

An analysis of acetylcholinesterase sequence for predicting mechanisms of its non-catalytic actions

Malathi Srivatsan*

Department of Biological Sciences, Arkansas State University, Jonesboro, AR 72401, USA; Malathi Srivatsan* -

Email: msrivatsan@astate.edu; Phone: +870 972 3167; Fax: +870 972 2168; * Corresponding author

received November 17, 2006; revised November 28, 2006; accepted December 04, 2006; published online December 06, 2006

Abstract:

The enzyme acetylcholinesterase (AChE) which belongs to the family of α/β hydrolases is well known for hydrolyzing the neurotransmitter acetylcholine (ACh). In addition to its catalytic function, AChE appears to play a significant non-catalytic role in development, regeneration and modulation of properties of neurons. However the mechanisms underlying these important actions of AChE are unknown. It prompted the analysis of the sequence of fetal bovine serum AChE to seek clue(s) for the mechanisms of AChE non-catalytic actions. The searches for motifs, finger prints and domains demonstrated the presence of a highly conserved carboxylesterase type B signature in AChE from slime molds to human. Interestingly, the presence of finger prints for a metabotropic glutamate receptor, gap junction protein connexin, a calcium binding motif, several phosphorylation sites, and a motif similar to nicotinic acetylcholine receptor were also found in the sequence of AChE suggesting a role for AChE in cell signalling. The similarity search using gapped BLAST retrieved previously known similar proteins, namely neurotactin, neuroligin and thyroglobulin. When aligned neurotactin, neuroligin and bovine AChE showed the presence of three stretches of consensus, one of which is carboxylesterase B signature. The identity of the other two are not known. Amino acids in these stretches in AChE may be involved in possible growth factor like action and awaits experimental verification. Thus sequence analysis can be very useful for biologists while searching for a possible mechanism of action of a protein. Further, these results underscore the significance of performing both local and global searches to identify possible functional domains as well as consensus sequence.

Keywords: motifs; domain; finger prints

Background:

The enzyme acetylcholinesterase (AChE) is the most efficient of all known enzymes. [1] AChE has received much attention due to its important action of rapidly catalyzing the hydrolysis of the neurotransmitter acetylcholine (ACh) at the cholinergic synapse. This knowledge has led to the development of anticholinesterases to treat cholinergic disorders and to use as pesticides in agriculture or nerve gas weapons in war.

AChE belongs to the family of α/β hydrolases and each subunit core consists of eight β sheets connected by α helices. The active site is described as a triad consisting of serine, aspartic acid and histidine residues. [2] The sequences of more than five hundred cholinesterases are available now. The primary structure of a mature catalytic subunit consists of a major common domain (motif common in all cholinesterases, other esterases and lipases) followed by variable C-terminal peptides.

Research in the last twenty years has revealed that in addition to its catalytic function, AChE appears to have several significant non-catalytic actions in the nervous system. AChE is present both as a membrane anchored and as a circulating secretory form in sites not requiring hydrolysis of ACh, suggesting that this molecule has additional, non-catalytic functions. [1] AChE directly modulates the electrical properties of neurons and has been reported to reduce an inward rectifier K^+ current. [3] Infusion of AChE induces behavioral changes in rats involving non-cholinergic pathway. [4] Subsequent studies have shown that a circulating form of AChE promotes neurite regeneration and growth in primary cultures of adult *Aplysia* dopaminergic neurons via a non-catalytic mechanism. [5] Overexpressed AChE promotes

neurite growth in DRG neurons of rat. [6] AChE purified from fetal bovine serum promotes neurite growth in sympathetic neurons of rat. [7] However the mechanisms underlying these important, non-catalytic actions of AChE are unknown. Since the sequence of a protein determines its structure and structure in turn usually determines its function, we hypothesized that the sequence of AChE has functional domains other than its catalytic site. Analysis of the sequence of AChE can provide clue(s) to possible mechanisms of non-catalytic actions of AChE by revealing the presence of motif, functional domain or sequence stretch similar to those present in other proteins known to have similar actions. The growth in recent years in the availability of easily accessible public domain data bases with robust search tools has made such analyses faster and easier. Since fetal bovine serum AChE is neurotrophic to neurons via a non-catalytic mechanism [7], we analyzed the sequence of bovine serum AChE to identify the presence of functional domains or sequence similarity to other proteins. Our findings which showed the presence of several interesting functional domains in the sequence of AChE are described here.

Methodology:

The secreted, circulating form of AChE appears to mediate most of its non-catalytic actions. Since bovine serum AChE is a secreted, circulating protein and its sequence is known, NCBI [8] protein database was searched using key words bovine serum acetylcholinesterase and returned five sequences. Of these only three were complete sequences of which two were obtained by sequencing the isolated proteins, one consisting of 583 amino acids and the other consisting of 613 amino acids. The third one (gi:14916950) (Figure 1) was the most recent and updated sequence. This sequence also

consisted of 613 amino acids and was the deduced amino acid sequence from the gene cloned from genomic DNA. The recombinant protein expressed using this clone had similar

characteristics as the naturally occurring protein in the serum. Hence for all the searches reported here, the FASTA format of this sequence as given here was used.

Figure 1

```
P23795. Acetylcholinesterase [gi:14916950]
>gi|14916950|sp|P23795|ACES_BOVIN Acetylcholinesterase precursor (AChE)
MRPPWCPLHTPSLTPPLLLLLLFLIGGGAEAEAGPEDELLVMVRGGRLRGLRLMAPRGPVSAFLGIPFAEP
PVGPRRFLPPEPKRPWPGVLNATAFQSVCYQYVDTLYPGFEGTEMWNPNSRESDCLYLNWVTPYPRPSS
PTPVLVWIYGGGFYSGASSLDVYDGRFLTQAEGTVLVSMNYRVGAFGFLALPGSREAPGNVGLLDQRLAL
QWVQENVAAFGGDPTSVTLFGEAGAAASVGMHLLSPPSRGLFHRAVLQSGAPNGPWATVGVGGEARRRATL
LARLVGCPPGGAGGNDTELVACLRRAPAQDLVDHEWRVLPQESVFRFSFVPVVDGDFLSDTPEALINAGD
FHGLQVLVGVVKDEGSYFLVYGAPGFSKDNESLISRAQLAGVRVGVQPASDLAAEAVALHYTDWLHPED
PARLREALSDVVDHNVVCPVAQLAGRLAAQGARVYAYIFEHRASLTSWPLWMGVPHGYEIEFIFGLPLE
PSLNYTIEERTFAQRLMRYWANFARTGDPNDPRDPKAPQWPPYTAGAQYVSLNLRPLEVRRGLRAQACA
FWRNRLPKLLSATDTLDEAERQWKAEFHRWSSYMVHWKNQFDHYSKQDRCSDL
```

To obtain clues for the existence of novel, non-catalytic functional domains in bovine serum AChE, motif analysis was performed at PROSITE [9] and eMOTIF. [10] Protein finger print and functional profile analyses were performed at PRINTS [11] and Pfam. [12] A search for related proteins through sequence alignment was performed at Gapped-Blast [13] and search for related proteins based on structure were performed at PDB. [14]

Results and Discussion

Motif, finger print and family analyses of bovine AChE

Analyses for Motifs (conserved pattern of amino acids that is found in two or more proteins, often having a similar biochemical activity) was performed at PROSITE [9] and at eMOTIFS. [10] Prosite scan was performed with and with out excluding patterns with high probability of occurrence. The results showed the presence of several phosphorylation sites (commonly occurring patterns) and two true hits for carboxylesterases type B1 and B2 and the number of occurrences of those patterns in the sequence.

PROSIE Entry	Description of site	Occurrences in sequence
PS00008	N-myristoylation site	13
PS00001	N-glycosylation site	4
PS00007	Tyrosine kinase phosphorylation site	1
PS00004	cAMP- and cGMP-dependent protein kinase phosphorylation site	2
PS00006	Casein kinase II phosphorylation site	6
PS00122	Carboxylesterases type-B serine active site	1
PS00941	Carboxylesterases type-B signature 2	1

Table 1: Commonly occurring patterns found in the sequence of bovine AChE by motif analysis at PROSITE

Higher eukaryotes have many distinct esterases. Among the different types are those which act on carboxylic esters (EC 3.1.1.). Carboxyl-esterases have been classified into three categories (A, B and C) on the basis of differential patterns of inhibition by organophosphates. Since carboxylesterases and cholinesterases have the common esterase domain, it is not surprising to find these hits. As the eMOTIFS are derived from the multiple sequence alignments in the BLOCKS+ and the PRINTS data base, an eMOTIF search with bovine AChE at an expectation of 0.01 or less found 90 matches comprising of cholinesterases or carboxylesterases. When a relaxed search with an expectation of 10 or less was performed, 143 matches including members of other enzyme families such as aldehyde dehydrogenase (5.89e+00, E- value is combined *p* value x number of alignments done between

query and data base) resulted. One interesting match found here was nicotinic acetylcholine receptor (6.89e+00) which might explain why some of the AChE inhibitors such as tacrine and physostigmine also bind to the nicotinic ACh receptor itself [15]

We then searched for known protein fingerprints in bovine AChE using the PRINTS [11] data base which is a collection of protein fingerprints or group of conserved motifs used to characterise a protein family; its diagnostic power is refined by iterative scanning of a SWISS-PROT/TrEMBL composite. Fingerprints can encode protein folds and domains more flexibly and powerfully than can single motifs and hence we searched PRINTS data base with the sequence of bovine AChE and the result is given in Table 1.

Fingerprint	No. of Motifs	Pvalue
<u>CHOLNESTRASE</u>	6 of 6	4.5e-46
<u>MTABOTROPIC GLU 7R</u>	2 of 7	1.6e-05
<u>PRICHEXTENSIN</u>	2 of 8	2.9e-05
<u>MATRIXIN</u>	2 of 5	0.00025
<u>ZETATUBULIN</u>	2 of 15	0.00028
<u>PEROPSIN</u>	2 of 11	0.00066
<u>ACRIFLAVINRP</u>	2 of 9	0.0013
<u>CLOACIN</u>	2 of 8	0.0016
<u>HPVCAPSIDL1</u>	2 of 11	0.0022
<u>CONNEXINB3</u>	2 of 4	0.0022

Table 2: The protein finger prints or conserved motifs found in bovine AChE by searching the PRINTS data base

The search at PRINTS also yielded the most significant match for the cholinesterase signature. Interestingly, the second most significant match was to the metabotropic glutamate receptor (GLU 7R). GLU R is coupled to G-proteins and stimulate the inositol phosphate/Ca²⁺ intracellular signalling pathway. [16] The amino acid sequence of the membrane-associated domain of this receptor which matches to AChE is also shared by the calcium-sensing receptors. It is possible that AChE-mediated neuromodulation occurs via this stretch of sequence of AChE using intracellular signaling. Connexin is a gap junction protein, performing an adhesive function.

The connexin finger print present in AChE suggests that the sequence stretch in AChE that resembles connexin may have an adhesive function and thus can promote neurite growth.

Pfam [12] is a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains. Search of Pfam library database of profiles hidden Markov models (HMM) with bovine AChE sequence in the local alignment mode yielded the following results (Table 2)

Model	Score	E-Value	Description
<u>Drf FH1</u>	-45.4	7	Formin Homology Region 1
<u>COesterase</u>	818.6	3.1e-243	Carboxylesterase
<u>LETM1</u>	-0.4	7.8	LETM1-like protein
<u>Esterase</u>	-74.4	0.5	Putative esterase
<u>Abhydrolase 3</u>	18.2	0.00011	alpha/beta hydrolase fold

Table 3: Protein domains found in bovine AChE obtained by searching Pfam library database of profiles hidden Markov models

In addition to the carboxylesterases domain, this search identified the presence of a Formin homology region, LETM1 and hydrolase domains. Mammalian Diaphanous-related formins act as Rho small GTPase effectors during growth factor-induced cytoskeletal remodeling and cell division. LETM1 encodes a putative member of the EF-hand family of Ca(2+)-binding proteins. These finding again suggests that AChE is involved in cell signalling, binds with calcium or utilizes the calcium signalling mechanism for its non-catalytic actions since calcium has a well known role in neurite growth as well as neuromodulation. [17]

Search for Related Proteins using sequence alignment

BLAST [13] (Basic Local Alignment Search Tool), provides a method for rapid searching of nucleotide and protein databases. Gaps enable the BLAST algorithm to detect local as well as global alignments efficiently and regions of similarity hidden in otherwise unrelated proteins can be detected. Both types of similarity may provide important clues to the function of uncharacterized proteins. Therefore with bovine AChE sequence were performed local as well as global alignments. An unfiltered Gapped-BLAST search of the unrestricted data base returned 445 hits. The most statistically significant matches were with 11 other cholinesterases (E value = 0). The last hit was that of Esterase-5A precursor (E value = 1e-29). A filtered search

also produced similar results. An important finding here was the retrieval of four different neuroligins including human neuroligin (XP_013051.2) with E value of 5e-78 which is highly significant. Neuroligin is reported to be present in synapses mediating cell adhesion and thus suggesting AChE may play a similar role in synapse formation. [18] The Gapped-BLAST search also identified thyroglobulin and neurotactin. Another biologically interesting protein identified in this search is insect juvenile hormone esterase (E = 3e-36) which plays a significant role during metamorphosis in insects. We then aligned neurotactin, neuroligin and bovine AChE and found three strings of residues showing consensus. Of these three strings of amino acid residues shown below, (2) from residues 123-133 is the carboxylesterase type B motif and (1) from residues 73-82 and (3) from residues 180-188 seem to match no known motifs. The striking consensus found were at (1) GPRRFLPPEP - (bov AChE 73-82); (2) EDCLYLNVTWP - (bov AChE 123-133); (3) NYRVGAFGL - (bov AChE 180-188)

Similarity Search based on Structure

Since 3D structure is strongly conserved in molecular evolution, one can often make valid inferences concerning structure and function by examining the structures of related proteins. Therefore we extended the analysis of bovine AChE to search for similar proteins based on structure at PDB [14]

which is a repository of solved structures of proteins / peptides. Our search yielded 137 protein structures, most of which were cholinesterases or lipases. There was an interesting protein, the fucose binding lectin that had structural similarity to AChE.

Conclusion:

The searches for motifs, finger prints and domains in bovine AChE as well as searches for related sequences all demonstrated the presence of a highly conserved carboxylesterase type B signature from slime molds to human. However the motif searches did provide clues to the neuromodulatory actions of AChE by underscoring its domain identity to GLU 7R, calcium binding motif, connexin and nicotinic receptor. The similarity search using gapped BLAST retrieved biologically significant similar proteins, namely neurotactin, neuroligin and thyroglobulin. Among them, neurotactin and neuroligin are known cell adhesive molecules involved in neurite fasciculation and synaptogenesis. [19] Sequence similarity among neurotactin, Torpedo acetylcholinesterase and neuroligin [20] were reported. Our analyses of bovine AChE show that neurotactin, neuroligin and bovine AChE exhibit three stretches of consensus, of which, one is the well known carboxylesterase B signature. The identity of the other two is not known. Either one or both of these may be involved in the growth factor like action and this needs to be experimentally tested.

Initially it appears logical to search for a motif since short segment of the sequence may have better chances of matching up with similar motif present in other proteins which may otherwise be very different in their sequence composition. Yet the results here show the importance of global as well as local similarity searches since all protein motifs are not yet identified and also due to differences among the data bases. It was evident throughout this study that the search results as well as their significance were highly dependent on the strength of the algorithm, the type of scoring matrix and the size of the data base. This underscores the importance of doing multiple searches and comparison of the results before arriving at any conclusion. The advances in technology and web access have made it possible to access all the different global facilities and programs for these searches reported here. These results demonstrate that a thorough search of data

bases can provide clues to the underlying mechanism of action of a protein which can be followed by experimentation and thus can be a very effective tool for biological and biomedical research.

Acknowledgement:

This research was partially supported by NIH Grant Number P20 RR-16460 from the IDeA Networks of Biomedical Research Excellence (INBRE) Program of the National Center for Research Resources.

References:

- [01] J. Massoulie, *et al.*, *Progress in Neurobiol.*, 41:31 (1993) [PMID: 8321908]
- [02] R. B. Ravelli, *et al.*, *Acta Crystallogr D Biol Crystallogr.*, 54:1359 (1998) [PMID: 10089512]
- [03] B. Peretz, *Netherlands J. Zool.*, 44:395 (1994)
- [04] S. A. Greenfield, *et al.*, *Exp Brain Res.*, 54:513 (1984) [PMID: 6202543]
- [05] M. Srivatsan & B. Peretz, *Neuroscience*, 77:921 (1997) [PMID: 9070763]
- [06] J. W. Bigbee, *et al.*, *Brain Res.*, 861:354 (2000) [PMID: 10760497]
- [07] D. H. Small, *et al.*, *J. Neurosci.*, 15:144 (1995) [PMID: 7823125]
- [08] <http://www.ncbi.nlm.nih.gov/>
- [09] <http://us.expasy.org/prosite/>
- [10] <http://motif.stanford.edu/emotif/emotif-search.html>
- [11] <http://umber.sbs.man.ac.uk/dbbrowser/PRINTS/>
- [12] <http://pfam.janelia.org/hmmsearch.shtml>
- [13] <http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>
- [14] <http://www.rcsb.org/pdb/search/searchSequence.do>
- [15] R. Zwart, *et al.*, *J Neurochem.*, 75:2492 (2000)[PMID: 11080202]
- [16] Y. Tanabe, *et al.*, *Neuron*, 8:169 (1992) [PMID: 1309649]
- [17] C. P. Gilman & M. P. Mattson, *Neuromolecular Med.*, 2:197 (2002) [PMID: 12428811]
- [18] R. A. Philibert, *et al.*, *Gene*, 246:303 (2000) [PMID: 10767552]
- [19] Y. Barthalay, *et al.*, *EMBO J.*, 9:3603 (1990) [PMID: 2120048]
- [20] K. Ichtchenko, *et al.*, *Cell*, 81:435 (1995) [PMID:7736595]

Edited by P. Kangueane

Citation: Srivatsan, Bioinformatics 1(8): 281-284 (2006)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.