Cholinesterase-like domains in enzymes and structural proteins: functional and evolutionary relationships and identification of a catalytically essential aspartic acid

E. Krejci, Nicolas Duval, Arnaud Chatonnet, P. Vincens, J. Massoulie

To cite this version:


HAL Id: hal-02711891
https://hal.inrae.fr/hal-02711891
Submitted on 1 Jun 2020
Cholinesterase-like domains in enzymes and structural proteins: Functional and evolutionary relationships and identification of a catalytically essential aspartic acid

(cholinesterase/esterase/lipase/neurotactin/glutabin)

ERIC KREJCÍ*, NATHALIE DUVAL*, ARNAUD CHATONNET†, PIERRE VINCEN§, AND JEAN MASSOULIÉ*

*Laboratoire de Neurobiologie, Centre National de la Recherche Scientifique Unité Associée 295, Ecole Normale Supérieure, 75005 Paris, France; †Laboratoire de Physiologie Animale, Institut National de la Recherche Agronomique, 34060 Montpellier, France; §Groupe de Bioinformatique, Laboratoire de Biochimie et Physiologie du Développement, Centre National de la Recherche Scientifique Unité Associée 668, Ecole Normale Supérieure, 75005 Paris, France

Communicated by David D. Sabatini, April 1, 1991 (received for review December 27, 1990)

ABSTRACT Primary sequences of cholinesterases and related proteins have been systematically compared. The cholinesterase-like domain of these proteins, about 500 amino acids, may fulfill a catalytic and a structural function. We identified an aspartic acid residue that is conserved among esterases and lipases (Asp-397 in Torpedo acetylcholinesterase) but that had not been considered to be involved in the catalytic mechanism. Site-directed mutagenesis demonstrated that this residue is necessary for activity. Analysis of evolutionary relationships shows that the noncatalytic members of the family do not constitute a separate subgroup, suggesting that loss of catalytic activity occurred independently on several occasions, probably from bifunctional molecules. Cholinesterases may thus be involved in cell-cell interactions in addition to the hydrolysis of acetylcholine. This would explain their specific expression in well-defined territories during embryogenesis before the formation of cholinergic synapses and their presence in noncholinergic tissues.

Acetylcholinesterases (AcChoEases, EC 3.1.1.7) and butyrylcholinesterases (BtChoEases, EC 3.1.1.8) (1–3) constitute a closely homologous family of proteins (4–13). These enzymes consist of a major common catalytic domain and a small C-terminal variable region. In vertebrate AcChoEase, two types of C-terminal regions are encoded by alternative exons, differentiating the subunits that participate in the collagen-associated, hydrophobic-tailed, or soluble tetramers (exon T) and in the glycolipid-anchored dimers (exon X) (refs. 14 and 15; for review, see ref. 3). A number of proteins have been shown to present homology with the common domain of cholinesterases (ChoEases). In addition to serine hydrolases such as carboxylesterases (16–23), pancreatic lysophospholipase or cholesterol esterase (24, 25), and lipase from Geotrichum candidum (26), this family includes noncatalytic proteins; the C-terminal domain of bovine (27), rat (28), and human (29) thyroglobulins (4, 30, 31); and two Drosophila proteins that seem to be involved in cellular interactions, glutacin (32) and neurotactin (33, 34).

Acetylcholine by AcChoEase operates near the theoretical limit set by the diffusion of the substrate (35). Its catalytic mechanism remains obscure, however; in particular, the presence of a charge–relax system, Asp/His/ Ser, is not established. We presently show that Asp-397 of Torpedo AcChoEase, which, to our knowledge, has not hitherto been considered to be involved in the catalytic mechanism of ChoEases, is strictly conserved among Choease-like esterases and is essential for activity.

ChoEases now appear as prototypes of a family of proteins. In this report, we compare the primary structures of the various elements of this family. We define an alignment of amino acid sequences, use it to identify common features, and discuss functional and evolutionary implications.

RESULTS AND DISCUSSION

Alignment of Peptide Sequences. Fig. 1 shows an alignment of ChoEases and homologous protein domains. This alignment was obtained by visual inspection on the basis of strongly conserved peptides or isolated amino acid motifs. Essentially the same peptide motifs were found by using the multialign program developed by Corpet (38). Some sequences present insertions, e.g., large hydrophilic peptides in Drosophila AcChoEase (RANGGEGHPNKQUADTHLH-NGPATTTG) (9) and Anopheles AcChoEase (GLNGF-SNDYFGQDDDFQRQHKSQG) (12). Such additional peptides are likely to be located at the periphery of the tertiary structure of the protein between distinct elements of secondary structure, so as not to disturb its organization. In Drosophila AcChoEase, this peptide is normally removed from the mature subunit, producing two fragments of 18 and 55 kDa from the 75-kDa precursor (39, 40).

Fig. 2 illustrates the position of the ChoEase-like domain in various proteins. This domain is represented as a grey bar, in which we have indicated the approximate position of residues that seem to play a role in the catalytic mechanism of esterases (see below). In addition, we have marked the position of intracatenary disulfide bridges, deduced from homology with Torpedo AcChoEase (36). Note that the ChoEase-like domain constitutes the predominant part of the protein in esterases but is associated with other large peptide domains in the nonenzymatic proteins (thyroglobulin, glutacin, and neurotactin).

Identification of an Essential Aspartic Acid Residue. The present analysis shows that Asp-397 (numbering of Torpedo AcChoEase) is present in all ChoEase-like esterases, except in the first primary sequence determined for Torpedo marmorata, which contained asparagine at this position (codon AAC) (5). This was probably due to an error in the construction of the cDNA since independent clones contained aspartic acid (codon GAC) at this position.

Transfections of COS cells with CDM8 constructions (41, 42) that contained either Asp-397 or Asn-397 but were otherwise identical showed that AcChoEase activity was produced exclusively in the first case, although the synthesis and the stability of the protein were equivalent (Fig. 3).

Abbreviations: ChoEase, cholinesterase; AcChoEase, acetylcho- linesterase; BtChoEase, butyrylcholinesterase.
Fig. 1. (Legend appears at the bottom of the opposite page.)
AcChoEase residues of microsomal 5, seem to and may have reported the existence of a ChoEase-like domain in sequence 2. In sequence 2, Asp-93 is identified. Enzymes active in ChoEases are BtChoEase proton, at H, histidine involved in catalysis with bovine AcChoEase. Thus, BtChoEase is predicted to have a catalytic triad (D,JH, H). Analysis of amino acid sequence of neurotactin is solid (TM). The alternative C-terminal region of ChoEases is related by homology of sequence 2, and H is a glutamyl histidine in sequence 2. Positions of the catalytic triad, Asp/His/Ser, in ChoEases and homologous esterases remains an open question. Apart from the immediate environment of the active serine (GXSXG), they show no homology with enzymes in which the elements of a catalytic triad have been formally identified. In Torpedo AcChoEase, the active serine is Ser-200, and an essential histidine has been identified by site-directed mutagenesis as His-440 by Giblye et al. (46). All the enzymes of the family, serine and histidine residues occur at homologous positions. For rat carboxylesterase, this histidine residue was reported to react with diisopropyl fluorophosphate (18).

Assuming the existence of a catalytic triad, Doctor et al. (47) have considered Asp-93 (numbering of Torpedo AcChoEase) as a possible candidate because (i) it seems to exist in all ChoEases, (ii) it is contained in a very well-conserved peptide sequence, and (iii) it is close to an intrachain disulfide bridge. In addition, these authors showed that an inhibitory monoclonal antibody, AE-2, binds to this region of fetal bovine serum AcChoEase. The absence of this residue in juvenile hormone esterase, unless due to a cloning error, precludes an essential role in catalysis. A similar reservation applies to Asp-172 (numbering of Torpedo AcChoEase), another possible candidate proposed by Soe and Prod (48); it is conserved in most ChoEase-like proteins but absent from rabbit esterase 1.

The alignment of Fig. 1 shows a very limited number of positions where all or most enzymes of the family possess aspartic or glutamic acids. In addition to the residues already discussed (Asp-93, Asp-172, and Asp-397), Glu-92, Glu(Asp)-327, and Glu(Asp)-443 (numbering of Torpedo AcChoEase) are the only other conserved acidic residues. The present demonstration that Asp → Asn at position 397 results in the loss of Torpedo AcChoEase activity makes it a good candidate for the third element of a charge–relay system. For trypsin, an Asp → Asn change in the triad reduced the activity by four orders of magnitude (49) but did not affect the geometry of the corresponding residues or the structure of the active site (50).

It should be noted, however, that both Asp-397 and His-440 are conserved in all ChoEase-like proteins, including non-esterolytic ones, suggesting that they may be important for structural reasons and not necessarily or exclusively as part of a catalytic triad.

**Comparison of ChoEase-Like Serine Esterases with Other Serine Hydrolases.** The enzymes that present a structural homology with ChoEases are all serine esterases and possess...
similar catalytic properties, as indicated by their EC numbers: lysophospholipase (EC 3.1.1.6) and esterases (EC 3.1.1.1; rat carboxylesterase, Drosophila esterase P and esterase 6, and insect juvenile hormone esterase). Note, however, that other esterases, included in the same group, appear entirely unrelated to this family.

In ChoEases and other esterases, lipases, and serine proteases, the peptides containing the active serine are similar (GESAG in ChoEases, GDSG in eukaryotic serine proteases, and GX1SX2G in other eukaryotic and prokaryotic esterases and lipases) (51). This is probably the result of convergent evolution, since this is the only recognizable feature common to these proteins. In particular, although the elements of a charge-relay system, Asp/His/Ser, are not located in equivalent positions along the primary sequence of ChoEases and serine proteases, convergent evolution certainly resulted in a similar spatial arrangement of this triad. Crystallography has shown, for example, that the prokaryotic enzymes subtilisin and Mucor miehei lipase possess Asp/His/Ser triads that are geometrically very similar to that of chymotrypsin (52), although the order of the residues in the sequence is different (their positions being 32/64/221, 203/257/144, and 102/57/195).

Evolutionary Relationships. From the alignment presented in Fig. 1, we calculated distances between each pair of sequences with the matrix of Dayhoff (58) and constructed a tree according to the method of Saitou and Nei (53) (Fig. 4). It is clear that the residues which are directly involved in specific functions (e.g., the catalytic triad of hydrolases or hormonogenic residues of thyroglobulin) are too few to determine these distances. We may, therefore, consider that these differences correspond to evolutionary drift among proteins that have conserved a common overall structure.

The results allow us to identify subgroups within the family of ChoEase-like domains: ChoEases form such a group that also includes mammalian microsomal carboxylesterases, Drosophila esterases, and Dictyostelium esterases (Fig. 4). The definition of these subsets does not depend on the distance parameters used or on the choice of fragments of the proteic sequence. In contrast, the arrangement of these branches into a complete phylogenetic tree is variable: considering different regions of the domain (N-terminal and C-terminal halves) leads to different patterns. It is clear that the divergence between the main branches illustrated in Fig. 4 (boxed) is very ancient, possibly comparable to the separation between the ancestor of Dictyostelium and those of invertebrates and vertebrates, which is considered to be 1200–1500 million years old. Fungi and animals diverged 1000–1200 million years ago, and the separation between vertebrates and invertebrates was about 700 million years ago. The relationships between catalytic and noncatalytic members of the ChoEase family suggest that the catalytic capacity was lost on several independent occasions during evolution.

Noncholinergic Functions of ChoEases? The presence of a ChoEase-like domain in noncatalytic proteins may reflect its capacity for protein–protein interactions in enzymes and in noncatalytic proteins. The ChoEase-like enzymes might thus possess a dual, catalytic and structural, role. ChoEases themselves, in addition to their catalytic function, may be involved in recognition and/or adhesion mechanisms. This was suggested on the basis of their very early appearance in well-defined territories during embryogenesis (54). In the chicken embryo, Layer (55) proposed that the expression of BtChoEase is correlated with cellular movements or neurite extension whereas the appearance of AcChoEase is one of the first signs of nervous or muscular differentiation. Moreover, Greenfield (56) showed that AcChoEase modifies the activity of hypothalamic neurons. This effect was totally unrelated to the ChoEase activity, since it persisted after irreversible inhibition or heat inactivation. The possibility of an adhesion function for ChoEases agrees well with the presence of an HNK1-like carbohydrate epitope, which is considered a hallmark of adhesion proteins (57), on some forms of AcChoEase from electric organs (43).
Note Added in Proof. The three-dimensional structure of AcChoEase from Torpedo, recently determined by crystallographic analysis, indicates that Glu-327 is probably part of the catalytic triad, whereas Asp-397 is not directly associated with the active site (J. L. Sussman, M. Harel, F. Frolov, C. Oefner, A. Goldman, L. Toker, and I. Silman, personal communication).

We thank Mrs. Jacqueline Pons and Rose Bouaziz for typing the manuscript. We thank our colleagues Drs. Fernando Jiménez, Bhubendra Das, Levanger Taylor, and William R. Randall for early communication of their sequence data and Dr. Michel Piovant, Prof. André Adoutte, and Prof. L. Pezzementi for fruitful discussions. This work was supported by grants from the Centre National de la Recherche Scientifique, the Direction des Recherches et Etudes Techniques, and the Association Française contre les Myopathies. N.D. was recipient of a fellowship from the Association Française contre les Myopathies.