Short communication

New friendly tools for users of ESTHER, the database of the α/β-hydrolase fold superfamily of proteins

Ludovic Renault, Vincent Nègre, Thierry Hotelier, Xavier Cousin, Pascale Marchot, Arnaud Chatonnet*

Ingénierie des Protéines CNRS, IFR Jean Roche, Université de la Méditerranée, Marseille, France
Unité Informatique de Centre INRA, INRA-UMII-AgroM, Montpellier, France
Différenciation Cellulaire et Croissance, INRA-UMII-AgroM, Montpellier, France

Available online 17 November 2005

Abstract

The structural α/β-hydrolase fold is characterized by a β-sheet core of five to eight strands connected by α-helices to form a α/β/α sandwich. The superfamily members, exemplified by the cholinesterases, diverged from a common ancestor into a number of hydrolytic enzymes displaying a wide range of substrate specificities, along with proteins with no recognized hydrolytic activity. In the enzymes, the catalytic triad residues are presented on loops of which one, the nucleophile elbow, is the most conserved feature of the fold. Of the other proteins, which all lack from one to all of the catalytic residues, some may simply be ‘inactive’ enzymes while others have been shown to be involved in heterologous surface recognition functions.

The ESTHER (for esterases, α/β-hydrolase enzymes and relatives) database (http://bioweb.ensam.inra.fr/esther) gathers and annotates all the published pieces of information (gene and protein sequences; biochemical, pharmacological, and structural data) related to the superfamily, and connects them together to provide the bases for studying structure–function relationships within the superfamily. The most recent developments of the database are presented.

© 2005 Elsevier Ireland Ltd. All rights reserved.

Keywords: Database; ESTHER; α/β-hydrolase fold; Acetylcholinesterase

1. Introduction

The cholinesterases belong to the superfamily of proteins with a α/β-hydrolase fold, as originally described by Ollis et al. [1]. The canonical α/β-hydrolase fold is composed of 5–11 β-strands, of which one only, strand β2, is antiparallel (Fig. 1). The resulting β-sheet defines the core of the folded protein structure, which is ornamented with loops and helices interspersed among the β-strands in the primary sequence. In the enzymes, the catalytic triad residues are presented on loops, of which one, the nucleophile elbow, is the most conserved feature of the fold. Almost all the catalytically active α/β-hydrolase fold proteins are thought to use a two-step mechanism, based on the catalytic triad residues that are adjacent in the tertiary structure [2,3]. Of the other proteins, which all lack from one to all three of the catalytic triad residues [4–6], some may simply be ‘inactive’ enzymes while others are known to be involved in surface recognition functions [7].

Since the primary description of the superfamily [1], the number of proteins found to adopt the α/β-hydrolase fold has increased dramatically. In 1994, the need for a specialized database dedicated to the superfamily became obvious, and ESTHER (for esterases, α/β-hydrolase enzymes and relatives;
http://bioweb.ensam.inra.fr/esther) was created. The ESTHER server gathers and annotates the published pieces of information (gene and protein sequences; biochemical, pharmacological, and structural data) related to the superfamily and connects them all to provide the bases for structure–function analysis.

Links to original entries in generalist databases (PDB, Pfam, Swiss-Prot, etc.), and kinetics data on natural and mutant enzymes of the cholinesterase and carboxylesterase subfamilies [8–10] are provided. Here, we present four of the new friendly tools recently added to the database: a section on human diseases related to members in the family; profile hidden Markov models (HMMs) that help characterize the subfamilies; a User Basket for selecting, importing and comparing sequences; and a new Query interface for easier use of ESTHER.

2. New friendly tools

2.1. Human genetic disease table (Fig. 2)

Fifteen genetic diseases due to mutations in members of the superfamily have been reported. Of these, mutations in the neuroligin genes have been found to be associated with autism disorders [11] and mutations in the maspardin gene to be important in Mast Syndrome [12]. Mutations in three other genes appear to
**Fig. 3.** Output from the HMMER tool using, as an input, the neuroligin 3 sequence.
be associated with risk factors for three more diseases, and mutations in two more genes are related to increased sensitivity to xenobiotics. These diseases, risk factors and xenobiotics sensitivities are summarized into a distinct table (http://bioweb.ensam.inra.fr/ESTHER/disease.table), and links to bibliography reports, gene mutations and the OMIM database [13] are provided.

2.2. Profile hidden Markov models (Fig. 3)

The $\alpha/\beta$-hydrolase fold subfamilies typically display little sequence homologies. Building profile HMMs is amongst the most successful procedures for detecting remote homology between proteins. HMMER, the profile HMM software suite for making and using HMMs of biological sequences, was used for sensitive database searching and statistical generation of a sequence consensus within each of the ESTHER subfamilies [14]. Users can now align their own sequences or queries with these sequence consensus family profiles and get back a table of hits associated with scores and E-values.

2.3. User Basket (Fig. 4)

A User Basket has been created to customize access to the database, in allowing a user to select ‘gene locus’ items and analyze them in various combinations using links (CLUSTAL W, bibliography report, etc.) also available through the database.

2.4. Database Queries

The original AceDB software initially used to run the database included simple queries based on keywords (Simple search, Text search, Class Browser) and complex queries interconnecting several database objects (Ace Query, AQL Query).

However, users had to learn specific syntax and know the structure of each object to use the database. To make access to the database easier, we implemented the Query builder interface, a step-by-step graphic interface to formulate Ace queries initially developed for the ParaDB database [15]. This interface is dynamically generated and fitted to the selected table model, and users with minimal knowledge of the database structure can formulate queries.

3. Conclusions

The ESTHER server, created 10 years ago, constantly gathers, annotates and connects the published information related to the $\alpha/\beta$-hydrolase superfamily of proteins [8–10]. Recently developed friendly tools include: a Disease Table providing information on diseases, risk factors, and xenobiotics sensitivities due to mutations in genes of the superfamily; consensus HMM profiles of subfamilies for detecting remote homologies between sequences; a User Basket for selecting ‘gene locus’ items and analyzing them in various combinations; and a new Query interface for easier use of the database.

This work was supported by the Génopôle Montpellier Languedoc-Roussillon (to A.C.) and the Programme Bioinformatique inter EPST 2002 (to A.C. and P.M.).

References

[1] D.L. Ollis, E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S.M. Franken, M. Harel, S.J. Remington, I. Silman, J. Schrag, et al., The $\alpha/\beta$-hydrolase fold, Protein Eng. 5 (1992) 197–211.
[2] P. Heikinheimo, A. Goldman, C. Jeffries, D.L. Ollis, Of barn owls and bankers: a lush variety of $\alpha/\beta$ hydrolases, Struct. Fold Des. 7 (1999) R141–R146.
[3] M. Holmgren, $\alpha/\beta$-Hydrolase fold enzymes: structures, functions and mechanisms, Curr. Protein Pept. Sci. 1 (2000) 209–215.
[4] M. Hortich, N.H. Patel, A.J. Bieber, Z.R. Traquina, C.S. Goodman, Drosophila neurotactin, a surface glycoprotein with homol-
ogy to serine esterases, is dynamically expressed during embryogenesis, Development 110 (1990) 1327–1340.

[5] V.J. Auld, R.D. Fetter, K. Brodker, C.S. Goodman, Glutaxin, a novel transmembrane protein on peripheral glia, is required to form the blood–nerve barrier in Drosophila, Cell 81 (1995) 757–767.

[6] K. Ichtchenko, Y. Hata, T. Nguyen, B. Ulrich, M. Miseler, C. Mosmann, C.T. Sudhof, Neuroligin 1: a splice site-specific ligand for β-neurexins, Cell 81 (1995) 435–444.

[7] H. Soreq, S. Sohman, Acetylcholinesterase—new roles for an old actor, Nat. Rev. Neurosci. 2 (2001) 294–302.

[8] A. Chatonnet, T. Hotelier, X. Cousin, Kinetic parameters of cholinesterase interactions with organophosphates: retrieval and comparison tools available through ESTHER database, ESTerases, s/n Hydrolase Enzymes and Relatives, Chem. Biol. Interact. 119–120 (1999) 567–576.

[9] A. Chatonnet, X. Cousin, A. Robinson, Links between kinetic data and sequences in the s/n-hydrolase fold database, Brief Bioinform. 2 (2001) 30–37.

[10] T. Hotelier, L. Renault, X. Cousin, V. Negro, P. Marchot, A. Chatonnet, ESTHER, the database of the s/n-hydrolase fold superfamily of proteins, Nucleic Acids Res. 32 Database Issue (2004) D45–D46.

[11] S. Jamain, H. Quach, C. Betancur, M. Rastam, C. Colineaux, J.C. Gillberg, B. Soderstrom, B. Giro, M. Leboyer, C. Gillberg, T. Bourgeron. Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism, Nat. Genet. 34 (2003) 27–29.

[12] M.A. Simpson, H. Cross, C. Prunakas, A. Pryde, R. Hershberger, A. Chatonnet, M.A. Patton, A.H. Crosby. Masparin is mutated in mast syndrome, a complicated form of hereditary spastic paraplegia associated with dementia, Am. J. Hum. Genet. 73 (2003) 1147–1156.

[13] A. Hamesh, A.F. Scott, J. Amberger, C. Bocchi, D. Valle. McKusick, Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders, Nucleic Acids Res. 30 (2002) 52–55.

[14] S.R. Eddy. Profile hidden Markov models, Bioinformatics 14 (1998) 755–766.

[15] M. Levengr, K. Prat, N. Prenier, D. Birmann, F. Coulter. ParaDB: a tool for paralogy mapping in vertebrate genomes, Nucleic Acids Res. 31 (2003) 63–67.