Biochimica et Biophysica Acta 1798 (2010) 1908–1912

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbamem

YTPdb: A wiki database of yeast membrane transporters

Sylvain Brohée [a,1], Roland Barriot [a,b,c,2], Yves Moreau [a,3], Bruno André [d,*]

[a] ESAT-SCD, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
[b] Université de Toulouse, UPS, Laboratoire de Microbiologie et Génétique Moléculaires, F-31000 Toulouse, France
[c] Centre National de la Recherche Scientifique, LMGM, F-31000 Toulouse, France
[d] Physiologie Moléculaire de la Cellule, Institut de Biologie et de Médecine Moléculaires, Université Libre de Bruxelles, Rue des Pr, Jeener et Brachet, B-6041 Gosselies, Belgium

ARTICLE INFO

Article history:
Received 19 February 2010
Received in revised form 17 May 2010
Accepted 7 June 2010
Available online 19 June 2010

Keywords:
yeast
transporter
database

ABSTRACT

Membrane transporters constitute one of the largest functional categories of proteins in all organisms. In the yeast *Saccharomyces cerevisiae*, this represents about 300 proteins (~5% of the proteome). We here present the Yeast Transport Protein database (YTPdb), a user-friendly collaborative resource dedicated to the precise classification and annotation of yeast transporters. YTPdb exploits an evolution of the MediaWiki web engine used for popular collaborative databases like Wikipedia, allowing every registered user to edit the data in a user-friendly manner. Proteins in YTPdb are classified on the basis of functional criteria such as subcellular location or their substrate compounds. These classifications are hierarchical, allowing queries to be performed at various levels, from highly specific (e.g. ammonium as a substrate or the vacuole as a location) to broader (e.g. cation as a substrate or inner membranes as location). Other resources accessible for each transporter via YTPdb include post-translational modifications, K_m values, a permanently updated bibliography, and a hierarchical classification into families. The YTPdb concept can be extrapolated to other organisms and could even be applied for other functional categories of proteins. YTPdb is accessible at http://homes.esat.kuleuven.be/ytpdb/.

1. Introduction

The *membrane transporter* functional class of proteins is typically one of the most abundantly represented in cellular proteomes. The estimated percentage of predicted membrane transporters ranges from 2% to 10% in bacteria and 2% to 7% in eukaryotes. It is 2.7% in the human species [1]. These proteins facilitate the transport of a very wide variety of small compounds (small ions, organic compounds, short peptides, water, etc.) across the lipid bilayers of the plasma and internal membranes. Most inventoried transporters belong to families conserved throughout evolution [2].

Among the 5690 protein-encoding genes of the yeast *Saccharomyces cerevisiae* [3], almost 300 code for established or predicted transmembrane transporters [4]. *Established* transporters are those for which experimental evidence of a direct role in transmembrane transport of specific compounds is available. *Predicted* transporters share significant sequence similarity with at least one experimentally characterized transporter from any organism, but their functions remain unknown: although the subcellular locations of many of them have been determined by large-scale analysis [5], the compounds they recognize have not been identified and their roles in cell metabolism thus remain unknown.

Specialized databases of membrane transporters have previously been implemented (links accessible via YTPdb). For instance, the Transport Classification Database (TCDB) provides a classification system for membrane transporters of all organisms into no less than 550 families [2], and the related TransportDB lists all transporters predicted from completely sequenced genomes [1]. Other databases categorize transporters or membrane proteins of a single or a small number of organisms, e.g. the Yeast Transport Information (YETI) [6], the Yeast Membrane Protein Library (YMPL) [7], or the Plant Membrane protein Database (Aramemnon) [8]. There also exist databases supplying information about one type of transporters in different organisms (e.g. P-type ATPases [9]). Lastly, information about membrane transporters is also accessible through databases collecting data on the complete sets of genes and proteins of specific organisms, e.g. the Saccharomyces Genome Database (SGD) [10] or the Comprehensive Yeast Genome Database (CYGD) [11].

Yet molecular biologists working in the membrane transport field also need databases providing classification systems based on the main functional properties of listed proteins, e.g. their substrate compounds and subcellular locations. Furthermore, these databases should ideally include annotations about the affinity constants (K_m), membrane topology, post-translational modifications, as well as an

---

[*] Corresponding author. Tel.: +32 2 6509958.
E-mail addresses: Sylvain.Brohee@esat.kuleuven.be (S. Brohée), roland.barriot@biotoul.fr (R. Barriot), yves.moreau@esat.kuleuven.be (Y. Moreau), bran@ulb.ac.be (B. André).
[1] Tel.: +32 16 328643.
[2] Tel.: +33 5 61335821.
[3] Tel.: +32 16 328645.

0005-2736/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbamem.2010.06.008

up-to-date bibliography. It seems obvious that such goal is only achievable if this database can be updated through a user-friendly interface by several users having expertise in the field. We have applied these criteria to implement the Yeast Transport Protein database (YTPdb: http://homes.esat.kuleuven.be/ytpdb/), a web-accessible database devoted to the precise classification and annotation of *S. cerevisiae* membrane transport proteins. Besides its main features described below, YTPdb is presented in the form of a wiki, i.e. on the model of the well-known web encyclopedia Wikipedia. Any registered user can thus freely edit data in the database and add or remove data. We think that the YTPdb concept is extensible to other organisms and even to other functional classes of proteins.

2. Results

2.1. YTPdb content

The Yeast Transporter Protein database provides manually annotated information on 299 yeast proteins classified as *established* (199) or *predicted* (100) membrane transporters (see Introduction). These 299 proteins correspond to the 337 proteins predicted to be transporters by the TransportTP server [12] from which 38 have not been considered because they correspond to subunits of protein translocation systems, of the mitochondrial ATP synthase, of the vacuolar H^+ pump, to non-membrane proteins containing the ABC (ATP-binding cassette) motif, and to several enzymes (e.g. HMG-CoA reductases, lipid modification enzymes, etc.) and membrane proteins which may not be classified as *established* or *predicted* transporters. For each protein, YTPdb provides a page including a short description, the list of substrates, and the subcellular location (Fig. 1). Clicking on a substrate gives the list of all proteins transporting this compound and clicking on a protein's location gives the list of proteins colocalized in this cell compartment. An originality of YTPdb is that functional classification criteria such as subcellular location and recognized substrates are organized in a tree-like fashion. For instance, a *glucose* transporter is also classified as a transporter of *hexose*, *monosaccharide*, and *carbohydrate* and a *cytosine* transporter as a protein transporting *pyrimidine* and *nucleobase*. Similarly, a protein localizing to the *Golgi compartment* is also classified as a protein present in the *Golgi and endosome membrane system*, in the *secretory pathway membrane system*, and in *internal membranes*. It is thus possible – using the "Browse by substrate", "Browse by substrate categories", and "Browse by subcellular location" query tools – to retrieve proteins on the basis of very specific or broader criteria. For instance, the user can list all calcium or leucine transporters but also those for cations, amino acids, etc. Similarly, the user can view all transporters present in the "late endosome" or in the less defined "Golgi and endosomal membrane" system. This mode of classification is convenient in the case of proteins whose properties have been characterized with only limited precision. For instance, the P-type ATPase Neo1 has been located in the Golgi and/or endosomal membrane [13,14]. YTPdb contains a total list of 315 chemical compounds (plus 65 synonyms) and 28 subcellular locations, corresponding to the branch extremities of the tree-like classification schemes. The affinity constants (K_m) for substrates (136 values in the May 2010 version of YTPdb) with the corresponding references are also accessible.

Transporters encoded in YTPdb are also classified according to the transport classification (TC) system [2] (Fig. 1). Using the "Browse by TC class" option, proteins at any level of the TC classification can be listed, e.g. all *channels/pores* (22 proteins) or the *Major Intrinsic Protein (MIP) Family* (4 proteins). We have developed another classification system (YTPdb classification) including more subdivisions and more suited to yeast transporters. For instance, P-type ATPases (16 proteins in *S. cerevisiae*) form a unique class of proteins in TC, whereas they are subdivided into six subfamilies in the YTP classification. The YTP classification also provides a multiple sequence alignment computed by the Muscle algorithm [15].

An important feature of transporters is the topology they adopt in the membrane. For each protein, YTPdb gives access to a web page recapitulating the data of topology prediction generated by seven distinct algorithms.

Many protein databases do not provide an updated bibliography for protein entries. In YTPdb, two systems have been implemented to associate each transporter with references. The "All references" option is based on an editable Boolean formula retrieving references from the PubMed.gov server, and the corresponding bibliography is thus permanently updated. Through the "Curated references" option, the user can generate a list of references classified into several categories. For instance, "Early studies" provides references about studies preceding the molecular characterization of the transporter gene, "Initial molecular characterization, general properties and function" include references about the functional characterisation of the transporter. Other categories (eight in all) list references about "Intracellular trafficking and its regulation", "Gene expression and its regulation", "Transport activity (including influence of mutations) and its regulation", etc. Curated lists of references are currently available for about 40 transporters. As mentioned below, registered curators are invited to associate references with their favorite transporters.

Post-translational modifications (PTMs) play a central role in regulation of protein function. This is well illustrated for yeast transporters, e.g., ubiquitylation typically targets these proteins for degradation in the vacuole [16] and many yeast transporters have been shown to be phosphorylated. A special effort has been devoted to including in YTPdb all known PTMs of yeast transporters deduced from studies centered on specific proteins and from large-scale mass spectrometry analyses [17–25]. In the current version of YTPdb (May 2010), 764 unique PTMs have been encoded for 123 transporters. For each protein, these PTMs are presented in a recapitulative table with references, and their relative positions are indicated in a figure showing the predicted transmembrane regions [26], global topology, and positions of conserved PFAM domains [27] of the protein (Fig. 2). This convenient presentation facilitates comparisons between proteins of the same family and detection of possible recurrent PTM profiles. For instance, many amino acid permeases are palmitoylated at their C-termini [28], and their cytosolic N-terminal tails often contain several phosphorylated serines and threonines as well as ubiquitylated lysines.

YTPdb also associates each transporter with direct links to well-known databases (MIPS [29], Ensembl [30], SGD [10]) to a repository of gene expression data (ArrayExpress) [31] and to the Yeast Gene Order Browser [32]. The latter is an online tool for visualizing the syntenic context of any gene from twelve hemiascomycete species; this is particularly useful for detecting the many pairs of transporter genes deriving from the whole genome duplication event that occurred in an ancestor of the *S. cerevisiae* lineage [33,34].

2.2. YTPdb is a collaborative database

There is no longer any need to introduce Wiki technology. The past few years have seen an explosion in the number of wiki-based web sites, ranging from very general (e.g. the well-known Wikipedia, wiktionnary) to more specialized or biological (wikiproteins [35], wikipathways [36], wikipedia for genes [37], etc). YTPdb is a wiki based on an extension of MediaWiki technology (running Wikipedia). This implies that almost all data in the database can be freely edited and extended by any registered user. For instance, specific and readily editable forms are available for adding K_m values, post-translational modifications, and references and for editing the basic information (substrate compounds, subcellular location, gene synonyms, Boolean formula to retrieve references, etc.) on each transporter. A similar form is available for adding or editing substrate compounds