4518.
72. Morita T, Kawamoto H, Mizota T, Inada T, Aiba H. Enolase in the RNA degradosome plays a crucial role in the rapid decay of glucose transporter mRNA in the response to phosphosugar stress in *Escherichia coli*. *Mol Microbiol*. 2004;54(4):1063–1075.
73. Donoghue PC, Benton MJ. Rocks and clocks: calibrating the tree of life using fossils and molecules. *Trends Ecol Evol*. 2007;22(8):424–431.
74. Zhao S, Choy BS, Kornblatt MJ. Effects of the G376E and G157D mutations on the stability of yeast enolase – a model for human muscle enolase deficiency. *FEBS J*. 2008;275(1):97–106.
75. Hyo JK, Jung S-K, Kim SJ, Chung SJ (2008) Structure of human alpha-enolase (hENO1), a multifunctional glycolytic enzyme. *Acta Crystall D.* 64:651–657.