1
Insights into Transporter Classifications: an Outline of Transporters as Drug Targets

Michael Viereck,[1] Anna Gaulton,[2] and Daniela Digles[1]

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

1.1
Introduction

Classifications are a useful tool to get an overview of a topic. They show instances grouped together that share common properties according to the creator of the classification, and as in the case of hierarchical classifications, they allow to draw conclusions on the relation of different classes. In the case of the identification of drug targets (including transporters as the main drug target), several publications show classifications as a helpful tool.

Imming *et al.* [1] categorized drugs according to their targets, to get an estimate on the number of known drug targets, including channels and transporters. Drugs on the market were connected to a target only if it was described as the main target in the literature. For transport proteins (including uniporters, symporters, and antiporters), this identified six different types of transporter groups that are relevant as drug targets. These are the cation-chloride cotransporter (CCC) family (SLC 12), Na^+/H^+ antiporters (SLC 9), proton pumps, Na^+/K^+ ATPases, the eukaryotic (putative) sterol transporter (EST) family, and the neurotransmitter/Na^+ symporter (NSS) family (SLC 6). These families mainly belong to either ATPases or solute carriers (SLC). Whether the EST family is treated as transport protein or not depends on the classification used.

Rask-Andersen *et al.* [2] used a manually curated and extended version of the DrugBank [3] data from 2009 to analyze drug targets. They identified 435 therapeutic effect-mediating targets, where the third largest group (67) is of transporter proteins (including 35 ion channels). These transport proteins are mainly targeted by antihypertensive drugs, diuretics, anaesthetics, and antiarrhythmic drugs. In a more recent study [4], they analyzed the Drugs in the Clinical Trials Database by CenterWatch to investigate the targets of new drug candidates.

Transporters as Drug Targets, First Edition. Edited by Gerhard F. Ecker, Rasmus P. Clausen, and Harald H. Sitte.

Transporter proteins in this data set were classified using the transporter classification (TC) system established by Saier *et al.* [5].

This chapter will first give an overview on selected existing transporter classifications and then describe our process of creating a combined classification scheme for the ChEMBL [6] database. Finally, the investigation of the counts of drugs and diseases for one example protein superfamily is provided, to show the usefulness of classifications in characterizing related proteins and to give a first overview on the topic of this book, focusing on transporters as drug targets.

1.2 Available Transporter Classifications

As a consequence of the significance of the transport process for every living organism, already numerous classification schemes for membrane transport proteins of several organisms exist. Quite a few focus only on specific families, for instance, the SLC superfamily or the ABC transporter superfamily. Table 1.1 shows a selection of membrane transport protein databases with a focus on human proteins. To keep the table concise, databases focusing on different organisms (e.g., the plant membrane protein database Aramemnon [7], the yeast transport protein database YTPdb [8], or ABCdb [9], a database about bacterial ABC systems) are not included, even though bacterial or protozoal channels and transporters can also be promising drug targets [10,11].

We were interested in classification schemes that not only try to cover the full variety of human membrane transport proteins but also provide their own complete classification. Therefore, we analyzed the functional and phylogenetic classification scheme of the Transporter Classification Database (TCDB), the more pharmacology-driven IUPHAR/BPS classification, and the mainly functional-driven classification of channels and transporters in the bioactivity database ChEMBL-16 [18]. These were recently reviewed in [19], which provides a more detailed discussion for the interested reader. In addition, we included the SLC series [12], which is a well-known nomenclature system.

1.2.1 TCDB

The transporter classification system developed in the laboratory of Saier is to some extent comparable to the Enzyme Commission (EC) classification. But while the EC system concentrates on the function of enzymes, the transporters in the TC system are classified according to function and phylogeny [20]. Both schemes are recommended by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology [21,22] (URL: http://www.chem.qmul.ac.uk/iubmb/mtp/). The TC is also included in several other databases such as UniProt [23] (URL: www.UniProt.org/) or the Protein Data Bank (PDB) [24] (URL: http://www.rcsb.org/pdb/). TCDB is an exhaustive classification

Table 1.1 Selection of transporter collections with a focus on human membrane transport proteins.

Database	Description	URL	Included proteins	Limited to organisms
Bioparadigms SLC Series [12]	Online resource of the 52 human solute carrier families	slc.bioparadigms.org/	Solute carriers	*Homo sapiens*
ChEMBL [6]	Large-scale bioactivity database for drug discovery	https://www.ebi.ac.uk/chembldb/target/browser	Proteins with bioactivity data	—
Human ABC transporters [13]	Basic information about human ABC transporters	www.nutrigene.4t.com/humanabc.htm	ABC transporters	*Homo sapiens*
IUPHAR/BPS Guide to PHARMACOLOGY [14]	Overview of human drug targets with their pharmacology	www.guidetopharmacology.org	Drug targets	*Homo sapiens*, and *Mus musculus*, *Rattus norvegicus*
TCDB [5]	Provides a classification scheme for all membrane transport proteins in all living organism	www.tcdb.org/	Channels, transporters, auxiliary transport proteins	—
TransportDB [15]	Genomic comparison of membrane transport proteins, prediction of their function, and classification according to TCDB	www.membranetransport.org/index.html	Channels, transporters, auxiliary transport proteins	Complete genome-sequenced organisms (365 total)
Transporter substrate database (TSdb) [16]	Provides a transporter substrate repository with mappings to KEGG pathways	http://tsdb.cbi.pku.edu.cn/home.cgi	Transporters	—
UCSF-FDA TransPortal [17]	Provides information on transporters important in the drug discovery process	http://bts.ucsf.edu/fdatransportal/	Transporters	*Homo sapiens*

including over 750 transporter families, and over 10 000 protein sequences are included in the Transporter Classification Database [5] (URL: www.tcdb.org/). The database stores sequences and information regarding all classified transport proteins. Transporters can be found in the database either by browsing the classification or by directly searching for the protein of interest (e.g., by UniProt accession number). For unclassified proteins, similar sequences can be found using a BLAST search.

The classification scheme contains five levels as exemplified for 2.A.22.6.3 in Table 1.2. The first level number indicates the class of membrane transport protein. This can be, for example, a channel or primary active transporter. Interestingly, also accessory factors involved in transport are included. Classes 6 and 7 are currently empty and serve as placeholders for yet undiscovered types of transport and class 9 contains not yet fully characterized transporters. Next, a letter indicates the transporter subclass (e.g., energy source for primary active transport), followed by a number for the transporter family or superfamily. The assignment of a transporter 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 than this degree of sequence similarity occurred by coincidence [25]. The fourth level indicates the transporter subfamily and the last level classifies a transport system according to its substrate or range of substrates. To summarize, the first two levels describe the function of the transporter, the next two classify according to phylogenetic similarity, and the last one defines the substrate or indicates the belonging to a transport system.

Table 1.2 Transport system 2.A.22.6.3 as an example of the TCDB classification.

TCDB level	TC number of the level	TCDB name of the exemplary level
Transporter class (level 1, functional):	2	Electrochemical potential-driven transporters
Transporter subclass (level 2, functional):	2.A	Porters (uniporters, symporters, and antiporters)
Transporter family/superfamily (level 3, phylogenetic):	2.A.22	The neurotransmitter:sodium symporter family
Transporter subfamily (level 4, phylogenetic):	2.A.22.6	No explicit level name
TCDB level 5 (examples for the transport system having the same substrate):	2.A.22.6.3	No explicit level name, given examples: • Sodium-dependent neutral amino acid transporter B(0)AT1 (human) • Sodium-dependent neutral amino acid transporter B(0)AT1 (mouse) • Transmembrane protein 27 aka TMEM27

1.2.2 IUPHAR/BPS

The Guide to PHARMACOLOGY database and web page (IUPHAR/BPS Guide to PHARMACOLOGY; URL: www.guidetopharmacology.org/) is created by cooperation between the British Pharmacological Society (BPS) and the International Union of Basic and Clinical Pharmacology (IUPHAR). Originally, they provided on their web page access to their two independent databases, the BPS Guide to Receptors and Channels [26] and the IUPHAR database [27]. Since 2014, the IUPHAR database is included in the BPS database and the web page was renamed from BPS Guide to Receptors and Channels to IUPHAR/BPS Guide to PHARMACOLOGY [28].

The IUPHAR/BPS Guide to PHARMACOLOGY is an expert-driven collection of pharmacological targets and the substances that act on them. It contains several different sections, including G protein-coupled receptors (GPCRs), ion channels, nuclear hormone receptors (NHRs), kinases, catalytic receptors, enzymes, other protein targets, and transporters.

The transporters are divided into the ATP-binding cassette family, F-type and V-type ATPases, P-Type ATPases, major facilitator superfamily (MFS) of transporters, and the SLC superfamily of solute carriers. The nomenclature follows mainly the HGNC gene families. In addition to the standard SLC nomenclature (SLC family 1–52), some of the SLC families are further divided according to commonality of the substrate.

The channels are divided into voltage-gated ion channels, ligand-gated ion channels, and other ion channels.

1.2.3 ChEMBL-16 and ChEMBL-18

The ChEMBL database [6] (URL: https://www.ebi.ac.uk/chembl/) provides large-scale bioactivity data, linking small molecules to the protein targets through which they exert their effects. In order to facilitate browsing and analysis of these data, it was necessary to provide a protein family classification system within the database.

The emphasis of the ChEMBL-16 classification was on a functional rather than a sequence-based classification. Since the ChEMBL-16 transporter hierarchy was heavily focused on protein function, inclusion of new proteins was a largely manual process, relying on the availability of significant knowledge around these proteins. However, the exact transport mechanism (e.g., antiporter, uniporter, Na-symporter, or H-symporter) for a number of transporters, such as the OATPs, is unknown or not completely understood [29]. This makes it difficult to include them in this classification scheme.

Therefore, ChEMBL decided, starting with ChEMBL-18 [30], to move to a more phylogenetic classification that is easier to maintain. The classification is derived mainly from IUPHAR/BPS and TCDB. The idea and schema of this combined classification is described below and depicted in Table 1.5.

1.2.4 SLC Series

In the 1990s, Matthias A. Hediger developed the nomenclature of solute carrier families in collaboration with Phyllis McAlpine [31]. The HUGO Gene Nomenclature Committee (HGNC, available from www.genenames.org) included this nomenclature for gene names of classical membrane transport proteins. Although originally introduced for human genes, the term is sometimes used for nonhuman species as well.

A collection of this 52 family-containing series is available from www.bioparadigms.org. Although the number of included transporters is limited to a specific type of membrane transport proteins, available reviews for each family by experts [12], and manually curated information on transport type, substrates, and expression, make this collection a valuable resource.

Members within an individual SLC family share more than 20% sequence similarity with each other, but the homology between the 52 families is often quite low or nonexistent [32]. The SLC members are treated in different ways in other classifications. IUPHAR/BPS summarizes the SLC families in a superfamily. Due to the fact that SLC families have only a vague definition in common (membrane transport proteins that are not driven by ATP) and their sometimes missing sequence similarity, TCDB has no single class that contains all SLC families. SLC families that are not found in class 2 of TCDB (electrochemical potential-driven transporters) are given in Table 1.3.

Table 1.3 SLC members counterparts in TCDB that do not belong to TCDB class 2: electrochemical potential-driven transporters.

SLC family	TC class	TC family
SLC42 Rh ammonium transporter family	Channels/pores	1.A.11 the ammonia transporter channel (Amt) family
SLC41 MgtE-like magnesium transporter family	Channels/pores	1.A.26 the Mg2+ transporter-E (MgtE) family
SLC14 urea transporter (UT) family	Channels/pores	1.A.28 the urea transporter (UT) family
SLC31 copper transporter family	Channels/pores	1.A.56 the copper transporter (Ctr) family
SLC27 fatty acid transporter family	Group translocators	4.C.1 the proposed fatty acid transporter (FAT) family
SLC3 Heavy subunits of the heteromeric amino acid transporters	Accessory factors involved in transport	8.A.9 the rBAT transport accessory protein (rBAT) family
SLC52 riboflavin transporter family RFVT/SLC52	Incompletely characterized transport systems	9.A.53 the eukaryotic riboflavin transporter (E-RFT) family
SLC50 sugar efflux transporters	Incompletely characterized transport systems	9.A.58 the sweet; PQ-loop; saliva; MtN3 (sweet) family

The SLC proteins belong to several different Pfam [33] clans, thus sharing specific sequence motifs. The largest ones are the major facilitator superfamily and the amino acid-polyamine-organocation (APC) superfamily, which contain members of 14 and 9 SLC families, respectively [32]. The affiliation of the SLC families to a superfamily is depicted in Figure 1.2.

1.3 Function versus Sequence Similarity

Regarding their function, channels and transporters are clearly distinguishable, but protein classifications in our days classify besides functional criteria often according to phylogenetic relationships. Therefore, ABC transporter proteins that act as channels (CFTR (ABCC7), SUR1 (ABCC8), and SUR2 (ABCC9)) or as translation factors (ABCE1, ABCF1, ABCF2, and ABCF3) are assigned as transporters in TCDB due to their sequence similarity to the functional ABC transporters. Also, the majority of proteins in the solute carrier superfamily act as secondary active transporter and, therefore, the 52 SLC families are included as transporters in IUPHAR/BPS. However, for more than a few solute carrier members, the transport mechanism is currently unrevealed and for some it is known that they function as channels. Table 1.4 shows the contradictory classification of some membrane transport proteins in TCDB, IUPHAR/BPS, and ChEMBL-16.

Figure 1.1 tries to give an overall impression of human proteins involved in transmembrane transport. Furthermore, basic differences between the more phylogenetic-driven classification TCDB and the more pharmacological-driven classification IUPHAR/BPS can be read out.

Figure 1.1 contains some simplification. Not for every IUPHAR/BPS level is an exact counterpart available in TCDB and vice versa. For instance, there is no SLC superfamily in TCDB. Equivalents of SLC families can be found in TCDB class 1, 2, 4, and 8, but they are classified mainly into class 2 (electrochemical potential-driven transporters). Table 1.3 lists the SLC counterparts in TCDB that are not labeled as electrochemical potential-driven transporters. Furthermore, the equivalent of an SLC family in TCDB can contain proteins related to this SLC family but not belonging to this SLC family. For instance, the Rh ammonium transporter family (SLC 42) comprises three Rh glycoproteins, namely, RhAG, RhBG, and RhCG. In red blood cells, the ammonium transport is mediated by a complex of RhAG, RhCE, and RhD [34]. The latter two are not included in SLC 42, but due to their function and sequence similarity all five proteins share the same family (1.A.11 the ammonia channel transporter (Amt) family) and subfamily (1.A.11.4) in TCDB.

1.4 Merged Top-Level Transporter Classification

For a combined overview on human transporters, Figure 1.2 shows a coarse classification into four major groups (solute carriers, ATPases, ABC proteins, and

Table 1.4 Contradictory classification in TCDB, IUPHAR/BPS, and ChEMBL-16.

Protein name	Gene name	TCDB	IUPHAR/BPS	ChEMBL-16
Ammonium transporter Rh type C	RHCG C15orf6 CDRC2 PDRC2 RHGK	Channels/pores	Transporters	—
Ammonium transporter Rh type B	RHBG	Channels/pores	Transporters	—
Ammonium transporter Rh type A	RHAG RH50	Channels/pores	Transporters	—
Solute carrier family 41 member 1	SLC41A1	Channels/pores	Transporters	—
Solute carrier family 41 member 2	SLC41A2	Channels/pores	Transporters	—
Solute carrier family 41 member 3	SLC41A3	Channels/pores	Transporters	—
Urea transporter 1	SLC14A1 HUT11 JK RACH1 UT1 UTE	Channels/pores	Transporters	—
Urea transporter 2	SLC14A2 HUT2 UT2	Channels/pores	Transporters	—
High-affinity copper uptake protein 1	SLC31A1 COPT1 CTR1	Channels/pores	Transporters	—
Probable low-affinity copper uptake protein 2	SLC31A2 COPT2 CTR2	Channels/pores	Transporters	—
ATP-binding cassette subfamily E member 1	ABCE1 RLI RNASEL1 RNASELI RNS4I OK/SW-cl.40	Primary active transporters	—	—
ATP-binding cassette subfamily F member 1	ABCF1 ABC50	Primary active transporters	—	—
ATP-binding cassette subfamily F member 3	ABCF3	Primary active transporters	—	—
ATP-binding cassette subfamily F member 2	ABCF2 HUSSY-18	Primary active transporters	—	—
Cystic fibrosis transmembrane conductance regulator	CFTR ABCC7	Primary active transporters	Ion channels	Transporter
ATP-binding cassette subfamily C member 9	ABCC9 SUR2	Primary active transporters	Transporters	Ion channel
ATP-binding cassette subfamily C member 8	ABCC8 HRINS SUR SUR1	Primary active transporters	Transporters	Ion channel

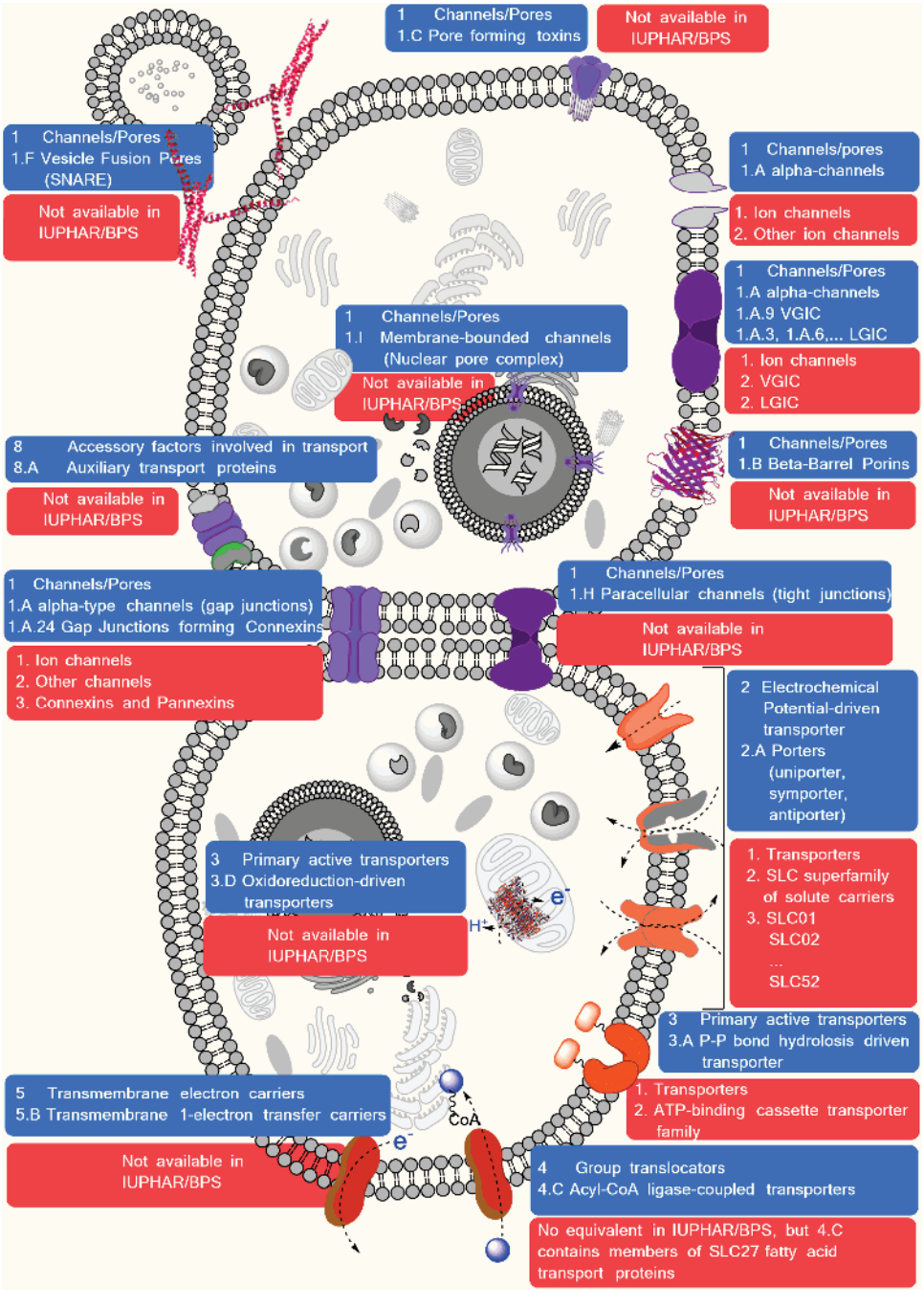

Figure 1.1 Overview of human membrane transport proteins.

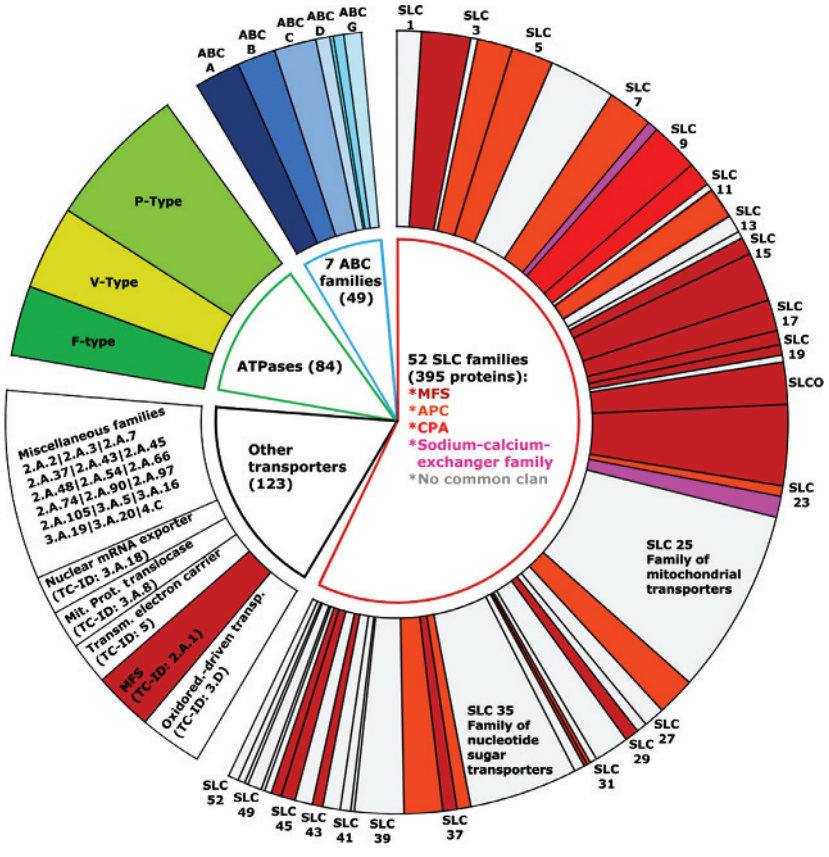

Figure 1.2 Simplified overview of human transport protein families. Abbreviations: MFS, major facilitator superfamily; APC, amino acid/polyamine/organocation superfamily; CPA, cation:proton antiporter superfamily.

other transporters). Figure 1.2 tries to reflect both human transporter classifications in TCDB and IUPHAR/BPS.

The reported names of the protein groups and the number of proteins in Figure 1.2 provide only a rough guide that can vary considerably between the actual classifications. For instance, in a pure functional transporter classification, you may find only 41 human ABC proteins and 381 solute carriers. In a pharmacology-driven classification like IUPHAR/BPS, only the target of levetiracetam (the synaptic vesicle glycoprotein 2A, SV2A) from the group of other transporters is included. Nevertheless, the splitting into four major groups is inspired by the IUPHAR/BPS Guide to PHARMACOLOGY classification of transporters. The assignment to a Pfam clan or family (e.g., the MFS) for the 52 SLC families is color-coded to show their phylogenetic heterogeneity. The MFS group in the group of other transporters is also color-coded, to show the connection to

proteins that are not assigned to an SLC family (e.g., SV2A). The subgroups of the last group in Figure 1.2 (other transporters) are derived from TCDB.

1.5
Choice and Design of the New ChEMBL Classification

Within the framework of the Open PHACTS project [35,36], we were interested to find a classification suitable for channels and transporters. For this, integrating the classification into the existing ChEMBL classification was chosen to facilitate maintainability. By querying different databases (UniProt: reviewed+human+keyword:transport (April 3, 2013); TCDB: all human proteins (May 30, 2013); HGNC: known channel and transporter gene families; GeneOntology: *Homo sapiens*+GO:0022857 transmembrane transport activity (June 15, 2013)), we compiled a list with 1144 human membrane transport proteins and additionally included 300 nonhuman transporters and channels of ChEMBL-16.

The comparison of the classifications for this list of proteins is shown in Figure 1.3. Each of the data sources contains proteins that are unique to this database. Even TCDB, which uses a comprehensive classification of all transport proteins, does not include all identified transporters. An explanation for this is that for

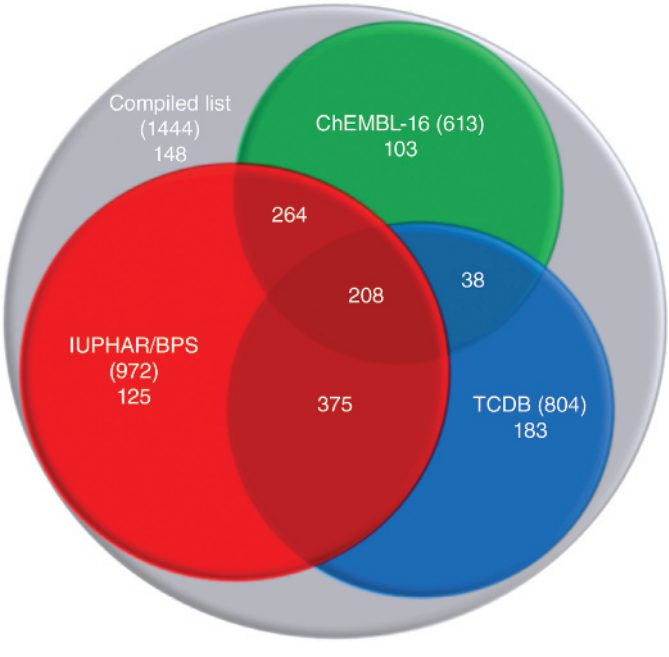

Figure 1.3 Overlap of classified membrane transport proteins in IUPHAR/BPS, TCDB, and ChEMBL-16.

each family, only some examples are provided but not an exhaustive list. On the other hand, TCDB covers even other drug targets such as proteins involved in endocytosis with the term "membrane transport protein." For instance, targets of the medically used botulinum neurotoxin A, for example, SNAP-25 (TC-ID 1.F.1: the synaptosomal vesicle fusion pore (SVF-Pore) family) [37], and a target of the cholesterol-lowering drug ezetimibe, the Niemann-Pick C-1-like protein (TC-ID 2. A.6: the eukaryotic (putative) sterol transporter family) [38] are included in TCDB. SNAP-25 is not contained in IUPHAR/BPS and Niemann-Pick C-1-like protein is classified into other protein targets and patched family. Nevertheless, IUPHAR/BPS includes the highest number of transport proteins from this list, reflecting its focus on pharmacologically relevant proteins.

To generate a classification that can include all proteins from the list, we first predicted, where possible, the classification of the proteins that were unclassified in IUPHAR/BPS (500) or TCDB (604). For proteins where this was not possible in IUPHAR/BPS, new groups were created or added from TCDB.

Finally, IUPHAR/BPS was used as the basis for ion channels, including some subclasses of TCDB. For the transporters, a combination of IUPHAR/BPS and TCDB was used, following TCDB for the first and second level, and afterward using an IUPHAR/BPS-based classification (including the concept of an SLC superfamily). This introduced some contradictions, which were accepted as the SLC classification is well known, thus increasing the usability. In addition, a top-level group of auxiliary transport proteins was introduced according to TCDB class 8.

Table 1.5 shows human transport protein containing classes and subclasses of TCDB and the equivalent groups in IUPHAR/BPS and ChEMBL-19. Text in italics indicates IUPHAR/BPS as source for the ChEMBL-19 classification.

1.6 Transporter as Drug Targets

To get the first insight into the topic of transporter as drug targets, Table 1.6 shows examples of approved drugs and the targeted transport protein group. The drugs in Table 1.6 are all derived from DrugBank. For some, the exact pharmacological mechanism is largely unknown, for example, artemisinin and derivates. Furthermore, the transporter may be one but not the main target for the indication. For instance, the diuretic effect of amiloride is mainly assigned to the inhibition of epithelial Na^+-channels. Table 1.6 also includes some nonclassical transporters (printed in *italics*). These were included because these may be found in phylogenetic transporter classifications like TCDB and we share with Ashcroft *et al.* the view of a blurred boundary between channels and transporters [39]. For instance, the cystic fibrosis transmembrane conductance regulator (CFTR) protein acts as a chloride channel but is often classified as ABC transporter due to phylogenetic reasons. Also, the target of ezetimibe, which is neither a functional channel nor a transporter, is included as it was one of the examples given by Imming *et al.* [1].

Table 1.5 Human membrane transport protein classification of TCDB, IUPHAR/BPS, and ChEMBL-19 in contrast.

TCDB	IUPHAR/BPS	ChEMBL-19
Channels/pores	Ion channels	*Ion channels*
• Alpha-type channels • Beta-barrel porins • Pore-forming toxins • Vesicle fusion pores • Paracellular channels • Membrane-bounded channels	• Ligand-gated ion channels • Voltage-gated ion channels • Other ion channels ◦ Aquaporins ◦ Chloride channels ◦ Connexins and pannexins ◦ Sodium-leak channel, nonselective	• *Ligand-gated ion channels* • *Voltage-gated ion channels* • Other ion channels ◦ *Aquaporins* ◦ *Chloride channels* ◦ *Connexins and pannexins* ◦ *Sodium-leak channel, nonselective* ◦ Vesicle fusion pores ◦ Annexins ◦ . . .
Electrochemical transporter	Transporters	*Transporters*
• Porters (uniporters, symporters, antiporters) Primary active transporters • P-P-bond-hydrolysis-driven transporters • Oxidoreduction-driven transporters Group translocators • Acyl CoA ligase-coupled transporters • Polysaccharide synthase/exporters Transmembrane electron carriers • Transmembrane 1-electron transfer carriers	• SLC-superfamily of solute carriers • Major facilitator superfamily (MFS) of transporters • ATP-binding cassette transporter family • P-Type ATPases • F-type and V-type ATPases	• Electrochemical transporter ◦ *SLC superfamily of solute carriers* ◦ Vesicular neurotransmitter transporter family • Primary active transporter ◦ *ATP-binding cassette* ◦ *P-Type ATPases* ◦ *F-type and V-type ATPases* ◦ Endoplasmic reticular retrotranslocon family ◦ Oxidoreduction-driven transporters • Group translocator • Transmembrane 1-electron transfer carriers
Accessory factors involved in transport • Auxiliary transport proteins	—	Auxiliary transport proteins
Incompletely characterized transport systems	—	—

Table 1.6 Approved drugs and targeted transport proteins.

Major group	Classification in IUPHAR/BPS (TC family and TC-ID of the targeted protein)	Example for approved drug	Drug group (ATC code)
Solute carrier	SLC2 facilitative GLUT transporter family (solute:sodium symporter (SSS) family; 2.A.21.3.16)	Dapaglifocin	Antidiabetic drug (A10BX09)
	SLC6 sodium- and chloride-dependent neurotransmitter transporter family (neurotransmitter:sodium symporter family; 2.A.22.1.1)	Fluoxetin	Antidepressant (N06AB03)
	SLC6 sodium- and chloride-dependent neurotransmitter transporter family (neurotransmitter:sodium symporter family; 2.A.22.3.2)	Tiagabine	Antiepileptics (N03AG06)
	SLC9 Na+/H+ exchanger family (the monovalent cation:proton antiporter-1 (CPA1) family; 2.A.36.1.13)	Amiloride	Diuretics, potassium-sparing (C03DB01)
	SLC12 electroneutral cation-coupled Cl cotransporter family (cation-chloride cotransporter family; 2.A.30.1.2)	Furosemide	Diuretics, high-ceiling (C03CA01)
	SLC18 vesicular amine transporter family (the drug:H+ antiporter-1 (12 spanner) (DHA1) family; 2.A.1.2.29)	Reserpine	Antihypertensives (C02AA02)
		Tetrabenazine	Hyperkinetic movement disorder (N07XX06)
	SLC22 organic cation/anion/zwitterion transporter family (organic cation transporter (OCT) family; 2.A.1.19.10|2.A.1.19.31|2.A.1.19.34)	Probenecid	Uricosuric drug (M04AB01)
	SLC25 mitochondrial carrier (MC) family (MC family; 2.A.29.1.2|2.A.29.1.1|2.A.29.1.10)	Clodronate	Osteoporosis, bone metastases (M05BA02)
	SLC52 riboflavin transporter family RFVT/SLC52 (E-RFT family; 9.A.53.1.3)	Gamma hydroxybutyric acid	Anesthetics (N01AX11)
ATPases	P-type ATPase (P-type ATPase superfamily; 3.A.3.1.1)	Digitoxin	Cardiac glycosides (c1AA04)
	P-type ATPase (P-type ATPase superfamily; 3.A.3.1.2)	Omeprazol	Proton pump inhibitors (A02Bc1)
	P-type ATPase (P-type ATPase superfamily; 3.A.3.2.?)	Lumefantrine, artemether (artemisinin derivates)	*Antimalarials (P01BF01)*

Table 1.6 (*continued*)

Major group	Classification in IUPHAR/BPS (TC family and TC-ID of the targeted protein)	Example for approved drug	Drug group (ATC code)
ABC protein	*ATP-binding cassette subfamily C member 7 (cystic fibrosis transmembrane conductance exporter (CFTR) family; 3.A.1.202.1)*	*Ivacaftor*	*Cystic fibrosis (R07AX02)*
	ATP-binding cassette subfamily C member 8 (the drug conjugate transporter (DCT) family; 3.A.1.208.4)	*Repaglinide*	*Antidiabetic drug (A10BX02)*
Other transport proteins	Major facilitator superfamily of transporters, non-SLC (vesicular neurotransmitter transporter (VNT) family; 2.A.1.22.?)	Levetiracetam	Antiepileptics (N03AX14)
	Other protein targets (*eukaryotic (putative) sterol transporter (EST) family; 2.A.6.6.6)*	*Ezetimibe*	*Lipid-modifying agent (C10AX09)*

Note: The transport proteins are not necessarily the target responsible for the reported drug indication.

Regarding prospective targets, various transporters are and were considered promising drug targets in chemotherapy, but so far the candidates have failed in clinical trials. At present, Winter *et al.* describe SLC35F2 as a prospective target in cancer therapy and Rask-Andersen *et al.* mention members of SLC2, SLC5, SLC7, and SLC9 as promising targets in cancer therapy and members of SLC10 as potential targets against constipation and hypercholesterolemia [40,41].

1.7 Drug Targets in the SLC Classification

Several solute carriers are reported targets of approved drugs [41]. Here, we use the SLC classification as a framework to give an overview on the diseases a transporter is connected to, known drug molecules that directly target the transporter, and a count of bioactivity data available in ChEMBL to give an estimate on the degree of interest in the target.

The data shown in the figures was collected from different sources. Molecules targeting transporters were retrieved from the DrugBank xml [42], using a modified KNIME [43] workflow from the available example workflows [44]. Disease information was downloaded from DisGeNET [45] using the curated gene–disease associations [46]. Protein/gene mappings were generated from UniProt [23]. Bioactivity data counts were retrieved from ChEMBL-19, using single proteins only. Cladograms were generated with FigTree [47] using SLC sequences retrieved from Uniprot. Multiple sequence alignments for the

members of one Pfam clan (e.g., the major facilitator superfamily) were generated with Clustal Omega [48] using the default parameters on the EBI Web server [49]. Counts for each target were added manually.

Figure 1.4 shows the counts for SLC members belonging to the amino acid-polyamine-organocation superfamily. Seven of the families show members with reported drugs, with three of them being previously reported by Rask-Andersen

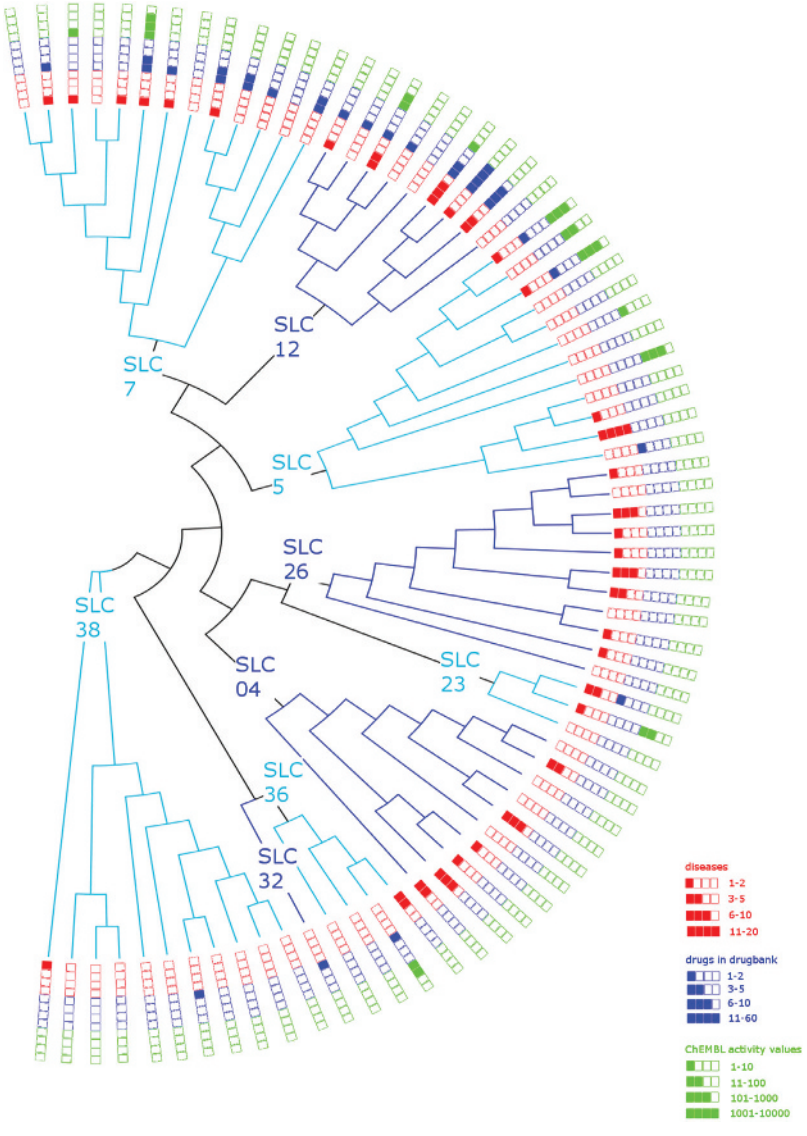

Figure 1.4 Cladogram of the 9 SLC families belonging to the APC superfamily and the number of connected diseases, drugs, and bioactivity values.

et al. [41] to be targets of approved drugs or under investigation (SLC5, SLC7, and SLC12). Closer investigation of the drugs for the remaining families shows that these are mostly vitamins or amino acids. Investigating the number of associated diseases for families without known drugs finds SLC4 and SLC26 as interesting families. Indeed, these are mentioned as potential new targets by Rask-Andersen *et al.*, however, as target for antineoplastic agents, which is not one of the associated diseases. Figure 1.4 thus shows their association with several other diseases as well.

1.8 Conclusions

In this chapter, we gave an overview of available transporter classification schemes. A new version of the ChEMBL classification was introduced. For this, we wanted to have a less complex, browsable classification and, therefore, merged TCDB with the IUPHAR/BPS classification. The advantage compared to the pure IUPHAR/BPS transporter classification is that you still easily find the main transporter groups (ABC transporter, SLC members, and ATPases) and, if new bioactivity data for less common human or nonhuman transporters/transporter families are reported, these transporters can be easily integrated in conformity with TCDB, which is more complicated with a classification following IUPHAR/BPS only. For ChEMBL, we wanted to use the well-known SLC families to have a less complex transport protein classification than TCDB but keep the possibility to extend the scheme with the corresponding TCDB classes if it becomes significant.

Classifications allow (semi)automatic clustering of information. We used the SLC families to give an overview of interacting drugs and associated diseases of members of the APC clan. One disadvantage of an automated approach, however, is that false positive connections can be drawn. For example, the only human member of SLC32, the vesicular inhibitory amino acid transporter (VIAAT), seems to have a target drug according to Figure 1.3. On closer inspection, this is glycine, which is one of the natural substrates of this transporter. A more detailed investigation will, therefore, be necessary to draw valid conclusions from these investigations.

Acknowledgment

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no. 115191, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in-kind contribution. We also acknowledge financial support provided by the Austrian Science Fund, grant F3502.

References

1 Imming, P., Sinning, C., and Meyer, A. (2006) Drugs, their targets and the nature and number of drug targets. *Nat. Rev. Drug Discov.*, **5** (10), 821–834.

2 Rask-Andersen, M., Almén, M.S., and Schiöth, H.B. (2011) Trends in the exploitation of novel drug targets. *Nat. Rev. Drug Discov.*, **10** (8), 579–590.

3 Wishart, D.S., Knox, C., Guo, A.C., Cheng, D., Shrivastava, S., Tzur, D., Gautam, B., and Hassanali, M. (2008) DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.*, **36** (Database issue), D901–D906.

4 Rask-Andersen, M., Masuram, S., and Schiöth, H.B. (2014) The druggable genome: evaluation of drug targets in clinical trials suggests major shifts in molecular class and indication. *Annu. Rev. Pharmacol. Toxicol.*, **54** (1), 9–26.

5 Saier, M.H., Reddy, V.S., Tamang, D.G., and Vastermark, A. (2013) The transporter classification database. *Nucleic Acids Res.*, **42** (D1), D251–D258.

6 Gaulton, A., Bellis, L.J., Bento, A.P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B., and Overington, J.P. (2011) ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.*, **40** (D1), D1100–D1107.

7 Schwacke, R., Schneider, A., van der Graaff, E., Fischer, K., Catoni, E., Desimone, M., Frommer, W.B., Flügge, U.-I., and Kunze, R. (2003) ARAMEMNON, a novel database for arabidopsis integral membrane proteins. *Plant Physiol.*, **131** (1), 16–26.

8 Brohée, S., Barriot, R., Moreau, Y., and André, B. (2010) YTPdb: a wiki database of yeast membrane transporters. *Biochim. Biophys. Acta*, **1798** (10), 1908–1912.

9 Fichant, G., Basse, M.-J., and Quentin, Y. (2006) ABCdb: an online resource for ABC transporter repertories from sequenced archaeal and bacterial genomes. *FEMS Microbiol. Lett.*, **256** (2), 333–339.

10 Kumar, S., Mukherjee, M.M., and Varela, M.F. (2013) Modulation of bacterial multidrug resistance efflux pumps of the major facilitator superfamily. *Int. J. Bacteriol.*, **2013**, e204141.

11 Slavic, K., Krishna, S., Derbyshire, E.T., and Staines, H.M. (2011) Plasmodial sugar transporters as anti-malarial drug targets and comparisons with other protozoa. *Malar. J.*, **10** (1), 165.

12 Hediger, M.A., Clémençon, B., Burrier, R.E., and Bruford, E.A. (2013) The ABCs of membrane transporters in health and disease (SLC series): introduction. *Mol. Aspects Med.*, **34** (2–3), 95–107.

13 M. Müller. ABC-Transporter Proteins and other Transporters. Available at http://nutrigene.4t.com/translink.htm (accessed 22 Sep 2016).

14 Pawson, A.J., Sharman, J.L., Benson, H.E., Faccenda, E., Alexander, S.P.H., Buneman, O.P., Davenport, A.P., McGrath, J.C., Peters, J.A., Southan, C., Spedding, M., Yu, W., and Harmar, A.J. (2014) The IUPHAR/BPS Guide to PHARMACOLOGY: an expert-driven knowledgebase of drug targets and their ligands. *Nucleic Acids Res.*, **42** (Database issue), D1098–D1106.

15 Ren, Q., Chen, K., and Paulsen, I.T. (2007) TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. *Nucleic Acids Res.*, **35** (Database), D274–D279.

16 Zhao, M., Chen, Y., Qu, D., and Qu, H. (2011) TSdb: a database of transporter substrates linking metabolic pathways and transporter systems on a genome scale via their shared substrates. *Sci. China Life Sci.*, **54** (1), 60–64.

17 Morrissey, K.M., Wen, C.C., Johns, S.J., Zhang, L., Huang, S.-M., and Giacomini, K.M. (2012) The UCSF-FDA TransPortal: a public drug transporter database. *Clin. Pharmacol. Ther.*, **92** (5), 545–546.

18 Hersey, A. (Sep. 2013) ChEMBL database release 16, EMBL-EBI.

19 Viereck, M., Gaulton, A., Digles, D., and Ecker, G.F. (2014) Transporter taxonomy: a comparison of different transport

protein classification schemes. *Drug Discov. Today Technol.*, **12**, e37–e46.
20 Saier, M.H. (2000) A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol. Mol. Biol. Rev.*, **64** (2), 354–411.
21 Tipton, K. and Boyce, S. (2000) History of the enzyme nomenclature system. *Bioinforma. Oxf. Engl.*, **16** (1), 34–40.
22 Busch, W. and Saier, M.H., Jr. (2003) The IUBMB-endorsed transporter classification system. *Methods Mol. Biol.*, **227**, 21–36.
23 The UniProt Consortium (2013) Activities at the universal protein resource (UniProt). *Nucleic Acids Res.*, **42** (D1), D191–D198.
24 Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000) The protein data bank. *Nucleic Acids Res.*, **28** (1), 235–242.
25 Saier, M.H., Yen, M.R., Noto, K., Tamang, D.G., and Elkan, C. (2009) The transporter classification database: recent advances. *Nucleic Acids Res.*, **37** (Database), D274–D278.
26 Alexander, S., Mathie, A., and Peters, J. (2011) Guide to receptors and channels (GRAC). 5th edition, *Br. J. Pharmacol.*, **164**, S1–S2.
27 Sharman, J.L., Benson, H.E., Pawson, A.J., Lukito, V., Mpamhanga, C.P., Bombail, V., Davenport, A.P., Peters, J.A., Spedding, M., Harmar, A.J., and NC-IUPHAR (2013) IUPHAR-DB: updated database content and new features. *Nucleic Acids Res.*, **41** (D1), D1083–D1088.
28 Alexander, S.P.H., Benson, H.E., Faccenda, E., Pawson, A.J., Sharman, J.L., McGrath, J.C., Catterall, W.A., Spedding, M., Peters, J.A., Harmar, A.J., and CGTP Collaborators (2013) The Concise Guide to PHARMACOLOGY 2013/14: overview. *Br. J. Pharmacol.*, **170** (8), 1449–1458.
29 Hagenbuch, B. and Stieger, B. (2013) The SLCO (former SLC21) superfamily of transporters. *Mol. Aspects Med.*, **34** (2–3), 396–412.
30 Hersey, A. (Aug. (2014)) CHEMBL database release 18, EMBL-EBI.
31 Hediger, M.A., Romero, M.F., Peng, J.-B., Rolfs, A., Takanaga, H., and Bruford, E.A. (2004) The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteins. *Pflügers Arch.*, **447** (5), 465–468.
32 Höglund, P.J., Nordström, K.J.V., Schiöth, H.B., and Fredriksson, R. (2011) The solute carrier families have a remarkably long evolutionary history with the majority of the human families present before divergence of bilaterian species. *Mol. Biol. Evol.*, **28** (4), 1531–1541.
33 Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E.L.L., Eddy, S.R., Bateman, A., and Finn, R.D. (2011) The Pfam protein families database. *Nucleic Acids Res.*, **40** (D1), D290–D301.
34 Nakhoul, N.L. and Lee Hamm, L. (2013) Characteristics of mammalian Rh glycoproteins (SLC42 transporters) and their role in acid–base transport. *Mol. Aspects Med.*, **34** (2–3), 629–637.
35 Gray, A.J.G., Groth, P., Loizou, A., Askjaer, S., Brenninkmeijer, C., Burger, K., Chichester, C., Evelo, C.T., Goble, C., Harland, L., Pettifer, S., Thompson, M., Waagmeester, A., and Williams, A.J. (2014) Applying linked data approaches to pharmacology: architectural decisions and implementation. *Semantic Web*, **5**, 101–113.
36 Williams, A.J., Harland, L., Groth, P., Pettifer, S., Chichester, C., Willighagen, E.L., Evelo, C.T., Blomberg, N., Ecker, G., Goble, C., and Mons, B. (2012) Open PHACTS: semantic interoperability for drug discovery. *Drug Discov. Today*, **17** (21–22), 1188–1198.
37 Frevert, J. (2015) Pharmaceutical, biological, and clinical properties of botulinum neurotoxin type a products. *Drugs RD*, **15**, 1–9.
38 Altmann, S.W., Davis, H.R., Zhu, L., Yao, X., Hoos, L.M., Tetzloff, G., Iyer, S.P.N., Maguire, M., Golovko, A., Zeng, M., Wang, L., Murgolo, N., and Graziano, M.P. (2004) Niemann-Pick C1 like 1 protein is critical for intestinal cholesterol absorption. *Science*, **303** (5661), 1201–1204.
39 Ashcroft, F., Gadsby, D., and Miller, C. (2009) Introduction. The blurred boundary between channels and

transporters. *Philos. Trans. R. Soc. B*, **364** (1514), 145–147.

40 Winter, G.E., Radic, B., Mayor-Ruiz, C., Blomen, V.A., Trefzer, C., Kandasamy, R.K., Huber, K.V.M., Gridling, M., Chen, D., Klampfl, T., Kralovics, R., Kubicek, S., Fernandez-Capetillo, O., Brummelkamp, T.R., and Superti-Furga, G. (2014) The solute carrier SLC35F2 enables YM155-mediated DNA damage toxicity. *Nat. Chem. Biol.*, **10** (9), 768–773.

41 Rask-Andersen, M., Masuram, S., Fredriksson, R., and Schiöth, H.B. (2013) Solute carriers as drug targets: current use, clinical trials and prospective. *Mol. Aspects Med.*, **34** (2–3), 702–710.

42 DrugBank: Downloads," DrugBank. Available at http://www.drugbank.ca/system/downloads/current/drugbank.xml.zip (accessed 10 October 2014).

43 Berthold, M.R., Cebron, N., Dill, F., Gabriel, T.R., Kötter, T., Meinl, T., Ohl, P., Sieb, C., Thiel, K., and Wiswedel, B. (2007) KNIME: the Konstanz information miner, in *Studies in Classification, Data Analysis, and Knowledge Organization Springer*.

44 KNIME | DrugBank," KNIME. Available at http://tech.knime.org/book/network-examples/drugBank (accessed 1 July 2014).

45 Bauer-Mehren, A., Bundschus, M., Rautschka, M., Mayer, M.A., Sanz, F., and Furlong, L.I. (2011) Gene-disease network analysis reveals functional modules in mendelian, complex and environmental diseases. *PLoS ONE*, **6** (6), e20284.

46 DisGeNET: a database of gene-disease associations," *DisGeNET*. Available at http://www.disgenet.org/web/DisGeNET/v2.1/downloads#curated (accessed 2 October 2014).

47 FigTree v1.4.2," *Molecular Evolution, Phylogenetics and Epidemiology*. Available at http://tree.bio.ed.ac.uk/software/figtree/ (accessed 11 November 2014).

48 Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., and Higgins, D.G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.*, **7** (1), 539.

49 Clustal Omega < Multiple Sequence Alignment < EMBL-EBI." Available at http://www.ebi.ac.uk/Tools/msa/clustalo/ (accessed: 12 October 2014).