Drug Discovery Today: Technologies
Vol. 12, 2014

Transporter assays

Transporter taxonomy – a comparison of different transport protein classification schemes

Michael Viereck[1], Anna Gaulton[2], Daniela Digles[1,*], Gerhard F. Ecker[1]

[1]University of Vienna, Department of Pharmaceutical Chemistry, Althanstrasse 14, 1090 Vienna, Austria
[2]European Molecular Biology Laboratory – European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK

Currently, there are more than 800 well characterized human membrane transport proteins (including channels and transporters) and there are estimates that about 10% (approx. 2000) of all human genes are related to transport. Membrane transport proteins are of interest as potential drug targets, for drug delivery, and as a cause of side effects and drug–drug interactions. In light of the development of Open PHACTS, which provides an open pharmacological space, we analyzed selected membrane transport protein classification schemes (Transporter Classification Database, ChEMBL, IUPHAR/BPS Guide to Pharmacology, and Gene Ontology) for their ability to serve as a basis for pharmacology driven protein classification. A comparison of these membrane transport protein classification schemes by using a set of clinically relevant transporters as use-case reveals the strengths and weaknesses of the different taxonomy approaches.

Section editors:
Gerhard F. Ecker – University of Vienna, Vienna, Austria.

Introduction

In biology, taxonomy is the practice and science of classifying things or concepts. Most often this is used for organisms, which can be classified into different taxa. Relationships between these groups can be used to build up a hierarchical classification, for example a phylogenetic tree which shows the evolutionary relationship of organisms. Similarly proteins can be classified, with one of the most prominent examples being the Enzyme Commission (EC) number classification scheme [1]. This classifies enzymes according to the chemical reactions they catalyze: enzymes that are diverse in sequence can be assigned to the same EC number if they participate in the same reaction, thus providing a functional grouping of proteins. Another example is the standard protein kinase classification scheme [2], which groups kinases according to substrate specificity, sequence similarity and functional similarity. In the framework of Open PHACTS [3,4] the implementation of the EC-classification already allows browsing of the hierarchy to retrieve pharmacology data for a whole family of enzymes. Of course it would be desirable to extend this feature also to other target classes, such as receptors and membrane transport proteins to allow the user to browse for targets, but also return classifications for targets of interest.

Transport is an essential feature of every living cell. Therefore, membrane transport proteins are of great interest in life sciences and in drug development. According to counts by Hediger *et al.* [5], based on data from the HUGO Gene Nomenclature Committee (HGNC), more than 800 human membrane transport proteins exist. They found 395 human protein coding SLC genes, 315 ion channels and ionotropic

*Corresponding author.: D. Digles (daniela.digles@univie.ac.at)

 http://dx.doi.org/10.1016/j.ddtec.2014.03.004

receptor genes, 48 ABC genes and 68 transport-related ATPases in HGNC. In drug development, membrane transport proteins are of interest as drug targets, for pharmacokinetics and drug delivery, and as a cause of adverse drug effects and drug–drug interactions [6]. Throughout this review, the phrase membrane transport proteins will be used, according to the Transporter Classification Database (TCDB) [7] definition, to summarize channels, transporters and accessory factors involved in transport. We will present and compare several membrane transport protein classification schemes, focusing on transporters of practical use in drug development. At the moment, eleven transporters are of interest to regulatory bodies such as the Food and Drug Administration [8] and the European Medicines Agency [9], because they have been shown to be involved in clinically significant drug–drug interactions (DDIs) [10]. We will use these transporters as a use-case to compare different available membrane transport protein classifications.

Available membrane transport protein classifications

Currently, there are a large number of membrane transport protein collections available on the web (Table 1 for a short overview). However, most of these collections do not provide their own complete classification, but use an already existing one. In addition, many sources focus on a specific group of membrane transport proteins. Collections like the Solute Carrier (SLC) tables on the BioParadigms webpage (SLC tables; URL: http://slc.bioparadigms.org/) provide excellent information about human solute carriers [5], and other web sites present a good overview of ABC-transporters. While, many of the human transporters important for drug-discovery are ABC-transporters or solute carriers, there are also human ATPases with medical relevance (e.g. Copper-transporting ATPase 2 is involved in Wilson's disease [11]).

Therefore, we concentrate here on comprehensive classifications that are able to classify all types of human membrane transport proteins. The focus of our work is on the classification systems implemented in the Transporter Classification Database (TCDB), the IUPHAR/BPS Guide to Pharmacology (previously Guide to Receptors and Channels, GRAC) of the British Pharmacological Society (BPS) [12], the ChEMBL bioactivity database [13] and the Gene Ontology (GO) [14]; all of which are capable of classifying the majority of human membrane transport proteins and could be added to the Open PHACTS Discovery Platform to allow the retrieval of pharmacology data for a whole family of membrane transport proteins.

TCDB

The Transporter Classification (TC) system, developed in the laboratory of Saier, is to some extent equivalent to the Enzyme Commission (EC) classification, using similar notation. But while the EC system, which was designed before comprehensive sequence data was available, concentrates on the function of enzymes (with each number specifying an enzyme-catalyzed reaction), the membrane transport proteins in the TC system are classified according to both function and phylogeny. The Transporter Classification number considers protein homology; therefore the 48 human ABC-proteins, for example share the same class (3: Primary Active Transporters), subclass (3.A: P-P-bond-hydrolysis-driven transporters) and family (3.A.1: The ATP-binding Cassette (ABC) Superfamily), even though some of the ABC-proteins function as ion channels (ABCC7(CFTR) and ABCC9(SUR2)), channel regulators (ABCC8(SUR1)) or as translation regulators (ABCE1 and ABCF1).

The scheme is recommended by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB; URL: http://www.chem.qmul.ac.uk/iubmb/mtp/) and the classification is also included in several other databases such as UniProt (UniProt; URL: http://www.uniprot.org/) and the Protein Data Bank (PDB; URL: http://www.rcsb.org/pdb/). It is an exhaustive classification with approximately 750 transport protein families, and more than 10,000 non-redundant proteins included in the Transporter Classification Database (TCDB; URL: http://www.tcdb.org) [7]. TCDB was created with the intention of covering all membrane transport systems (channels, transporters, accessory factors) and not being limited to one, or to a particular organism [15]. From a drug-discovery point of view, it therefore includes a large number of less relevant proteins, but in contrast to purely mammalian transport protein classifications, it also contains bacterial [16,17] or protozoan transporters [18], that are significant for drug therapy. The database stores sequence and information regarding all classified transport proteins. Membrane transport proteins can be found in the database by browsing the classification or by directly searching for the protein of interest (e.g. by UniProt accession number). In addition, the user can perform a BLAST search to find a potential classification for a currently unclassified protein of interest.

The classification scheme contains five levels. The first level number indicates the class of transport protein, for example 'Channels/Pores' or 'Primary Active Transporters'. Interestingly, 'Accessory Factors Involved in Transport' are also included. Classes six and seven are currently empty and reserved for yet undiscovered types of transport and class nine contains 'Incompletely Characterized Transport Systems'. Next, a letter indicates the transport protein subclass (e.g. energy source for primary active transport), followed by a number for the transport protein family or superfamily. The assignment of a transport protein to a specific family follows strict statistical criteria of homology, requiring comparison over a region of at least 60 residues and a probability of 10^{-19} or less that this degree of sequence similarity occurred by coincidence [19]. The fourth level indicates the transport