Published in final edited form as:
Comp Biochem Physiol Part D Genomics Proteomics. 2008 September ; 3(3): 195–204. doi:10.1016/j.cbd.2008.05.002.

Mammalian Carboxylesterase 5: Comparative Biochemistry and Genomics

Roger S Holmes[1,2,3,4], **Laura A Cox**[1,2], and **John L VandeBerg**[1,2]

[1]Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX, USA

[2]Southwest National Primate Research Center, Southwest Foundation for Biomedical Research, San Antonio, TX, USA

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

Abstract

C*arboxylesterase 5* (*CES5*) (also called cauxin or *CES7*) is one of at least five mammalian CES gene families encoding enzymes of broad substrate specificity and catalysing hydrolytic and transesterification reactions. *In silico* methods were used to predict the amino acid sequences, secondary structures and gene locations for *CES5* genes and gene products. Amino acid sequence alignments of mammalian CES5 enzymes enabled identification of key CES sequences previously reported for human CES1, as well as other sequences that are specific to the *CES5* gene family, which were consistent with being monomeric in subunit structure and available for secretion into body fluids. Predicted secondary structures for mammalian CES5 demonstrated significant conservation with human CES1 as well as distinctive mammalian CES5 like structures. Mammalian *CES5* genes are located in tandem with the *CES1* gene(s), are transcribed on the reverse strand and contained 13 exons. CES5 has been previously reported in high concentrations in the urine (cauxin) of adult male cats, and within a protein complex of mammalian male epididymal fluids. Roles for CES5 may include regulating urinary levels of male cat pheromones; catalysing lipid transfer reactions within mammalian male reproductive fluids; and protecting neural tissue from drugs and xenobiotics.

Keywords

Mammals; amino acid sequence; genomics; carboxylesterase; CES5; drug detoxification

Background

Carboxylesterases (CES; E.C.3.1.1.1) catalyse hydrolytic and transesterification reactions using a broad range of substrates, detoxify organophosphates, carbamate compounds and insecticides (Ahmad & Forgash, 1976; Leinweber, 1987; Satoh and Hosokawa, 1998; Satoh et al., 2002; Redinbo and Potter, 2005), catalyse several cholesterol and fatty acid metabolic reactions (Tsujita and Okuda, 1993; Becker et al., 1994; Diczfalusy et al., 2001; Dolinsky et

4Corresponding Author: Roger S Holmes, D.Sc., Department of Genetics, Southwest National Primate Research Center, Southwest Foundation for Biomedical Research, San Antonio, TX, USA 78227, Email: rholmes@sfbrgenetics.org, Phone: 210-258-9687, Fax: 210-258-9600.

al, 2001), and facilitate the conversion of lung alveolar surfactant (Krishnasamy et al., 1998; Ruppert et al., 2006). CES is also involved in the biotransformation of many drugs and prodrugs (Ahmad et al., 1999; He et al., 1995; Imai et al., 2003; Ohtsuka et al., 2003; Imai, 2006; Mutch et al., 2007), anti-tumor drugs (Humerickhouse et al., 2000; Xu et al., 2002; Tabata et al., 2004) and narcotics, such as cocaine and heroine (Pindel et al., 1997), and has been linked with the assembly of very-low density lipoproteins in liver (Wang et al., 2007).

At least five families of mammalian CES have been described which show 39–45% sequence identities (Holmes et al., 2008): CES1 and CES2, the major liver and intestinal enzymes, respectively, which are both widely distributed in mammalian tissues (Shibata et al., 1993; Schewer et al., 1997; Pindel et al., 1997; Ghosh, 2000); CES3, expressed in liver, colon and brain (Sanghani etal., 1997); CES6, a predicted CES-like enzyme in brain (Clark et al., 2003); and CES5, a major urinary protein of the domestic cat (Miyazaki et al., 2003). The latter enzyme is also called cauxin (for carboxylesterase-like urinary excreted protein) or CES7 and is present in high levels in the urine of male domestic cats (*Felis catus*), bobcat (*Lynx rufus*) and lynx (*Lynx lynx*) (Miyazaki etal., 2003; 2006a) and also found in mammalian male reproductive fluids (Ecroyd etal., 2006). Domestic cat CES5 apparently functions in the production of felinine, a putative pheromone precursor, which has been proposed to play a role in attracting females and as a territorial marker (Miyazaki et al., 2006b), whereas mammalian fluid CES5 may assist with lipid transport within high-molecular weight complexes in male reproductive tracts (Ecroyd etal., 2006).

This review summarizes current knowledge about mammalian CES5 and describes the predicted amino acid sequences and secondary structures and the structural, phylogenetic and evolutionary relationships for this CES gene family. The paper also reviews the major roles for mammalian CES1, CES2 and CES5 family members, and discusses potential roles for CES5 as a body fluid carboxylesterase and in contributing to a protective role from drugs and xenobiotics for the brain.

In silico mammalian *CES5* gene and protein identification

BLAST (Basic Local Alignment Search Tool) *in silico* studies were undertaken using the National Center for Biotechnology Information (NCBI) web site [http://www.ncbi.nlm.nih.gov/blast/Blast.cgi] for nucleotide and protein blasts using known CES sequences (Table 1) (Altschul et al., 1990). BLAT (BLAST-Like Alignment Tool) *in silico* studies were also undertaken using the UC Santa Cruz Genome Browser [http://genome.ucsc.edu/cgi-bin/hgBlat] (Kent *et al*. 2002) with the default settings. Genomes examined included: human (*Homo sapiens*) (International Genome Sequencing Consortium, 2004; chimpanzee (*Pan troglodytes*); (The Chimpanzee Sequencing and Analysis Consortium, 2005); orangutan (*Pongo abelii*) (Orangutan Genome Sequencing project, 2008); rhesus monkey (*Macaca mulatta*) (Rhesus Macaque Genome Sequencing and Analysis Consortium, 2007); mouse (*Mus musculus*) (Mouse Sequencing Consortium, 2002); rat (*Rattus norvegicus*) (Rat genome Sequencing project Consortium, 2004); cat (*Felis catus*) (Pontius et al., 2007); dog (*Canis familiaris*) (Dog Genome Project., 2005); cow (*Bos taurus*) (Bovine Genome Project, 2008); and horse (*Equus caballus*) (Horse Genome project, 2008). UniProtKB/Swiss-Prot Database [http://au.expasy.org] and GenBank [http://www.ncbi.nlm.nih.gov/Genbank/] sequences for human, mouse (Miyazaki et al., 2006a), rat, cat (Miyazaki et al., 2003), dog and sheep (Ecroyd et al., 2006) CES5 were used to interrogate known mammalian genome sequences (Table 1). Gene locations, predicted gene structures and CES protein subunits were observed for each CES sequence examined for those regions showing identity with the respective mammalian *CES5* gene products.

Predicted Secondary Structures for Mammalian CES5

Predicted secondary structures for human CES2 and various mammalian CES5 subunits were obtained using the PSIPRED v2.5 web site tools [http://bioinf.cs.ucl.ac.uk/psipred/psiform.html] (McGuffin *et al.*, 2000).

Phylogenetic Studies and Sequence Divergence

Phylogenetic trees were constructed using an amino acid alignment from a ClustalW-derived alignment of CES protein sequences, obtained with default settings and corrected for multiple substitutions (Chenna *et al* 2003) [http://www.ebi.ac.uk/clustalw/]. Alignment ambiguous regions, including the amino and carboxyl termini, were excluded prior to phylogenetic analysis yielding an alignment of 526 residues. Amino acid sequence divergences for mammalian CES5 were determined using the SIM-Alignment tool for Protein Sequences [http://au.expasy.org/tools/sim-prot.html] (Pietsch 1995; Schwede *et al.* 2003).

Alignments of human CES1, CES2 and CES5 amino acid sequences

The amino acid sequences for human CES1 (Shibata et al., 1993; Becker et al., 1994; Ghosh, 2000) and CES2 (Schewer et al., 1997; Pindel et al., 1997) and the predicted sequence for human CES5 (International Human Genome Sequencing Consortium, 2004) are shown in Figure 1 (see Table 1). Alignments of human CES1 with human CES2 and CES5 showed 45% and 42% sequence identities respectively, while the alignment of human CES2 with CES5 showed 43% sequence identity (Table 2). This comparison of human CES1, CES2 and CES5 subunit sequences suggests that they are products of separate families of CES genes, which supports earlier proposals (Satoh and Hosokawa, 2006; Holmes et al., 2008).

The predicted amino acid sequence for human CES5 was eight residues longer (575 amino acids) than for human CES1 (567), and 16 residues longer than for human CES2 (559 amino acids) (Figure 1). Comparisons of key residues are of particular interest, which have been previously shown to contribute to the catalytic, subcellular localization, oligomeric and regulatory functions for human CES1 (sequence numbers refer to human CES1). These include the active site catalytic triad (Ser228; Glu345; His458) (Cygler et al., 1993); microsomal targeting sequences, including the hydrophobic N-terminus signal peptide (von Heijne 1983;Zhen et al., 1995;Potter et al., 1998) and the C-terminal endoplasmic reticulum (ER) retention sequence (His-Ile-Glu-Leu) (Robbi & Beaufay 1983; Munro & Pelham 1987;Zhen et al., 1995); disulfide bond forming residues (Cys95/Cys123 and Cys280/Cys291) (Lockridge et al., 1987); and ligand binding sites, including the ‘Z-site’ (Gly358), the ‘side door’ (Val424-Met425-Phe426) and the ‘gate’ (Phe551) residues (Redinbo and Potter, 2005). Other CES class 1 specific residues included the ‘charge clamp’ residues contributing to the oligomeric subunit structures for human CES1 (Lys78..Glu183 and Glu72 ..Arg193), and the N-glycosylation site at Asn79-Ala80-Thr81 (Ozols, 1989;Kroetz et al, 1993;Bencharit et al., 2003;2006;Fleming et al., 2005). The predicted human CES5 sequence contained essential catalytic and structural residues, such as the active site triad, the disulfide bridges and the ‘Z-site’, and exhibited conservative changes in sequences for other key residues, such as the ‘side door’ and ‘gate’, which are proposed to facilitate product (acyl group) release from the CES active site (Redinbo and Potter, 2005).

Human CES5 showed a distinct predicted N-terminal ‘signal peptide’ sequence to that reported for human CES1 and CES2, and lacked relevant charge clamp residues, which contribute to the human CES1 oligomeric structure (trimer-hexamer) (Fleming et al., 2005). Human and baboon CES2 have been shown to be monomeric enzymes (Pindel et al., 1997; RS Holmes, J Glenn, JL Vandeberg and LA Cox, Baboon carboxylesterases 1 and 2: sequences, structures and phylogenetic relationships with human and other primate carboxylesterases] and it is likely

that human CES5 is also monomeric due to the absence of charge clamps reported for human CES1 (Fleming et al, 2005). The predicted C-terminus for human CES5 lacked the reported endoplasmic reticulum retention sequence human CES1 and CES2 (HIEL and HTEL, respectively), which may contribute to this enzyme being secreted, rather than being localized within the endoplasmic reticulum (Clark et al., 2003). Four potential N-glycosylation sites were observed for human CES5 (Table 3) which may assist in stabilizing the enzyme in support of its location in body fluids such as urine (observed in cats) (Miyazaki et al., 2006a) and in male reproductive fluids (observed in sheep) (Ecroyd et al., 2005). N-glycosylation at the single Asn79-Ala80-Thr81 human CES1 site has been previously reported to contribute to the stability of the enzyme and in maintaining maximal catalytic for this enzyme (Kroetz et al., 1993; Fleming et al., 2005), and similar roles for the multiple N-glycosylation human CES5 sites are proposed.

Comparative Mammalian CES Genomics

Structures for several mammalian *CES* genes have been determined, including human (Becker et al., 1994; Langmann et al., 1997; Pindel et al., 1997; Ghosh 2000; Marsh et al., 2004) and rodent *CES1* and *CES2* 'like' genes (Ghosh et al., 1995; Dolinsky et al., 2001; Hosokawa et al., 2007; Furihata et al., 2003; 2005; 2006). Six human *CES* genes have been described on chromosome 16 (Figure 2). The human *CES1* gene, encoding the major liver CES but also expressed in lung, tumours, oesophagus, breast, uterus and most other tissues of the body (Shibata etal., 1993; Becker et al., 1994), is defined by > 600 GenBank entries (Thierry-Mieg and Thierry-Mieg,, 2006) and maps at 16q13-q22.1 on the reverse strand (Table 1). The *CES1* gene is transcribed into 5 alternatively spliced mRNAs, of which 3 are derived from 14 exons, encoding CES protein subunits that appear to be fully functional, forming CES1a, CES1b and CES1c isoforms (Figure 3). The human *CES2* gene encodes the major intestinal CES (Schewer etal., 1997; Pindel etal., 1997) and is also expressed in most other tissues of the body, including liver, colon, kidney, brain and lung (Thierry-Mieg and Thierry-Mieg, 2006). The *CES2* gene sequence is defined by > 500 GenBank entries and maps at 16q22.1 on the positive strand (Table 1). *CES2* transcription produces 11 alternatively spliced mRNAs, of which at least 3 are derived from 12 exons and form apparent fully functional CES subunits, namely CES2a, CES2b and CES2c (Figure 2).

The human *CES5* gene sequence is currently defined by 17 GenBank entries and encodes at least 4 alternatively spliced mRNAs which are derived from several tissues of the body, including brain, testis and lung (International Human Genome Sequencing Consortium, 2004; Thierry-Mieg and Thierry-Mieg, 2006) (Figure 3). Human *CES5* maps on chromosome 16 at 16q12.2 in a tandem location to human CES1 and encodes 4 isoforms which appear to be fully functional, derived from 13 exons: CES5a, CES5b, CES5c and CES5d. Figure 2 describes the gene structures for human CES1, CES2 and CES5, based on the 5' to 3' alignments of multiple mRNAs respectively on the human genome and obtained using AceView (Thierry-Mieg and Thierry-Mieg, 2006). These genes are transcribed into several multiple splice variants in each case, although only those isoforms with capped 5' ends and validated 3' ends are shown: CES1 encodes at least 3 such isoforms with 14 exons in each case; human CES2 encodes at least 3 mature isoforms although these are derived from 12 exons; while human CES5 encodes at least 4 mature isoforms with the CES5a isoform being derived from 13 exons. Human CES1 and CES5 are encoded on the minus strand whereas human CES2 is encoded on the plus strand.

Figure 1 shows the locations of the intron-exon boundaries for human CES1, CES2 and CES5 gene products and their positioning within the aligned amino acid sequences. Exon 1 corresponds to the encoded signal peptide in each case, with the last exon encoding the endoplasmic reticulum targeting sequence (for human CES1 and CES2) or the C-terminal

sequence involved in a proposed role in the secretion of human CES5. There is identity or near identity for the intron-exon boundaries for each of the human *CES1*, *CES2* and *CES5* genes, with the exception of an additional human CES1 exon boundary (forming exon 9), and the absence of an intron-exon boundary for human CES2 within exon 5.

At least 3 additional human CES genes have been described which are also located on chromosome 16 within two segments of CES gene clusters (Figure 2; Table 1). The human *CES3* gene sequence is currently defined by 65 GenBank entries which span 13.93 kb on chromosome 16 downstream from CES2 (Sanghani etal., 2004;Thierry-Mieg and Thierry-Mieg, 2006); the human *CES4* gene (Yan etal., 1999) is complex in nature, being transcribed into 4 alternatively spliced mRNAs, which may form multiple CES proteins with no sequence overlap and located near the *CES1* gene on chromosome 16 (Thierry-Mieg and Thierry-Mieg, 2006); whereas the human *CES6* gene (also called *FLJ37464*) sequence (Clark etal., 2003) is defined by 86 GenBank entries which are derived from mRNAs isolated from >40 tissues of the body and is located downstream of CES2 and CES3 on chromosome 16 (Table 1;Figure 2) (Thierry-Mieg and Thierry-Mieg, 2006).

Secondary Structures for Human CES1, CES2 and CES5

Figure 1 shows the comparative secondary structures previously reported for human CES1 (Bencharit et al., 2003;2006;Fleming et al, 2005) or predicted for human CES2 and CES5. Similar α-helix β-sheet structures were observed for the three human CES gene products examined, which was readily apparent near key residues or functional domains, including the α-helix within the N-terminal signal peptide and the β-sheet and α-helix structures surrounding the active site Ser228 (human CES1). In addition, a large random coil region (residues 55–117 for human CES1) was predominantly retained within human CES2 and CES5. Other sites however showed differences in predicted secondary structures for human CES5, in comparison with the reported structure for human CES1, including the predicted ‘Z-site’ for CES5 (human CES1 Gly356), which has a larger helix involving both this site and the proximate active site residue (human CES1 Glu354); the predicted ‘side door’ for CES5 (human CES1 Val424-Met425-Phe426), for which a large helix (human CES1 residues 410–439) was divided into 2 smaller predicted helices for human CES5; and an additional α-helix which was observed at the extended C-terminus for human CES5. It should be recognized however that predictions of CES secondary structure may not fully reflect CES structures *in vivo* and may serve only as a guide as to the comparative structures for CES5 and other CES subunits.

Comparative Amino Acid Sequences for Mammalian CES5. A Stop Codon within the Cat CES5 Gene

Figure 4 examines the comparative amino acid sequences for CES5 from 11 mammalian species. These CES5 sequences were complete with the exception of sheep epididymus CES5 (Ecroyd et al., 2005) which aligned with residues 118–498 of human CES5; and cat CES5, reported by Miyazaki et al (2003) and based on a cat kidney RT-PCR cDNA clone, which had a sequence aligning with human CES5 residues 1–544. The cat CES5 sequence, which lacked 30 residues at the C-terminus end, has apparently arisen from a gene mutation which introduced a stop codon (TGA) at a position encoding residue 545 (normally Ile545) for other mammalian CES5 sequences. A comparison of dog and cat CES5 exon 13 DNA sequences (Figure 5) indicated that a GT insertion at this position in the cat genome created a stop codon that resulted in a shortened (544 residues) cat CES5. The reduction in length for the sheep CES5 sequence may have arisen from either posttranslational modification of the male reproductive secreted enzyme; from the translation of a shorter spliced isoform CES5 mRNA isoform; or from gene sequence changes for the sheep CES5 gene.

The following mammalian CES5 residues were consistent with functions and aligned positions to those described for human CES1 (Redinbo and Potter, 2005; Bencharit et al., 2003; 2006): a 26 residue N-terminus hydrophobic signal peptide containing a large α-helix; a ten residue (amino acids 56–65) trans-membrane sequence; four cysteine residues forming disulfide bonds (Cys94; Cys102; Cys278; Cys291); CES5 active site triad residues : 226Ser; 345Glu; and 437His; a predicted 'Z-site' for CES7 at 347Gly, which may play a role in cholesterol binding; the predicted CES5 'side door' at residues 410Val-411Phe-412Phe; and a predicted human CES5 active site 'gate' at 543Ile, which may play roles in facilitating the release of acyl groups following hydrolysis or transesterification.

Mammalian CES5 sequences differed in the number and distribution of potential N-glycosylation sites which may play a role in enhancing the stabilities and catalytic activities for these enzymes (Table 3). Seven such sites were observed although only three of these were found in most CES5 sequences examined, including Site3: 281Asn-282Ala/Val/Ser-282Ser; Site 4: 363Asn-364Lys/Glu-365Ser; and Site 7: 522Asn-523Met/Ile/Val-524Ser. Most mammalian CES5 sequences contained 3 or 4 such sites which compares with the single N-glycosylation site reported for human and baboon CES1 and CES2 (Kroetz et al.,Pindel et al, 1997; RS Holmes, J Glenn, JL VandeBerg and LA Cox, Baboon carboxylesterases 1 and 2: sequences, structures, and phylogenetic relationships with human and other primate carboxylesterases, unpublished). This additional capacity for N-glycosylation and binding of carbohydrate residues may contribute significantly to increasing the stability and activity of these enzymes, particularly in light of their distribution in fluids of the body.

Predicted secondary structures for the mammalian CES5 sequences examined showed similar patterns of α-helices and β-strands, although some species differences were observed, including an additional helix (residues 242–251) for orangutan and horse CES5; the absence of predicted helices following the 345Glu active site residue for mouse and rat CES5 (residues 348–354) and rat, human and chimp CES5 (residues 356–359), respectively; and the lack of a predicted C-terminal helix for rat CES5.

Phylogeny of Human CES Gene Products and of Mammalian CES5

Phylogenetic and sequence alignments studies on vertebrate and invertebrate CES have reported evidence for at least five distinct CES gene family clusters, namely *CES1*, *CES2*, *CES3*, *CES5* and *CES6*, arising from rapid gene duplication events between 328–378 million years ago (Holmes etal., 2008). In addition, further gene duplication events have apparently occurred during mammalian evolution among common ancestors for rodents, generating at least four CES1 and CES2 like genes in each case; and the opossum, where three CES2 like genes have been described (RS Holmes, J Glenn, JL VandeBerg & LA Cox: Baboon carboxylesterases 1 and 2: sequences, structures and phylogenetic relationships with human and other primate carboxylesterases, unpublished; Holmes et al., 2008).

A phylogenetic tree (Figure 6) was calculated by the progressive alignment of five human CES amino acid sequences and eleven mammalian CES5 sequences which shows clustering into five main groups (or families) for the CES1, CES2, CES3, CES5 and CES6 'like' gene products. The eleven mammalian CES5 sequences clustered together within a single group which supports the proposal that these form part of a single enzyme class. Table 2 summarizes the percentages of identity for these enzymes and shows that mammalian CES5 sequences are > 66% identical which contrasts with the 42–45% identities observed for human CES1, CES2 and CES5. In addition, more closely related species show higher levels of sequence identity for CES5, including > 95% identical for primate species, human, chimpanzee, orangutan and rhesus monkey. Holmes et al. (2008) have recently reported evidence for a rapid early diversification into at least five CES gene family clusters, namely CES1; CES2; CES3; CES5;

and CES6, and reported an estimated time for the origin of these CES gene duplications at around 328–378 million years ago. Based on this report, we have concluded that the CES7 primordial gene predates the eutherian mammalian common ancestor (estimated at around 105 million years ago) (Murphy et al., 2001;Woodburne et al., 2003) by > 200 million years. In addition, it is also apparent that the mammalian CES5 gene has been derived from a common ancestor shared by all eutherian mammals ~100 MY ago.

CES5 Functional Aspects

Mammalian carboxylesterase 5 (*CES5*) (also called cauxin and CES7) is a distinct CES gene family and a member of the *CES* super-family of genes encoding enzymes of broad substrate specificity and responsible for the detoxification and metabolism of a wide range of carboxylesters, thioesters and aromatic amides, including xenobiotics, narcotics and clinical drugs, and which have the capacity to catalyze a number of cholesterol and lipid metabolic reactions (Becker et al., 1994; Satoh and Mizakawa, 1998; Satoh et al., 2002; Redinbo and Potter, 2005; Tsujita and Okuda, 1993; Diczfalusy et al., 2001; Dolinsky et al, 2001). More specific roles for mammalian CES have been described, including the activation of lung surfactant (Krishnasamy et al., 1998); the detoxification of organophosphate and carbamate poisons (Jakonivic et al., 1996; Satoh and Hosokawa, 1998); the activation of several prodrugs used in treating various diseases such as influenza (He et al., 1999), cancer (Humerickhouse et al., 2000; Ohtsuka et al., 2003; Tabata et al., 2004) and high blood pressure (Takai et al., 1997); providing additional metabolic capacity as lipid hydrolases for neutral fats and phopholipids in tissues of the body (Mentlein et al., 1988; Ghosh et al., 1995; Tsujita and Okuda, 1993; Okazaki et al., 2006); and contributing to glucocorticosteroid drug activation within bronchial cells following prodrug inhalation (Mutch et al., 2007).

The differential tissue distribution and microlocalization of CES family members may provide an important clue as to their roles within mammalian organisms. Mammalian liver is the predominant site for drug metabolism in the body and is also the major source of CES1 and CES2 (with CES1 > CES2) and is also the predominant site of drug metabolism in the body, where these enzymes play major roles in drug clearance from the body, following absorption of drugs and xenobiotics into the circulation (Pindel et al., 1997; Imai, 2006). CES1 and CES2, which are found with high activities in mammalian intestine (with CES2 > CES1), have been implicated in the first pass clearance of several drugs in rats. The predominant intestinal drug metabolism activity is located in the ileum and jejunum and processed via CES2 (Imai et al., 2006). Differential roles for CES1 and CES2 in drug metabolism have been investigated for the anti-cancer drug irinotecan (CPT-11) which is converted to its active form SN-38 predominantly by CES2 which has a 12 to 26 fold higher affinity for CPT-11 as compared with CES1 (Humerickhouse et al., 2000; Xu et al., 2002)

In contrast with mammalian CES1 and CES2, which are predominantly localized within the liver and intestine endoplasmic reticulum and are strongly membrane bound, mammalian *CES5* is predominantly expressed in peripheral tissues, including brain, kidney, lung and testis (Thierry-Mieg and Thierry-Mieg, 2006), and is a secreted form of CES enzyme due to the absence of the microsomal targeting sequence found at the carboxy-terminus (Figure 3; Figure 5). CES1 HIEL and CES2 HTEL C-terminal sequences, and similar tetrapeptide sequences, have been previously shown to direct the microlocalization of mammalian liver CES within the endoplasmic reticulum (Robbi and Beaufay, 1991). Miller and coworkers (1999) have modified this sequence for human CES1 using site-directed mutagenesis to that of HIER, substituting the C-terminal leucine with arginine, and observed that the enzyme was secreted rather than retained within the cultured cells examined. Human and other mammalian CES5 sequences normally have a long hydrophobic C-terminal sequence which apparently provides the trigger for secretion. The cat CES5 amino acid sequence lacked the elongated hydrophobic

C-terminal sequence observed for all other mammalian CES5 sequences examined and the C-terminus tetrapeptide motif observed for human CES1 and CES2 sequences.

Recent studies on mammalian CES5 have supported at least two major roles for this secreted enzyme within mammalian fluids. For example, domestic cat CES5 (also called CES7 and cauxin for carboxylesterase-like urinary excreted protein) represents a major protein in urine, particularly from adult males, being secreted from epithelial cells of kidney distal tubules where it is apparently functions in regulating the production of a pheromone precursor (Miyazaki *et al.* 2003; 2006). CES5 has also been identified in mammalian male reproductive fluids, specifically in the epididymal fluids of sheep, pigs, mice, horses and cats, where it was associated with a soluble form of the prion protein (Ecroyd et al., 2005). The physiological role for epididymyl CES5 in mammals remains to be determined however it is likely that the enzyme performs several roles given its broad substrate specificity, and may contribute significantly to lipid and cholesterol transfer processes within male reproductive fluids. With the exception of cat, mammalian CES5 sequences retained the Z-site glycine (347Gly for human CES7). This residue facilitates cholesterol analogue binding by human CES1 (Redinbo and Potter, 2005; Bencharit et al., 2003; 2006), and its retention may indicate that this property is shared by mammalian CES5 which may play a role in cholesterol ester metabolism. In addition, human CES5 ‘side door’ (410Val-411Phe-412Phe) and ‘gate’ (543Ile) residues have retained the hydrophobic property observed for the corresponding human CES1 residues, and this may reflect similar roles for mammalian CES5 in regulating acyl release following hydrolysis or transesterification (Redinbo and Potter, 2005; Bencharit et al., 2003; 2006).

CES5 has also been identified in human brain (Ota et al., 2004) and is likely to be localized in the cerebrospinal fluid or other fluids of the brain. This expression and likely distribution in neural fluids may provide a guide as to its role, such as protecting the brain and other neural tissues from drugs via the blood brain barrier or the cerebrospinal fluid. We conclude that CES5 plays a distinctly different role to that of the major liver (CES1) and intestine (CES2) enzymes, and may serve as a soluble fluid form of CES within peripheral tissues of the body, with specialized functions in drug, xenobiotic and lipid metabolism, which remain to be fully determined.

Conclusions

The results of the present study have demonstrated that mammalian CES5 from several species share key conserved sequences and structures that have been reported for human CES1, which are consistent with this family of enzymes also serving a broad role in carboxyl ester hydrolyis and transesterification reactions. Mammalian CES5 also exhibited family specific sequences which are distinct from the CES1 and CES2 gene families previously investigated (Bencharit et al., 2003; 2006; Schewer et al., 1997; Pindel et al., 1997), suggesting specialized roles in metabolism. Moreover, all of the mammalian CES5 sequences studied lacked the C-terminus sequence which ensures retention of the enzyme within the endoplasmic reticulum (HIEL and HVEL for human CES1 and CES2 respectively) enabling secretion of this enzyme into fluids of the body where it may perform specific roles in drug, pheromone and lipid metabolism.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This project was supported by NIH Grants P01 HL028972 and P51 RR013986. In addition, this investigation was conducted in facilities constructed with support from Research Facilities Improvement Program Grant Numbers 1 C06 RR13556, 1 C06 RR15456, 1 C06 RR017515.

REFERENCES

Ahmad S, Forgash AJ. Nonoxidative enzymes in the metabolism of insecticides. Drug Metab. Rev 1976;5:141–164.

Ahmed F, Vyas V, Cornfield A, Goodin S, Ravikumar TS, Rubin EH, Gupta E. In vitro activation of irinotecan to SN-38 by human liver and intestine. Anticancer Res 1999;19:2067–2071. [PubMed: 10470149]

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J. Mol. Biol 1990;215:403–410. [PubMed: 2231712]

Becker A, Bottcher A, Lackner KJ, Fehringer P, Notka F, Aslandis C, Schmitz. Purification, cloning and expression of a human enzyme with acyl coenzyme A: cholesterol acyltransferase activity, which is identical to liver carboxylesterase. Arterioscler. Thromb 1994;14:1346–1355. [PubMed: 8049197]

Bencharit S, Edwards CC, Morton CL, Howaard-Williams EL, Kuhn P, Potter PM, Redinbo MR. Multisite promiscuity in the processing of endogenous substrates by human carboxylesterase 1. J. Mol. Biol 2006;363:201–214. [PubMed: 16962139]

Bencharit S, Morton CL, Xue Y, Potter PM, Redinbo MR. Structural basis of heroin and cocaine metabolism by a promiscuous human drug-processing enzyme. Nat. Struct. Biol 2003;10:349–356. [PubMed: 12679808]

Berning W, De Looze SM, von Deimling O. Identification and development of a genetically closely linked carboxylesterase gene family of the mouse liver. Comp Biochem Physiol B 1985;30:859–865. [PubMed: 3995927]

Bovine Genome project. 2008. http://www.hgsc.bcm.tmc.edu/projects/bovine/

CES1 Gene Card. GC16M054395. 2008. http://www.genecards.org/cgi-bin/carddisp.pl?gene=CES1

CES2 Gene Card. GC16P065525. 2008. http://www.genecards.org/cgi-bin/carddisp.pl?gene=CES2

CES7 Gene Card. 2008. GC16P065525 http://www.genecards.org/cgi-bin/carddisp.pl?gene=CES7

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DJ, Thompson JD. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res 2003;31:3497–3500. [PubMed: 12824352]

Clark HF, Gurney AL, Abaya E, Baker K, Baldwin D, Brush J, Chen J, Chow B, Chui C, Crowley C, Currell B, Deuel B, Dowd P, Eaton D, Foster J, Grimaldi C, Gu Q, Hass PE, Heldens S, Huang A, Kim HS, Klimowski L, Jin Y, Johnson S, Lee J, Lewis L, Liao D, Mark M, Robbie E, Sanchez C, Schoenfeld J, Seshagiri S, Simmons L, Singh J, Smith V, Stinson J, Vagts A, Vandlen R, Watanabe C, Wieand D, Woods K, Xie MH, Yansura D, Yi S, Yu G, Yuan J, Zhang M, Zhang Z, Goddard A, Wood WI, Godowski P, Gray A. The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins:a bioinformatics assessment. Genome Res 2003;13:2265–2270. [PubMed: 12975309]

Cygler M, Schrag JD, Sussman JL, Harel M, Silman I, Gentry MK, Dostor BP. Relationship between sequence conservation and three-dimensional structure in a large family of esterases, lipases and related proteins. Protein Sci 1993;2:366–382. [PubMed: 8453375]

Diczfalusy MA, Bjorkkem I, Einarsson C, Hillebrant CG, Alexson SE. Characterization of enzymes involved in formation of ethyl esters of long-chain fatty acids. J Lipid Res 2001;42:1025–1032. [PubMed: 11441128]

Dog Genome Project. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 2005;438:803–819. [PubMed: 16341006]

Dolinsky VW, Sipione S, Lehner R, Vance DE. The cloning and expression of murine triacylglycerol hydrolase cDNA and the structure of the corresponding gene. Biochim. Biophys. Acta 2001;1532:162–172. [PubMed: 11470237]

Ecroyd H, Belghazi M, Dacheux J-L, Miyazaki M, Yamashita T, Gatti J-L. An epididymal form of cauxin, a carboxylesterase-like enzyme, is present and active in mammalian male reproductive fluids. Biol. Reprod 2006;74:439–447. [PubMed: 16251497]

Fleming CD, Bencharit S, Edwards CC, Hyatt JL, Tsurkan L, Bai F, Fraga C, Morton CL, Howard-Williams EL, Potter PM, Redinbo MR. Structural insights into drug processing by human carboxylesterase 1: tamoxifen, Mevaststin, and inhibition by Benzil. J. Mol. Biol 2005;352:165–177. [PubMed: 16081098]

Furihata T, Hosokawa M, Nakata F, Satoh T, Chiba K. Purification, molecular cloning and functional expression of inducible acylcarnitine hydrolase in C57BL/6J mouse belonging to the carboxylesterase gene family. Arch. Biochem. Biophys 2003;416:101–109. [PubMed: 12859986]

Furihata T, Hosokawa M, Fujii A, Derbel M, Satoh T, Chiba K. Dexamethosone-induced methylprednisolone hemisuccinate hydrolase: its identification as a member of the rat carboxylesterase 2 family and its unique presence in plasma. Biochem. Pharm 2005;69:1287–1297. [PubMed: 15794950]

Furihata T, Hosokawa M, Masuda M, Satoh T, Chiba K. Hepatocyte nuclear factor-4α plays pivotal roles in the regulation of mouse carboxylesterase 2 gene transcription in mouse liver. Arch. Biochem. Biophys 2006;447:107–117. [PubMed: 16527247]

Ghosh S. Cholesteryl ester hydrolase in human monocyte/macrophage: cloning, sequencing and expression of full-length cDNA. Physiol. Genomics 2000;2:1–8. [PubMed: 11015575]

Ghosh S, Mallonee DH, Hylemon PB, Grogan WM. Molecular cloning and expression of rat hepatic neutral cholesteryl ester hydrolase. Biochim. Biophys. Acta 1995;1259:305–312. [PubMed: 8541339]

He G, Massarella J, Ward P. Clinical pharmacokinetics of the prodrug oseltamivir and its active metabolite Ro 64-0802. Clin. Pharmacokinet 1999;37:471–484. [PubMed: 10628898]

Holmes RS, Chan J, Cox LA, Murphy WM, VandeBerg JL. Opossum carboxylesterases: sequences, phylogeny and evidence for CES duplication events predating the marsupial-eutherian common ancestor. BMC Evol. Biol 2008;8:54. [PubMed: 18289373]

Horse Genome Project. 2008. http://www.uky.edu/Ag/Horsemap/

Hosokawa M, Furihata T, Yaginuma Y, Yamamoto M, Koyano N, Fijii A, Nagahara Y, Satoh T, Chiba K. Genomic structure and transcriptional regulation of the rat, mouse and human carboxylesterase genes. Drug Metab. Rev 2007;39:1–15. [PubMed: 17364878]

Humerickhouse R, Lohrbach K, Li L, Bosron WF, Dolan ME. Characterization of CPT-11 hydrolysis by human liver carboxylesterase isoforms h-CE1 and hCE-2. Cancer Res 2000;60:1189–1192. [PubMed: 10728672]

Imai T, Yoshigae Y, Hosokawa M, Chiba K, Oragiri M. Evidence for the involvement of a pulmonary first-pass effect via carboxylesterase in the disposition of a propanolol ester derivative after intravenous administration. J. Pharmacol. Exp. Ther 2003;307:1234–1242. [PubMed: 14534358]

Imai T. Human carboxylesterase isozymes: catalytic propertires and rational drug design. Drug Metab. Pharmacogenet 2006;21:173–185.

International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature 2004;431:931–945. [PubMed: 15496913]

Jakonovic M, Kosanovic M, Maksimovic M. Interaction of organophosphorus compounds with carboxylesterases in the rat. Arch. Toxicol 1996;70:444–450. [PubMed: 8740539]

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. Genome Res 2002;12:994–1006.

Krishnasamy R, Teng AL, Dhand R, Schultz RM, Gross NJ. Molecular cloning, characterization and differential expression patter of mouse lung surfactant convertase. Am J Physiol 1998;275:L969–L975. [PubMed: 9815115]

Kroetz DL, McBride OW, Gonzalez FJ. Glycosylation-dependent activity of Baculovirus-expressed human liver carboxylesterases: cDNA cloning and characterization of two highly similar enzyme forms. Biochemistry 1993;32:11606–11617. [PubMed: 8218228]

Langmann T, Becker A, Aslanidis C, Notka F, Ulrich H, Schwer H, Schmitz G. Structural organization and characterization of the promoter region of a human carboxylesterase gene. Biochim. Biophys Acta 1997;1350:65–74. [PubMed: 9003459]

Leinweber FJ. Possible physiological roles of carboxyl ester hydrolases. Drug Metab. Rev 1987;18:379–439. [PubMed: 3286170]

Lockridge O, Adkins S, La Due BN. Location of disulfide bonds within the sequence of human serum cholinesterase. J. Biol. Chem 1987;262:12945–12952. [PubMed: 3115973]

Marsh S, Xiao M, Yu J, Ahluwalia R, Minton M, Freimuth RR, Kwok PY, McLeod HL. Pharmacogenomic assessment of carboxylesterases 1 and 2. Genomics 2004;84:661–668. [PubMed: 15475243]

McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics 2000;16:404–405. [PubMed: 10869041]

Mentlein R, Rix-Matzen H, Heymann E. Subcellular localization of non-specific carboxylesterases, acylcarnitine hydrolase, monoacylglycerol lipase and palmitoyl-CoA hydrolase. Biochim. Biophys. Acta 1988;964:319–328. [PubMed: 2894861]

Miller AD, Scott DF, Chacko TL, Maxwell DM, Schlager JJ, Lanclos KD. Expression and partial purification of a recombinant secretory form of human liver carboxylesterase. Protein Exp. Purif 1999;17:16–25.

Miyazaki M, Kamiie K, Soeta S, Taira H, Yamashita T. Molecular cloning and characterization of a novel carboxylesterase-like protein that is physiologically present at high concentrations in the urine of domestic cats (Felis catus). Biochem. J 2003;370:101–110. [PubMed: 12401131]

Miyazaki M, Yamashita T, Suzuki Y, Saito Y, Soeta S, Taira H, Suzuki A. A major urinary protein of the domestic cat regulates the production the production of felinine, a putative pheromone precursor. Chem. Biol 2006a;13:10171–10179.

Miyazaki M, Yamashita T, Hosokawa M, Taira H, Suzuki A. Species-, sex-, and age-dpendent urinary excretion of cauxin, a mammalian carboxylesterase. Comp. Biochem. Physiol. B 2006b;145:270–277. [PubMed: 17045831]

Mouse Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. Nature 2002;420:520–562. [PubMed: 12466850]

Munro S, Pelham HR. A C-terminal signal prevents secretion of luminal ER proteins. Cell 1987;48:899–907. [PubMed: 3545499]

Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ. Molecular phylogenetics and the origins of placental mammals. Nature 2001;409:614–618. [PubMed: 11214319]

Mutch E, Nave R, McCracken N, Zech K, Williams FM. The role of esterases in the metabolism of ciclesonide to desisobutyrl-ciclesonide in human tissue. Biochem. Pharmacol 2007;73:1657–1664. [PubMed: 17331475]

Ohtsuka K, Inoue S, Kameyama M, Kanetoshi A, Toru F, Kazuo T, Yoshikazu A, Akira S. Intracellular conversion of irinotecan to its active form, SN-38, by native carboxylesterase in human non-small cell lung cancer. Lung Cancer 2003;41:187–198. [PubMed: 12871782]

Orangutan Genome Sequencing Project. 2008. http://www.hgsc.bcm.tmc.edu/projects/orangutan/

Ota T, Suzuki Y, Nishikawa T, Otsuki T, Sugiyama T, Irie R, Wakamatsu A, Hayashi K, Sato H, Nagai K, Kimura K, Makita H, Sekine M, Obayashi M, Nishi T, Shibahara T, Tanaka T, Ishii S, Yamamoto J, Saito K, Kawai Y, Isono Y, Nakamura Y, Nagahari K, Murakami K, Yasuda T, Iwayanagi T, Wagatsuma M, Shiratori A, Sudo H, Hosoiri T, Kaku Y, Kodaira H, Kondo H, Sugawara M, Takahashi M, Kanda K, Yokoi T, Furuya T, Kikkawa E, Omura Y, Abe K, Kamihara K, Katsuta N, Sato K, Tanikawa M, Yamazaki M, Ninomiya K, Ishibashi T, Yamashita H, Murakawa K, Fujimori K, Tanai H, Kimata M, Watanabe M, Hiraoka S, Chiba Y, Ishida S, Ono Y, Takiguchi S, Watanabe S, Yosida M, Hotuta T, Kusano J, Kanehori K, Takahashi-Fujii A, Hara H, Tanase TO, Nomura Y, Togiya S, Komai F, Hara R, Takeuchi K, Arita M, Imose N, Musashino K, Yuuki H, Oshima A, Sasaki N, Aotsuka S, Yoshikawa Y, Matsunawa H, Ichihara T, Shiohata N, Sano S, Moriya S, Momiyama H, Satoh N, Takami S, Terashima Y, Suzuki O, Nakagawa S, Senoh A, Mizoguchi H, Goto Y, Shimizu F, Wakebe H, Hishigaki H, Watanabe T, Sugiyama A, Takemoto M, Kawakami B, Yamazaki M, Watanabe K, Kumagai A, Itakura S, Fukuzumi Y, Fujimori Y, Komiyama M, Tashiro H, Tanigami A, Fujiwara T, Ono T, Yamada K, Fujii Y, Ozaki K, Hirao M, Ohmori Y, Kawabata A, Hikiji T, Kobatake N, Inagaki H, Ikema Y, Okamoto S, Okitani R, Kawakami T, Noguchi S, Itoh T, Shigeta K, Senba T, Matsumura K, Nakajima Y, Mizuno T, Morinaga M, Sasaki M, Togashi T, Oyama M, Hata H, Watanabe M, Komatsu T, Mizushima-Sugano J, Satoh T, Shirai Y, Takahashi Y, Nakagawa K, Okumura K, Nagase T, Nomura N, Kikuchi H, Masuho Y, Yamashita R, Nakai K, Yada T, Nakamura Y, Ohara O, Isogai T, Sugano S. Complete sequencing and characterization of 21,243 full-length human cDNAs. Nat. Genet 2004;36:40–45. [PubMed: 14702039]

Ozols J. Isolation, properties, and the complete amino acid sequence of a second form of 60-kDa glycoprotein esterase. Orientation of the 60-kDa proteins in the microsomal membrane. J. Biol. Chem 1989;264:12533–12545. [PubMed: 2745458]

Peitsch MC. Protein modeling by E-mail. BioTechnology 1995;13:658–660.

Pindel EV, Kedishvili NY, Abraham TL, Brzezinski MR, Zhang A, Dean RA, Bosron WF. Purification and cloning of a broad substrate specificity human liver carboxylesterase that catalyzes the hydrolysis of cocaine and heroin. J. Biol. Chem 1997;272:14769–14775. [PubMed: 9169443]

Pontius JU, Mullikin JC, Smith DR, Agencourt Sequencing Team; Lindblad-Toh K, Gnerre S, Clamp M, Chang J, Stephens R, Neelam B, Volfovsky N, Schäffer AA, Agarwala R, Narfström K, Murphy WJ, Giger U, Roca AL, Antunes A, Menotti-Raymond M, Yuhki N, Pecon-Slattery J, Johnson WE, Bourque G, Tesler G, NISC Comparative Sequencing Program. O'Brien SJ. Initial sequence and comparative analysis of the cat genome. Genome Res 2007;17:1675–1689. [PubMed: 17975172]

Potter PM, Wolverton JS, Morton CL, Wierdl M, Danks MK. Cellular localization domains of a rabbit and human carboxylesterase: influence on irinotecan (CPT-11) metabolism by the rabbit enzyme. Cancer Res 1998;58:3627–3632. [PubMed: 9721871]

Rat genome Sequencing Project Consortium. Genome sequence of the brown Norway rat yields insights into mammalian evolution. Nature 2004;428:493–521. [PubMed: 15057822]

Redinbo MR, Potter PN. Mammalian carboxylesterases: from drug targets to protein therapeutics. Drug Disc. Today 2005;10:313–320.

Rhesus Macaque genome Sequencing and Analysis Consortium. Evolutionary and biomedical insights from the rhesus macaque genome. Science 2007;316:222–234. [PubMed: 17431167]

Robbi M, Beaufay H. The COOH terminus of several liver carboxylesterases targets these enzymes to the lumen of the endoplasmic reticulum. J. Biol. Chem 1991;266:20498–20503. [PubMed: 1939102]

Ruppert C, Bagheri A, Markart P, Schmidt R, Seeger W, Gunther A. Liver carboxylesterase cleaves surfactant protein (SP-B) and promotes surfactant subtype conversion. Biochem. Biophys. Res. Commun 2006;348:1449–1454. [PubMed: 16919595]

Sanghani SP, Quinney SK, Fredenberg TB, Davis WI, Murry DJ, Bosron WF. Hydrolysis of irinotecan and its oxidative metabolites, 7-ethyl-10-[4-N(5-aminopentanoic acid)-1-piperidino] carbonyloxycampothecin and 7-ethyl-10-[4-(1-piperidino)-1 amino]-carbonyloxycamptothecin, by human carboxylesterases CES1A1, CES2, and a newly expressed carboxylesterase isoenzyme, CES3. Drug Metab. Dispos 2004;32:505–511. [PubMed: 15100172]

Satoh T, Hosokawa M. The mammalian carboxylesterases: from molecules to functions. Annu. Rev. Pharmacol. Toxicol 1998;38:257–288. [PubMed: 9597156]

Satoh T, Hosokawa M. Structure, function and regulation of carboxylesterases. Chemico-Biol. Interact 2006;162:195–211.

Satoh H, Taylor P, Bosron WF, Sanghani P, Hosokawa M, Du PB. Current progress on esterases: from molecular structure to function. Drug Metab. Dispos 2002;30:488–493. [PubMed: 11950776]

Schewer H, Langmann T, Daig R, Becker A, Aslandis C, Schmitz G. Molecular cloning and characterization of a novel putative carboxylesterase, present in human intestine and liver. Biochem. Biophys. Res. Commun 1997;233:117–120. [PubMed: 9144407]

Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res 2003;31:3381–3385. [PubMed: 12824332]

Shibita F, Takagi Y, Kitajima M, Kuroda T, Omura T. Molecular cloning and characterization of a human carboxylesterase gene. Genomics 1993;17:76–82. [PubMed: 8406473]

Tabata T, Katoh M, Tokudome S, Nakajima M, Yokoi T. Identification of the cytosolic carboxylesterase catalyzing the 5'-deoxy-5-fluorocytidine formation from capecitabine in human liver. Drug Metab. Dispos 2004;32:1103–1110. [PubMed: 15269188]

Takai S, Matsuda A, Usami Y, Adachi T, Sugiyama T, Katagiri Y, Tatematsu M, Hirano K. Hydrolytic profile for ester- or amide- linkage by carboxylesterases 5.3 and 4.5 from human liver. Biol. Pharmacol. Bull 1997;20:869–873.

The Chimpanzee Sequencing and Analysis Consortium. Nature 2005;437:69–87. [PubMed: 16136131]

Thierry-Mieg D, Thierry-Mieg J. AceView: A comprehensive cDNA-supported gene and transcripts annotation. Genome Biol 2006;7:S12. [PubMed: 16925834] http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/av.cgi?exdb=AceView&db=36a&term=CES7&submit=Go

Tsujita T, Okuda H. Palmitoyl-coenzyme A hydrolyzing activity in rat kidney and its relationship with carboxylesterase. J. Lipid Res 1993;34:1773–1781. [PubMed: 7902406]

von Heijne G. Patterns of amino acids near signal-sequence cleavage sites. Eur. J. Biochem 1983;133:17–21. [PubMed: 6852022]

Wang H, Gilham D, Lehner R. Proteomic and lipid characterization of apo-lipoprotein B-free luminal lipid droplets from mouse liver microsomes: implications for very low density lipoprotein assembly. J. Biol. Chem 2007;282:33218–33226. [PubMed: 17848546]

Woodburne MO, Rich TH, Springer MS. The evolution of tribospheny and the antiquity of mammalian clades. Mol Phylogenet. Evol 2003;28:360–385. [PubMed: 12878472]

Xu G, Zhang W, Ma MK, McLeod HL. Human carboxylesterase 2 is commonly expressed in tumor tissue and is correlated with the activation of irinotecan. Clin. Cancer Res 2002;8:2605–2611. [PubMed: 12171891]

Yan D, Matoney L, Yang D. Human carboxylesterases in term placentae: enzyme characterization, molecular cloning and evidence for the existence of multiple forms. Placenta 1999;20:599–607. [PubMed: 10452915]

Zhen L, Rusiniak ME, Swank RT. The beta-glucuronidase propeptide contains a serpin-related octamer necessary for complex formation with egasyn esterase and for retention within the endoplasmic reticulum. J. Biol. Chem 1995;270:11912–11920. [PubMed: 7744842]

Figure 1. Amino acid sequence alignments for human CES1, CES2 and CES5

See Table 1 for sources of CES sequences; * shows identical residues for human CES1, CES2 and CES5; : 2 alternate residues. Residues involved in endoplasmic reticulum processing at N- (Signal peptide) and C- termini (MTS-microsomal (endoplasmic reticulum) targeting sequence); **N-glycosylation residues** at 79NAT (Human CES1) and potential N-glycosylation sites; AS shows active site triad residues Ser; Glu; and His. '**Side door**', '**Gate**' residues and Cholesterol binding Gly residue (**Z site**) for human CES1 Disulfide bond ---- Cys residues for human CES1. **Charge clamp residues identified for human CES1**; Helix (Human CES1 or predicted helix; Sheet (Human CES1) or predicted sheet. **Large font shows known or**

predicted exon junctions. Exons are numbered for CES1, CES2 and CES5.
Predicted transmembrane regions are shown

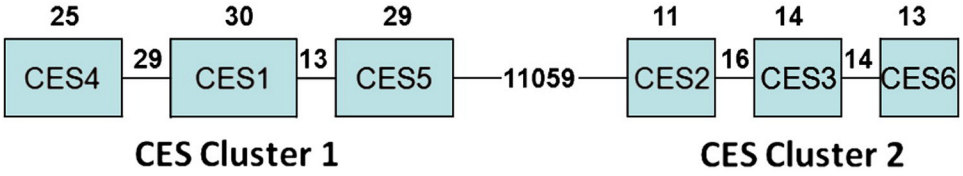

Figure 2. Locations of CES Genes on Human Chromosome 16
Numbers refer to kilobases of DNA.

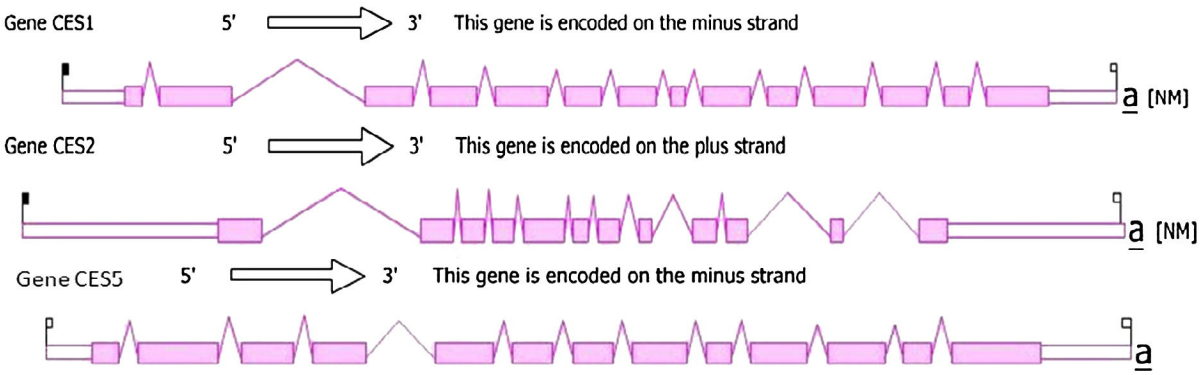

Figure 3. Gene Structures and Slicing Variants for Human CES1, CES2 and CES5 Genes

Derived from the AceView website (Thierry-Mieg and Thierry-Mieg, 2006). http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/av.cgi?exdb=AceView&db=36a&term=CES7&submit=Go Mature isoform variants (a, b, c etc) are shown with capped 5'- and validated 3'-ends for the predicted mRNA sequences. NM numbers refer to annotated RefSeq sequences for CES1, CES2 and CES5 genes. Scale refers to base pairs of nucleotide sequences.

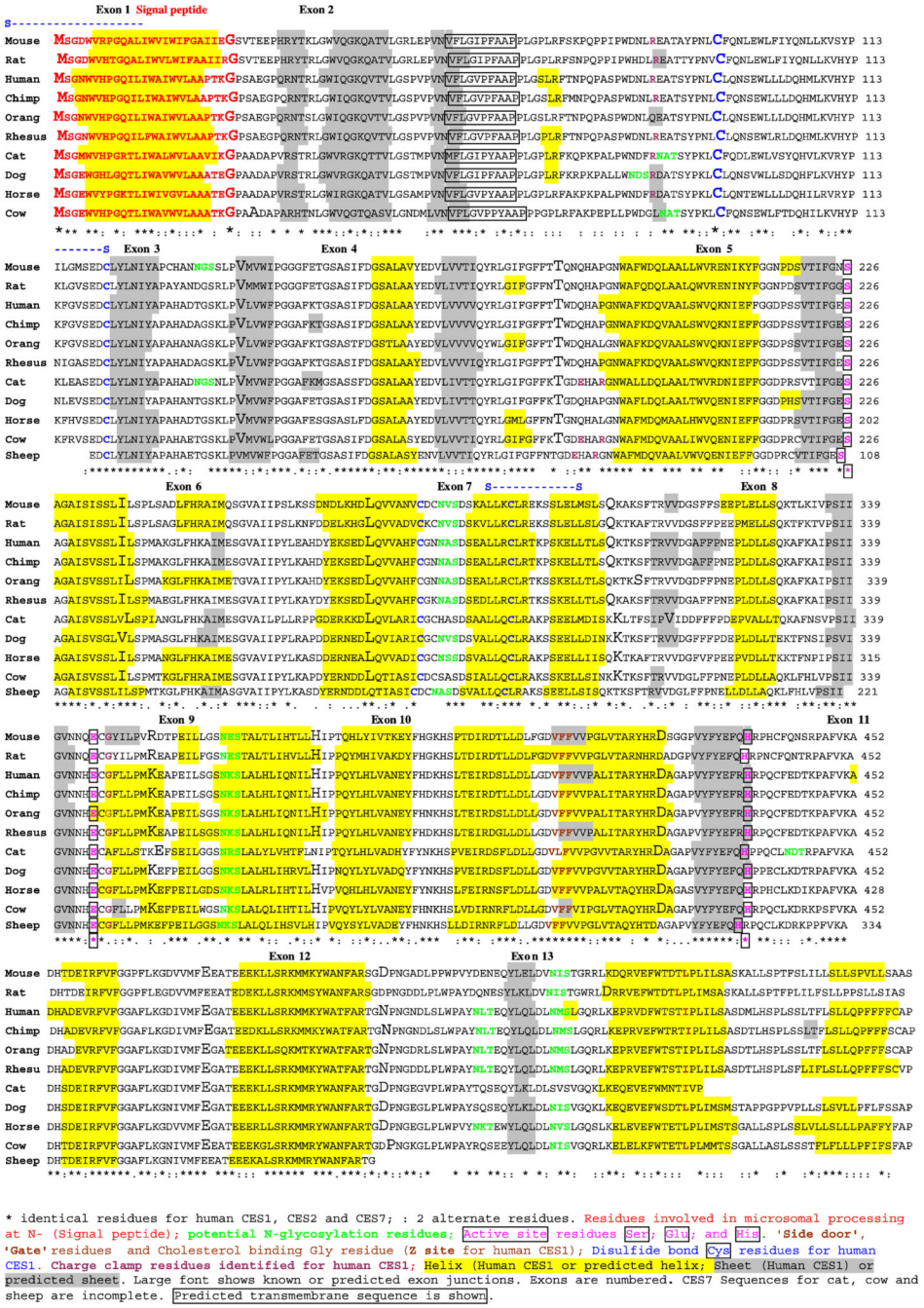

Figure 4. Amino acid sequence alignments for mammalian CES5

See Table 1 for sources of CES5 sequences; * show identical residues for mammalian CES5; : 2 alternate residues. Residues involved in microsomal processing at N- (Signal peptide); **N-glycosylation residues** at potential N-glycosylation sites; AS shows active site triad residues Ser; Glu; and His. '**Side door**', '**Gate**' residues and Cholesterol binding Gly residue (**Z site**) (for human CES1) Disulfide bond ---- Cys residues. **Charge clamp residues previously identified for human CES1**; Predicted helix; Predicted Sheet. **Large font shows known or predicted exon junctions. Exons are numbered for CES1, CES2 and CES5.**

Predicted transmembrane regions are shown

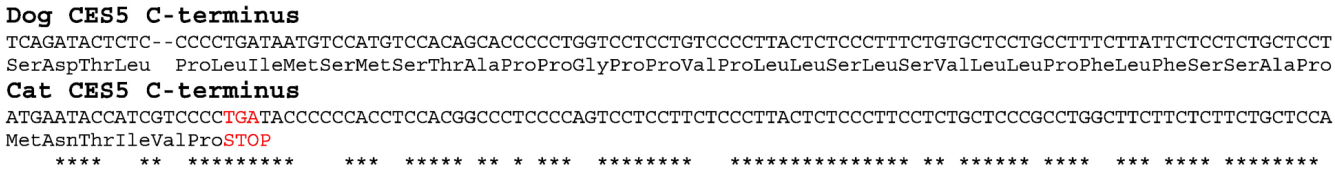

Figure 5. Carboxy-terminus nucleotide sequences and predicted amino acid sequences for dog and cat CES5 genes and proteins

Sequences were derived from BLAT analyses using known sequences for dog and cat CES5 sequences and the UCSC web site (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start) (see Table 1 for sources). * shows identical nucleotide residues following alignment of sequences. Stop indicates STOP codon for C-terminus observed in the cat CES7 gene. Note the apparent GT INDEL (insertion deletion) within the CES5 sequence generates an earlier termination of translation.

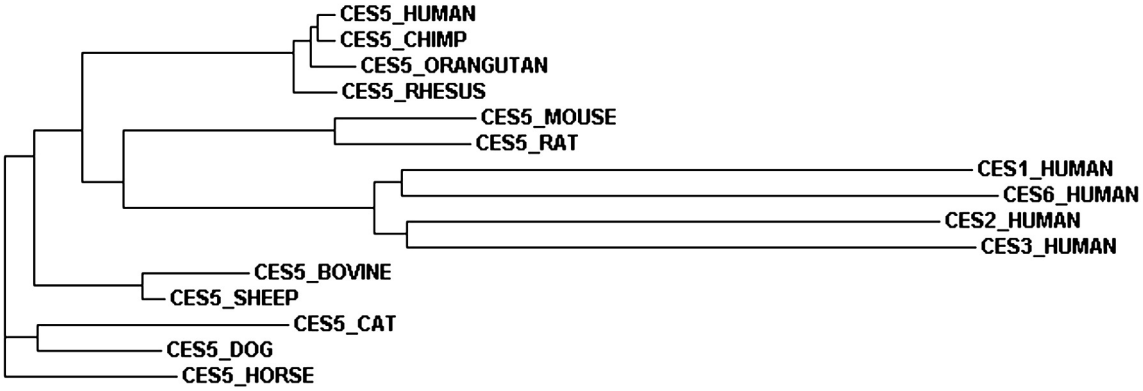

Figure 6. Phylogenetic tree of mammalian CES5 and of human CES1, CES2, CES3 and CES6 sequences

The tree is labeled with the gene name and the species name. Note the separation of distinct CES1, CES2, CES3 and CES6 gene families from the mammalian CES5 family cluster. The gene duplication events generating five distinct gene families (CES1, CES2, CES3, CES5 and CES6) have been previously estimated to have occurred ~ 328–378 million years ago (Holmes et al., 2008).

Comp Biochem Physiol Part D Genomics Proteomics. Author manuscript; available in PMC 2009 September 1.

Table 1

Mammalian CES7 Genes and Human CES Genes.

Mammal	CES Gene	GenBank ID *N-scan ID "clone fragment	UNIPROT ID	No of Amino Acids	Chromosome Location (*Scaffold)	Strand	CES Gene order	Gene Size (bps)
Human	*CES5*	BC069501	**Q6NT32**	575	16:54,437,867-54,466,634	Negative	CES4-CES1-**CES5**	28767
Chimp	*CES5*	*chr16.56.003		575	16: 55,250,783-55,279,556	Negative	CES4-CES1-**CES5**	28773
Orangutan	*CES5*	*chr16.470.1		575	16: 43,074,925-43,106,130	Negative	CES1-**CES5**	31205
Rhesus	*CES5*	*chr20.55.004		575	20:54,155,029-54,189,003	Negative	CES1-**CES5**	34074
Mouse	*CES5*	AB186393	**Q6AW46**	575	8:96,038,096-96,059,607	Negative	CESN-CES3-CES22-CES1-**CES5**	21511
Rat	*CES5*	AF479659	**Q5GRG2**	575	19:11,910,831-11,938,412	Positive	**CES5**-CES4-CES1-CES3	27581
Cat	*CES5*	AB045377	**Q81O34**	545	*150208:218,403-249,533	Negative	CES1-**CES5**	31130
Dog	*CES5*	AB186392	**Q6AW47**	575	2:63,116,213-63,145,396	Positive	**CES5**-CES1	29183
Cow	*CES5*	"AAFC03022775		553	18:24,405,935-24,433,935	Negative	CES1A-**CES5**-CES1B	28000
Horse	*CES5*	XP1493477		575	3:8,287,880-8,318,561	Negative	CES1A-1B-1C-1D-1E-**CES5**	30681
Sheep	*CES5*	AM075621	**Q3T930**	381				
Human	*CES1*	L07765	**P23141**	567	16:54,394,265-54,424,576	Negative	CES4-**CES1**-CES5	30311
	CES2	BX538086	**O00348**	559	16:65,525,828-65,536,493	Positive	**CES2**-CES3-CES6	10665
	CES3	NM024922	**Q6UWW8**	571	16: 65,552,639-65,566,552	Positive	CES2-**CES3**-CES6	13913
	CES4	BC131699	**Q8TDZ9**	505	16:54,352,012-54,364,924	Negative	**CES4**-CES1-CES5	12912
	CES6	FLJ37464	**Q6UX55**	545	16:65,580,134-65,602,403	Positive	CES2-CES3-**CES6**	22269

Table 2

Predicted N-Glycosylation Sites for Mammalian CES5

Mammal	CES Gene	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	No of Sites
Orangutan	CES5			281-3 NAS	363-5 NKS		511-3 NLT	522-4 NMS	4
Human	*CES5*			281-3 NAS	363-5 NES		511-3 NLT	522-4 NMS	4
Chimp	*CES5*			281-3 NAS	363-5 NKS		511-3 NLT	522-4 NMS	4
Rhesus	*CES5*			281-3 NAS	363-5 NKS		511-3 NLT	522-4 NMS	4
Mouse	*CES5*		134-6 NGS	281-3 NVS	363-5 NES			522-4 NIS	4
Rat	*CES5*			281-3 NVS	363-5 NES			522-4 NIS	3
Horse	*CES5*			281-3 NSS	363-5 NKS			522-4 NVS	3
Cow	*CES5*	86-88 NAT						522-4 NIS	2
Cat	*CES5*	86-88 NAT	134-6 NGS		363-5 NKS	443-5 NDT			4
Dog	*CES5*	82-82 NDS		281-3 NVS	363-5 NKS			522-4 NIS	4
Sheep	*CES5*			263-5 NAS	235-7 NKS				2

Numbers refer to amino acids in CES5 Sequence: N-Asn; A-Ala; S-Ser; V-Val; T-Thr; G-Gly; K-Lys; L-Leu; E-Glu; D-Asp; I-Ile; M-Met.

Table 3

Percentage Identities for Mammalian CES5 Amino Acid Sequences

CES	hCES5	chCES5	orCES5	rhCES5	mCES5	raCES5	caCES5	doCES5	coCES5	hoCES5	shCES5
hCES5	100	98	96	95	69	69	70	75	74	76	80
chCES5	98	100	96	96	70	70	70	75	74	76	80
orCES5	96	96	100	94	68	69	69	74	73	75	78
rhCES5	95	96	94	100	70	70	74	75	69	75	80
mCES5	69	70	68	70	100	86	64	68	67	66	73
raCES5	69	70	69	70	86	100	65	69	67	67	72
caCES5	70	70	69	69	64	65	100	83	74	78	75
doCES5	75	75	74	75	68	69	83	100	81	85	83
coCES5	74	74	73	74	67	67	74	81	100	82	93
hoCES5	76	76	75	75	66	67	78	85	82	100	82
shCES5	80	80	78	80	73	72	75	83	93	82	100

h-human; ch-chimpanzee; or-Orangutan; rh-Rhesus monkey; m-mouse; ra-rat; ca-cat; do-dog; co-cow; ho-horse; sh-sheep.