DOI: 10.1093/nar/gkh079

STCDB: Signal Transduction Classification Database

Ming Chen*, Susana Lin and Ralf Hofestaedt

Department of Bioinformatics and Medical Informatics, Faculty of Technology, University of Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany

Received August 18, 2003; Revised and Accepted October 3, 2003

ABSTRACT

The Signal Transduction Classification Database (STCDB) is a database of information relative to the classification of signal transduction. It is based primarily on a proposed classification of signal transduction and it describes each type of characterized signal transduction for which a unique ST number has been provided. This document presents, in its first version, the classification of signal transduction in eukaryotic cells. Approved classifications are available for web browsing at http://www.techfak.uni-bielefeld.de/~mchen/STCDB.

INTRODUCTION

Signal transduction is of great interest to many academic and pharmaceutical scientists. Signal transduction pathways regulate basic biological processes, their malfuction can cause disease. With the widespread use of modern techniques in various subfields of biology, more and more cellular data are being accumulated, which has led to a proliferation of information and terminology. The complexity created by the crosstalk in the signal transduction network makes it virtually impossible to infer by hand all the consequences that follow after the modification of one part of the network. Fortunately, a number of databases have been constructed to bring signal transduction knowledge into a well-organized format, providing simple and fast access to the signal transduction system. As examples we mention only: SPAD (http://www.grt.Kyushu-u.ac.jp/spad/) (1), CSNDB (http://geo.nihs.go.jp/csndb/) (2), TransPath (http://www.biobase.de/) (3), LGICdb (http://www.pasteur.fr/recherche/banques/LGIC/LGIC.html) (4), SMART (http://smart.embl-heidelberg.de/) (5), SENTRA (http://www-wit.mcs.anl.gov/sentra/index.html) (6) and Proteome's BioKnowledge Library (http://www.incyte.com) (7). Moreover, signal molecules and pathways are classified and graphically illustrated by BioCarta (http://www.biocarta.com). However, no reference has been made to the classification and nomenclature of transductions involving signal reception, transportation and function.

This document presents classifications concerned with signal transduction in eukaryotic cells, and brings order into a classification recommendation for them. The STCDB is useful to anyone working with signal transduction; moreover it may be helpful in the development of computer programs involved in the manipulation of signalling pathways.

CLASSIFICATION

In general, cellular signal transduction is the mechanism by which a signal encountered at a cell's surface (i.e. an extracellular signal) is transformed into an intracellular signal that in turn invokes cellular responses such as proliferation, differentiation, secretion and apoptosis within the responding cells. The movement of signals can be simple, like that associated with receptor molecules of the acetylcholine class: receptors are also selective channels which, upon ligand interaction, allow signals to be passed in the form of small ion movement, either into or out of the cell. These ion movements result in changes in the electrical potential of the cells, which in turn propagate the signal along the cell. More complex signal transduction involves a complex network of interwoven signalling cascades (e.g. phosphorylation by tyrosine kinases and/or serine/threonine kinases).

A systematic classification scheme is given for the various types of signal transduction and related reactions currently available:

(i) starting with the arrival of a signalling molecule, which is typically a hormone or a neurotransmitter on the cell surface;

(ii) the signalling molecules bind to specific membrane proteins, the receptors, which are activated;

(iii) these receptors activate proteins, which stimulate other proteins in the cytosol;

(iv) the active proteins bind to the transcription factors, which when activated regulate gene expression;

(v) finally, changes in gene expression initiate the biological response of the cell to the original signal.

An important first step toward acquiring understanding of molecular and cellular function is to build systems for organizing and categorizing functions of bioprocesses. Biochemical reactions that are normally catalysed by enzymes can be easily inferred from the enzymes involved. For example, the transformation of L-arginine to L-ornithine is normally catalysed by arginase, EC 3.5.3.1. According to the Classification and Nomenclature of Enzymes (IUBMB recommendation), it is clear that the reaction belongs to the hydrolysation category (EC 3.*.*.*), acting on 'carbon–nitrogen bonds, other than peptide bonds' (EC 3.5.*.*), and so on. A similar strategy is employed to classify signal transduction. Below is an overview listing the recommended

*To whom correspondence should be addressed. Tel: +49 521 1065284; Fax: +49 521 1065288; Email: mchen@techfak.uni-bielefeld.de

classification. A full listing is on the web page: http://www.techfak.uni-bielefeld.de/~mchen/STCDB.

A four-digit ST number $d_1.d_2.d_3.d_4$ denotes a particular signal transduction with main class:

d_1 = location of transduction
d_2 = type of interaction
d_3 = signal molecule's nature
d_4 = ID

The subclass notations are briefly described as:

d_1 = 1: extracellular signal reception events
- d_2 = 1: physical stimulation of receptors
- d_2 = 2: binding with hormones
- d_2 = 3: binding with non-growth factor cytokines
- d_2 = 4: binding with growth factors
- d_2 = 5: binding with neuronal receptors
- d_2 = 6: binding with other ligands

d_1 = 2: plasma membrane transduction events
- d_2 = 1: channel operation
- d_2 = 2: ion channel transduction
- d_2 = 3: G-protein transduction
- d_2 = 4: other Ser/Thr phosphorylation
- d_2 = 5: Tyr phosphorylation
- d_2 = 6: cleavage
- d_2 = 7: others

d_1 = 3: plasma membrane to cytoplasm transduction events
- d_2 = 1: membrane receptor releasing
- d_2 = 2: protein–protein interaction
- d_2 = 3: others

d_1 = 4: intracellular signal transduction events
- …

d_1 = 5: cytoplasm to nucleoplasm transduction events
- …

d_1 = 6: nucleoplasm to nucleoplasm transduction events
- ….

STCDB DESCRIPTION

Data source

The main source for the data in the STCDB comes from the CSNDB. A minor part of the data has been extracted from TransPath and BioCarta as well as from literature. Additionally, a web-based submission form is available for users' contributions.

Database structure

The STCDB contains data for each type of characterized signal transduction for which an ST number has been provided. It includes a short description, a full relational and hierarchical classification structure, and a map of all classified signal transductions. The entries in the database data file (ST number HTML) are structured so as to be legible to humans as well as to computer programs. The general structure of an entry is the following:

ST number
Recommended name
Alternative names (if any)
Reference
Pointers to the CSNDB entry(ies) that correspond to the signal transduction (if any)
Pointers to the BioCarta entry(ies) that correspond to the signal transduction (if any).

Latest data update

STCDB is regularly updated to reflect updates and additions to the classification. We also update the CSNDB and BioCarta pointers, correct eventual errors and complete the information concerning synonyms using the literature. We welcome and encourage any type of feedback.

The latest data release of STCDB was frozen on September 2003. Sequence corrections, mainly frameshift errors, led in most cases to the modification of classification. In a few cases, more dramatic changes, such as merging several entries or adding/removing entries, were required. Furthermore, additional corrections of signal transduction classifications resulted from a revised analysis of source data. Currently STCDB contains over 486 entries/pages, 400 cited references and 700 external hyperlinks. The numbers of entries of the main classes of signal transduction are shown in Table 1.

We would like to encourage users to submit any requests for a new classification via the web-based submission form (http://www.techfak.uni-bielefeld.de/~mchen/STCDB/submit.html) or to contact us directly by email if they have large data sets. Further analyses and database searching would validate every record that is entered in this way. We will also rely on assistance from a number of specialist advisers and communication with the scientific community in general to maintain accuracy.

ACCESS

The database and associated information files are freely accessible for research purposes at the URL http://www.techfak.uni-bielefeld.de/~mchen/STCDB. Suggestions should be directed to M.Chen at mchen@techfak.uni-bielefeld.de. The other authors can be reached at the email addresses slin@cmefcm.uncor.edu and hofestae@techfak.uni-bielefeld.de. We ask that this article be cited in research projects assisted by the use of the STCDB.

Table 1. Summary of signal transduction classification entries in the STCDB latest release (September 2003)

Signal transduction classification	Description (location of signalling events)	Entries
ST 1.*.*.*	Extracellular signal reception events	176
ST 2.*.*.*	Plasma membrane transduction events	53
ST 3.*.*.*	Plasma membrane to cytoplasm transduction events	22
ST 4.*.*.*	Intracellular signal transduction events	201
ST 5.*.*.*	Cytoplasm to nucleoplasm transduction events	4
ST 6.*.*.*	Nucleoplasm to nucleoplasm transduction events	31

FUTURE PROSPECTS

We plan to add more entries by tracking the development of signal transduction research. The full list of classifications and links will be continuously updated. Despite all the care taken in the construction of STCDB, mistakes in the data and defects in the program may still remain. We welcome information and comments about this classification system, as well as possible improvements concerning the WWW server. The participation of any academic institute in the future development of STCDB is also encouraged.

ACKNOWLEDGEMENTS

This work is supported in part by the German Research Foundation (DFG) and the German Academic Exchange Service (DAAD).

REFERENCES

1. Tateishi,N., Shiotari,H., Kuhara,S., Takagi,T. and Kanehisa,M. (1995) An integrated database SPAD (Signaling PAthway Database) for signal transduction and genetic information. *Genome Informatics Workshop*, *GIW95*, December 11–12 Yokohama, Japan.
2. Takai-Igaraschi,T. and Kaminuma,T. (1999) A pathway finding system for the cell signaling networks database. *In Silico Biol.*, **1**, 129–146.
3. Krull,M., Voss,N., Choi,C., Pistor,S., Potapov,A. and Wingender,E. (2003) TRANSPATH®: an integrated database on signal transduction and a tool for array analysis. *Nucleic Acids Res.*, **31**, 97–100.
4. Le Novère,N. and Changeux,J.-P. (2001) LGICdb: the ligand-gated ion channel database. *Nucleic Acids Res.*, **29**, 294–295.
5. Letunic,I., Goodstadt,L., Dickens,N.J., Doerks,T., Schultz,J., Mott,R., Ciccarelli,F., Copley,R.R., Ponting,C.P. and Bork,P. (2002) Recent improvements to the SMART domain-based sequence annotation resource. *Nucleic Acids Res.*, **30**, 242–244
6. Malsev,N., Marland,E., Yu,G.X., Bhatnagar,S. and Lusk,R. (2002) Sentra, a database of signal transduction proteins. *Nucleic Acids Res.*, **30**, 349–350.
7. Hodges,P.E., Carrico,P.M., Hogan,J.D., O'Neill,K.E., Owen,J.J., Mangan,M., Davis,B.P., Brooks,J.E. and Garrels,J.I. (2002) Annotating the human proteome: the Human Proteome Survey Database (HumanPSD) and an in-depth target database for G protein-coupled receptors (GPCR-PD) from Incyte Genomics. *Nucleic Acids Res.*, **30**, 137–141.