

## Database of cell signaling enzymes

Ragul Gowthaman\*, Ramya Venugopal and Dhiviya Dayanandam

Department of Bioinformatics, School of Chemical and Biotechnology, Shanmugha Arts Science Technology & Research Academy (SASTRA), SASTRA UNIVERSITY, Tanjore, Tamilnadu, India; Ragul Gowthaman\* - Email:

ragul@bioinfo.sastra.edu; \* Corresponding author

received November 06, 2006; accepted November 21, 2006; published online November 28, 2006

### Abstract:

This paper describes a database for cell signaling enzymes. Our web database offers methods to study, interpret and compare cell-signaling enzymes. Searching and retrieving data from this database has been made easy and user friendly and it is well integrated with other related databases. We believe the end user will be benefited from this database.

**Availability:** <http://www.sastra.edu/dcse/index.html>

**Keywords:** cell signaling; enzymes; web database; pBLAST

### Background:

A cell communicates with its neighbors and environment by sending and receiving information in the form of chemical signals. These signals are converted into intracellular second messenger signals that ultimately make the cells respond appropriately by dividing, moving or even dying. The external signals may enter the cell through enzymes, G-protein coupled receptors, hydrophobic molecules and ion channels. When the receptor sensing the signal is a catalyst such as enzyme, the response is amplified. Thus cell signaling governs the basic cellular activities and coordinates the cell action. Errors in cell signaling is the cause for many serious diseases/disorders including cancer, autoimmune diseases, diabetes etc. By understanding cell signaling we can treat these diseases effectively and potentially. [1] However, no attempts have been made so far to curate and catalog the enzymes involved in cell signaling.

DCSE, the Database of Cell Signaling Enzymes covers a gamut of cell signaling enzymes, which includes kinases, phosphatases, adenylyl cyclases, caspases, phosphodiesterases, phospholipases, prenyltransferases etc. As mentioned above the defect in cell signaling mechanism is the major cause for diseases and hence the cell signaling enzymes are considered to be potential target in rational drug design approach.

### Methodology:

DCSE was developed using MySQL [2], a relational database management system, at the back-end for storing data. The database can be regularly updated. The data for the database were collected from SwissProt [3], the repository of protein sequence information. PHP [4], a widely used general purpose scripting language that is especially suited for web development, was used to design the database interface.

The database can be accessed over the Internet. A screenshot of the home page is shown in Figure 1. The database can be searched by specifying keywords such as name, Enzyme Commission Number and species. For each enzyme in the database a unique identifier called DCSE ID has been assigned. The ID consists of two parts. The first part tells about the enzyme class and the second part indicates the number of that enzyme in that class. The two fields are separated by an underscore (\_). Each enzyme has been defined with its name, Enzyme commission number, and the species from which it has been sequenced. Cross-reference details to other databases are also provided. The functional and other sequence related informations are provided by SwissProt, the domain classification and function is provided by InterPro [5], the protein family classification is provided by Pfam [6] and the structural details are provided by PDB. [7] A hyperlink has been provided to corresponding entry page in the above-mentioned databases.

The sequence is also displayed in raw-text format. One can retrieve the sequence in FASTA format using the 'Retrieve in Fasta format' option available with each entry. An advanced search can also be performed by filling an advanced query form that takes input as DCSE id, name, EC number, Species, other database id such as SwissProt, InterPro, Pfam and PDB. The fields in the search form are joined together by AND operator. BLAST (Basic Local Alignment Search Tool) [8] is an algorithm for comparing biological sequences. Given a library or database of sequences, a BLAST search enables a researcher to look for sequences that resemble a given sequence of interest. A standalone BLAST has been installed and it can be used for similarity search purpose. BLAST can be done against SwissProt, DCSE or PD.

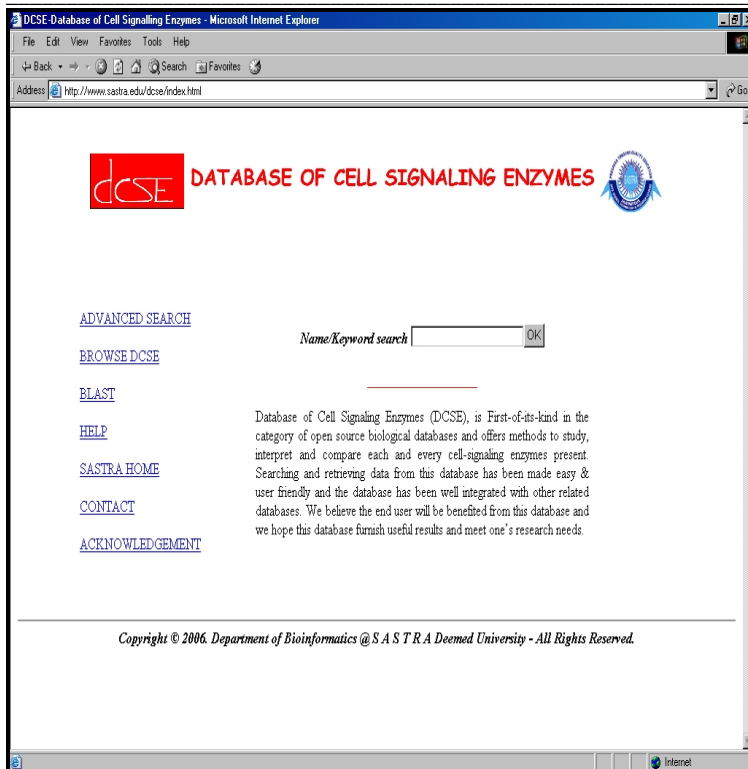


Figure1: DCSE home page

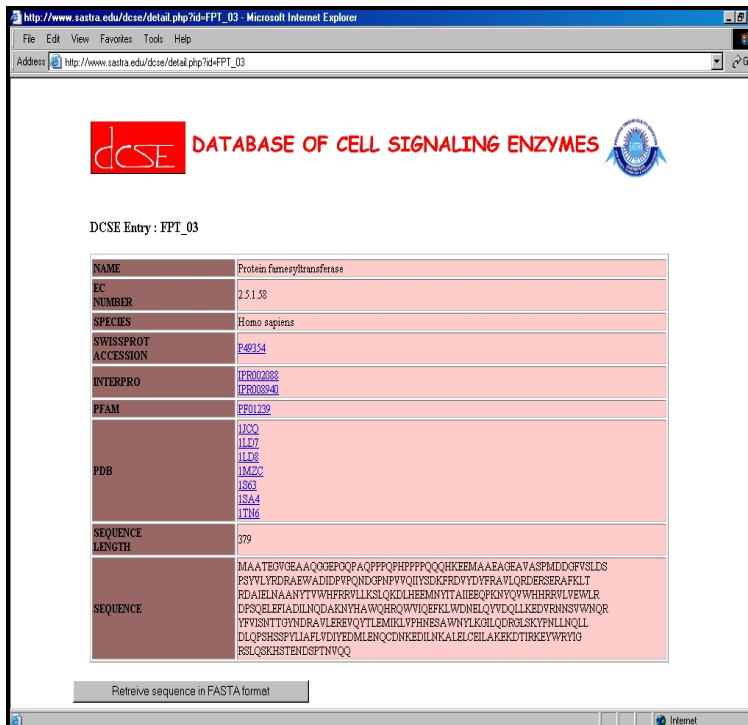


Figure2: An DCSE entry



Figure3: BLAST input page

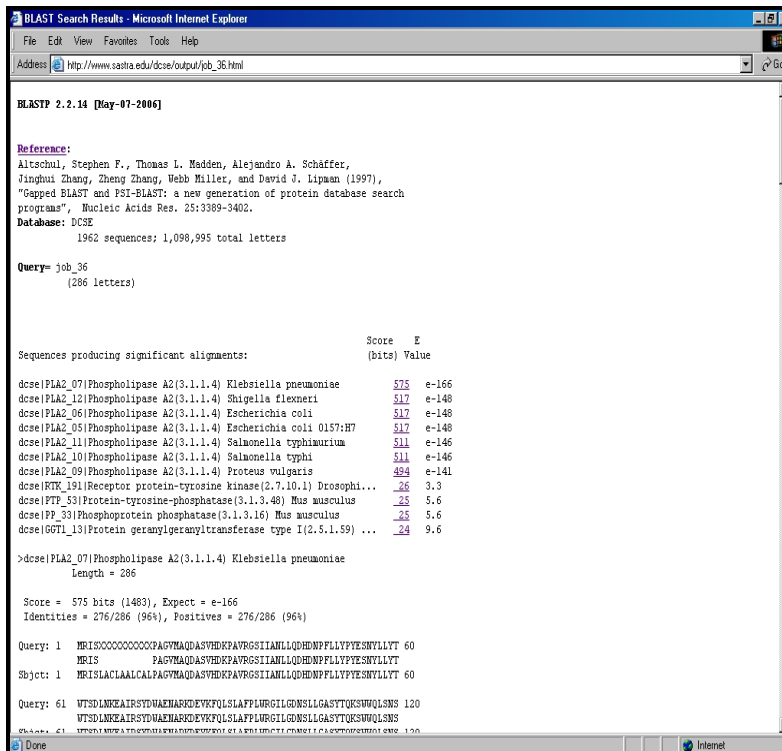


Figure 4: BLAST result page

**Usability/ Accessibility:**

There are three ways in which a user can query the database. The first is the 'keyword search' that can be done by specifying exact or likely keywords such as name, Enzyme Commission Number and species. The second search option is the 'Advanced Search' wherein the user can fill a form by specifying the details for various fields such as DCSE id, Enzyme name, Enzyme commission number, SwissProt accession number, InterPro domain ID, Pfam id or PDB id. The fields are joined together by AND operator. Alternatively the users can also browse the database by the 21 different categories of cell signaling enzymes. Whenever the database is searched the search returns the result with number of hits found for that query along with a summary of details for each entry with its id, name and species. The user can then select the appropriate hit by following the link on it. This displays a page with all the available details of the enzyme in the database such as name, EC number, sequence information and cross-reference to already existing databases as shown in Figure 2. A hyperlink has been provided to corresponding entry page in the cross-referenced databases for easy access. The sequence is displayed in raw-text format. One can retrieve the sequence in FASTA format using the 'Retrieve in Fasta format' option available with each entry. A BLAST search can also be done to search for similar sequences within the database or in general. A link has been provided in the tools page. The BLAST input page and result page are shown in figure 3 and 4.

**Utility to the biological community:**

Since DCSE gather together all the information on biological molecules, sequences, structures, functions, and biological reactions, which transfer the cellular signals, this database has the potential of becoming a major hub of resource for the biological community. This database provides a mechanism by which researchers and students can transform information about interactions between biomolecules into knowledge about a cellular process.

**Caveats:**

The database needs to be updated from time to time as the

data has been obtained from other sources. Despite our efforts to collect information from various sources and check them for consistency, the quality of the information depends heavily on the original source.

**Future development:**

The coiled-coil structure is important for protein interaction. Since the cell signaling also involves protein interaction, a tool to predict coiled-coil structure is of importance.

There are few tools for this purpose like MultiCoil. The Multicoil program just gives the probability for a residue to be in coiled-coil region. We intend to develop a coiled-coil prediction tool based on artificial neural networks that could tell whether a residue is in coiled-coil region if it crosses a threshold computed from a training set consisting of proteins with known coiled-coil structure.

**Acknowledgement:**

The authors are grateful to Dr. K.N. Somasekharan, Dean - School of Chemical & Biotechnology, Dr. S. Swaminathan, Dean - Sponsored Research & Director CeNTAB and Prof. K. Usha Devi, Dean - School of Computing, SASTRA University for their support.

**References:**

- [01] A. Persidis, *Nat Biotechnol.*, 16:1082 (1998) [PMID: 9831041]
- [02] <http://www.mysql.com/>
- [03] <http://www.expasy.org/sprot>
- [04] <http://www.php.net/>
- [05] <http://www.ebi.ac.uk/interpro/>
- [06] <http://www.sanger.ac.uk/Software/Pfam/>
- [07] <http://www.rcsb.org>
- [08] S. F. Altschul, *et al.*, *Nucleic Acids Res.*, 25:3389 (1997) [PMID: 9254694]

Edited by P. Kanguane

Citation: Gowthaman *et al.*, *Bioinformatics* 1(7): 265-268 (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.