

COMMENTARIES

SLC Classification: An Update

A Schlessinger^{1,2}, SW Yee^{3,4}, A Sali³⁻⁵ and KM Giacomini^{3,4,6}

The 386 human SLC superfamily members are diverse in sequence, structure, and function. Using sequence similarity, we previously classified the SLC superfamily members and identified relationships among families. With the recent determination of new SLC structures and identification of previously unknown human SLC families, an update of our previous classification is timely. Here, we comprehensively compare the SLC sequences and structures and discuss the applicability of structure-based ligand discovery to key SLC members.

The human SLC superfamily comprises 386 member proteins that transport a broad spectrum of substrates, such as nutrients, toxins, and prescription drugs.^{1,2} The superfamily members are classified into 52 families, based on the number of predicted or observed transmembrane α -helices (usually 10–14) and sequence similarity, in which members of each family share sequence identity of at least 20% with at least one other family member.² Members within an SLC family can have substrates with different physicochemical properties despite their common evolutionary origin. For example, the SLC22 family includes transporters of organic anions, cations, or zwitterions. Conversely, SLC families such as the amino acid transporter families SLC1 and SLC7 can be unrelated evolutionarily but still have chemically similar substrates.

Atomic structures of SLC transporters can guide computations and experiments to describe the mechanism of

transport and its clinical implications. In particular, structures can be used to rationalize the effect of nonsynonymous variants on drug transport, pharmacokinetics, or pharmacodynamics, as well as to predict these functional consequences of uncharacterized variants. Moreover, the transporter structure can be used to predict small-molecule ligands such as endogenous metabolites, fragment molecules, and prescription drugs via virtual screening of compound libraries, which can contain millions of purchasable molecules, against the transporter binding site.^{3,4} The ligand prediction is usually followed by experimental testing, and the confirmed hits can then (i) guide drug–drug interaction studies, (ii) provide novel chemical tools to further characterize the transporters' functions, and even (iii) serve as leads for designing drugs targeting selected SLC transporters.

Because structures of membrane proteins, particularly those in humans, are

difficult and expensive to determine, the only experimentally determined human SLC structure is the X-ray structure of the human ammonium transporter Rh type C (RHCG, SLC43A3), a glycoprotein that plays a role in blood group incompatibility.⁵ When an experimentally determined structure is not available, a model for a target protein can be constructed computationally by homology or comparative modeling, which relies on structures of related proteins that serve as templates for the target sequence.⁶

Several structures of SLC transporters from various prokaryotic and eukaryotic organisms can serve as templates for modeling the human members. These structures confirm that some human SLC families are unrelated to each other, as predicted from sequence, and unlike other superfamilies such as the ABC transporters.³ For example, RHCG has channel-like properties whereas the structurally different lactose permease (homologue of the SLC37 family) undergoes substantial conformational changes during transport.⁷ In addition, these structures also indicate that despite their distant relationships (sequence identity of ~10%), some human SLC families have surprisingly similar structures. In particular, the largest two structural classes or folds of the human SLC families are the major facilitator superfamily (MFS) (e.g., SLC22) and the neurotransmitter:sodium symporter (NSS) or LeuT-like fold (**Table 1** and **Supplementary Table 1** in the **Supplementary Material** online).³ Identifying distant relationships between human sequences and other SLC members of known structure is therefore beneficial

¹Department of Pharmacology and Systems Therapeutics, Mount Sinai School of Medicine, New York, New York, USA; ²Tisch Cancer Institute, Mount Sinai School of Medicine, New York, New York, USA; ³Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, California, USA; ⁴California Institute for Quantitative Biosciences, University of California, San Francisco, San Francisco, California, USA; ⁵Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California, USA; ⁶Institute for Human Genetics, University of California, San Francisco, San Francisco, California, USA. Correspondence: A Schlessinger (avner.schlessinger@mssm.edu) or KM Giacomini (kathy.giacomini@ucsf.edu)

Table 1 Drug ADME SLC families that can be modeled based on atomic resolution structures from other organisms

Family ^a	Function ^b	Template structure ^c	Percent sequence identity ^d	Representative drug substrates ^e
SLC7 (14)	Cationic amino acid transporter/ glycoprotein-associated family	 AdiC [#]	21 (1.4×10^{-47})	Melphalan, gabapentin, levodopa, baclofen
SLC10 (7)	Na ⁺ bile salt cotransporters	 ASBT _{NM}	26 (1.8×10^{-42})	Rosuvastatin, atorvastatin, fluvastatin
SLC15 (4)	Proton oligopeptide cotransporters	 PepT ₅₀ [*]	34 (2.2×10^{-28})	Valacyclovir, cephalexin, cefadroxil, enalapril, captopril
SLC22 (26)	Organic cation/anion/zwitterion transporters	 PiPT [*]	20 (4.9×10^{-34})	Metformin, acyclovir, methotrexate, olmesartan, ipratropium, oxaliplatin, cimetidine
SLC28 (3)	Na ⁺ -coupled nucleoside transporters	 vcCNT	40 (6.4×10^{-130})	Fludarabine, gemcitabine, cytarabine
SLC47 (2)	Multidrug and toxin extrusion (MATE) transporters	 NorM	23 (4.8×10^{-51})	Metformin, trospium, fexofenadine

ADME, absorption, distribution, metabolism, excretion.

^aThe human SLC family, as annotated by the Bioparadigms database.² The number of human protein sequences in the family is provided in parentheses.^bThe function of the human family, as described in the Bioparadigms database.^cThe atomic structure most closely related to the family. Structures with the MFS and NSS folds are marked with * and #, respectively. Detailed descriptions of the structures, including the full names of the proteins and the corresponding references, are given in the **Supplementary Material**.^dThe percentage sequence identity of the best-scoring hit from each family; E-value is given in parentheses (**Supplementary Material**).^eExamples of key prescription drugs that are substrates of the transporter.

for characterizing the structures and functions of the human members.

Previously, using sensitive sequence alignment and sequence-based clustering programs, we created a “map” of the human SLC superfamily.¹ We then annotated different functional characteristics of the transporters onto the

sequence-based network and found new relationships among seemingly unrelated families. Since our original analysis of the SLC superfamily in 2010, four previously unknown human SLC families (SLC49 to SLC52), which include 10 transporters, have been discovered.² Furthermore, technological

advances have enabled the experimental determination of a number of SLC structures from diverse organisms. This increase in the structural coverage of the SLC superfamily can be largely attributed to structural genomics consortia,⁸ such as the National Institutes of Health (NIH)-supported Center of Structural

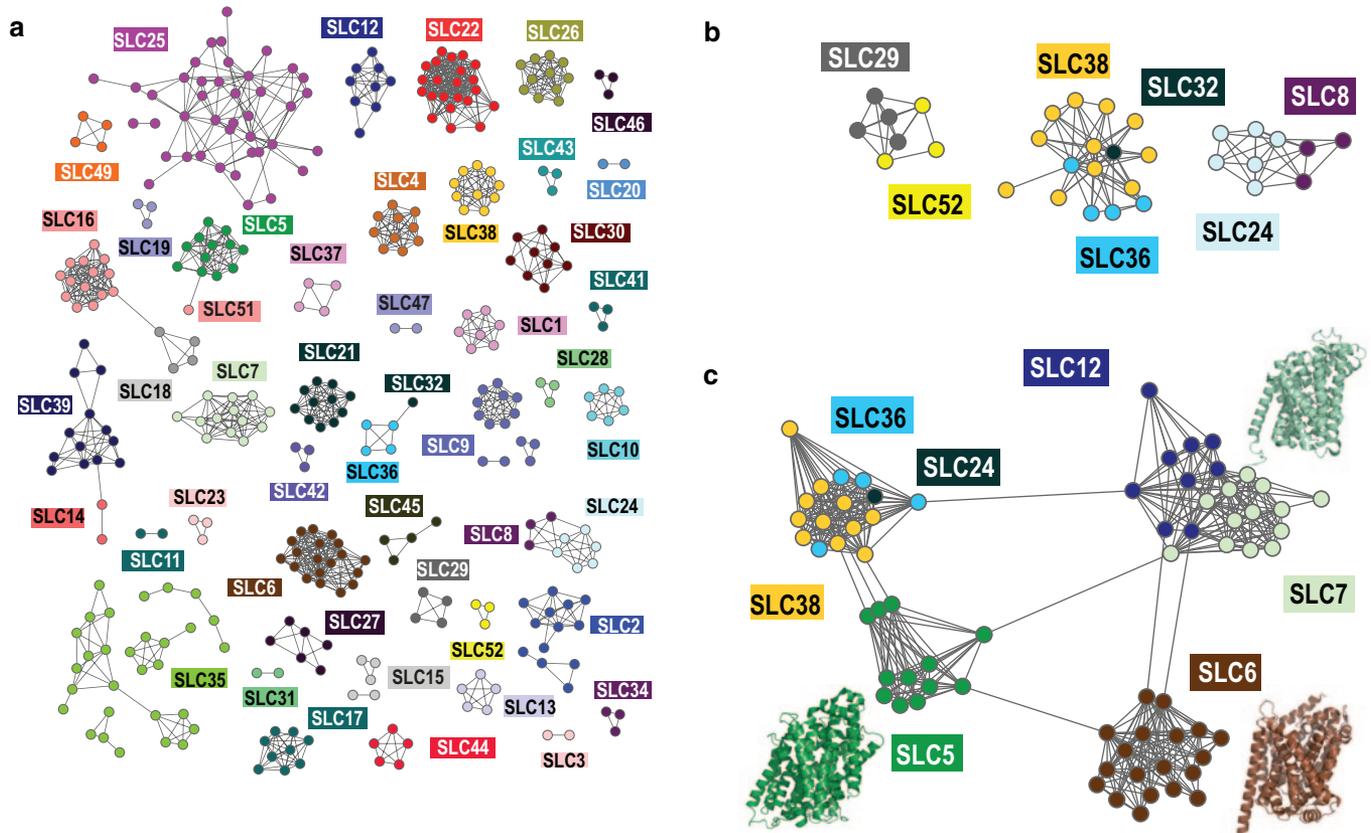


Figure 1 Network view of the human SLC superfamily. The relationships between SLC sequences are visualized in a sequence similarity network.¹ The nodes of the network correspond to SLC sequences, color-coded by their family.² The edges between the nodes correspond to pairwise sequence relationships better than a cutoff. We use sequence identity and E-value. The E-value of a match is the number of sequences in the queried database that are expected to match by chance the query sequence at least as well as the assessed match. Cutoffs are defined to be permissive enough to capture the majority of the known distant relationships among SLC members with minimal false positives. The edges correspond to alignments with sequence identity of at least 20% and (a) an E-value lower than 1; (b) an E-value lower than 0.01; and (c) sequence identity of at least 10% and an E-value lower than 5. In a, SLC sequences that are connected to at least one other sequence are shown. For example, SLC50A1, which is the only SLC50 member and is not related to any other SLC member, is not shown. In b, only highly interconnected families are shown for a more stringent E-value cutoff. In c, a cluster of distantly related families predicted to have the NSS fold is shown. Atomic structures of homologues of the human SLC families are visualized, and their colors correspond to the color of the most similar human family. The structures include the Na⁺/galactose symporter (vSGLT) (SLC5 homologue), Na⁺/amino acid transporter LeuT (SLC6 homologue), and amino acid transporter AdiC (SLC7 homologue). NSS, neurotransmitter:sodium symporter.

Genomics of Membrane Proteins, that aim at high-throughput experimental determination of membrane protein structures with biomedical relevance. In this Commentary, we update the SLC classification to create a new SLC map in view of the newly discovered human SLC families and the atomic structures of their homologues.

We visualize relationships among SLC sequences using similarity networks (Figure 1 and the **Supplementary Material**). In particular, we first gathered the sequences of 386 human SLC sequences, including the 10 recently discovered transporters. Briefly, these 10 transporters are grouped into the following fami-

lies: the SLC49 family of FLVCR-related transporters that transport heme (four members), the SLC50 family of sugar efflux transporters (one member), the SLC51 family of steroid-derivative transporters (two members that form a functional heterodimer), and the SLC52 family of riboflavin transporters (RFVT) (three members). We performed an “all-against-all” sequence comparison among the 386 sequences and employed sequence-based clustering to construct and visualize sequence-based similarity networks using different cutoffs (Figure 1 and **Supplementary Material**). When we use a cutoff similar to that for previously defined SLC families, the SLC49

and SLC50 families are not connected to any other family, confirming that they constitute distinct families (Figure 1a).

The SLC51 members SLC51A and SLC51B, which function as a heterodimer, are not related to each other in sequence. Previously, SLC members were grouped into families based on sequence relationships, and SLC members that function as heterodimers were grouped as distinct families (e.g., the SLC3 and SLC7 families). Therefore, the classification of the SLC50 members into one family can be confusing. Furthermore, SLC51B, a single membrane-spanning helix protein, is connected to SLC5A4, which is predicted to have 14

transmembrane helices (**Supplementary Figure 1**). Although their different topologies indicate they have different structures, their distant sequence similarity (sequence identity of 22%) suggests they might share functional features (e.g., a posttranslational modifications site). Importantly, the SLC52 family of riboflavin transporters is highly connected to the SLC29 family, even when using a more stringent expectation value (E-value) cutoff (**Figure 1b** and **Supplementary Material**) and also has the same number of predicted transmembrane helices (11) (**Supplementary Figure 1c**), suggesting that they are evolutionarily linked and might have similar structures, functions, or both.

To identify protein structures that are sufficiently similar to the human members for constructing useful comparative models, we compared the human sequences to all known experimentally determined structures in the Protein Data Bank (PDB), which contains human and nonhuman structures. Human sequences that share sequence identities of at least 20% (typically 20–40%) with the known structures, including most of the transmembrane region, can be suitable for comparative modeling. Because of the distant relationships between the SLC targets and their template structures, and because most modeling programs were not optimized for membrane proteins, models for these challenging SLC targets should be used with extra caution.⁹

Our analysis indicates that several transporter families with clinical importance can be modeled with the new structures as templates (**Figure 1**, **Table 1**, and **Supplementary Table 1**). They include transporter families important for drug absorption, distribution, metabolism, and excretion (ADME) (**Table 1**). For example, the organic ion transporters (SLC22) and the peptide transporters (PepT, SLC15) can be modeled based on the structures of the high-affinity phosphate importer PiPT and the peptide transporter PepT_{SO}, respectively (**Table 1**). Key drug targets (**Supplementary Table 1**), such as the glucose transporters (GLUT, SLC2), can be modeled based on the structure of the D-xylose–proton symporter (XyIE).

Furthermore, the PiPT, PepT_{SO}, and the XyIE structures indicate that human members of the SLC22, SLC15, and SLC2 families adopt the MFS fold, despite not sharing significant sequence similarity among them (**Table 1**, **Figure 1**).¹ The SLC22 and SLC2 families were previously predicted to belong to the MFS, thereby increasing our confidence in the sequence “map” as a guide for structural modeling. However, PepT_{SO} is not connected to any other family in our current and previous similarity network, indicating that some MFS families are more divergent than others and that the MFS fold might be more common among the human SLC members than expected.

In addition, several newly determined SLC structures are “novel folds” that are not detectably similar to any other structure in the PDB, covering previously unknown areas in the sequence space. They include the structures of the prokaryotic homologues of the Na⁺/Ca²⁺-K⁺ exchanger (NCKX, SLC24), the concentrative nucleoside transporter (CNT, SLC28), and the Na⁺-bile salt co-transporter (SLC10) families that share sequence identities of 28%, 40%, and 26% with the human members, respectively (**Table 1**). Our ability to model members of the SLC28 and SLC10 families, and possibly identify new ligands for these transporters with virtual screening, is of particular clinical importance, as these families include several pharmacologically important transporters that determine the distribution and elimination of many prescription drugs (**Table 1**). For example, CNT3 (SLC28A3) is responsible for transport of anticancer and antiviral nucleoside analog drugs to target tissues.

Notably, although speculative, the substrate spectrum of the transporter family can be correlated with its predicted fold. Several drug transporter families in the kidney and liver, such as the SLC21 (SLCO), SLC22, and SLC15, have broad substrate spectra and are also predicted to adopt the MFS fold (**Table 1**).¹ This fold has been associated with a wide array of substrates, which can partially be rationalized by its large pore and substantial conformational move-

ments during transport. Conversely, transporter families that are involved in tight regulation of signaling molecules, such as amino acids and neurotransmitters in the nervous system, usually have a narrow substrate spectrum. These transporters, including members of the SLC6, SLC7, SLC32, SLC36, and SLC38 families, are predicted to have the NSS fold (**Table 1** and **Figure 1c**).¹ Several proteins with the NSS fold have a narrower range of substrates than that of MFS, which can be rationalized by a smaller pore and smaller conformational changes during transport. Future studies will confirm or dispute this hypothesis, which is based on a small number of families.

Finally, prediction of SLC ligands with comparative modeling and virtual screening can be more accurate if the following three criteria are fulfilled: (i) high-quality template structure (e.g., resolution of 3 Å or higher); (ii) high sequence similarity between the target and template (e.g., sequence identity of 25% or higher), particularly in the binding site region; and (iii) the conformation of the template structure (ligand-bound conformations are likely to yield more accurate predictions because they typically represent high-affinity states and because the binding-site location can be transferred from the template structure to the model). Based on these criteria, several important human families, including the SLC2, SLC10, SLC22, and SLC28, can be targeted using structure-based ligand discovery. For example, the vcCNT¹⁰ structure is a relatively high-quality structure (e.g., resolution of 2.4 Å) that was determined in a ligand-bound conformation. vcCNT also shares sequence identity of 40% with the human CNT3 (SLC28A3), including almost identical binding-site residues (**Table 1**).

In conclusion, in this Commentary we have updated and extended the SLC classification. In particular, we analyzed an updated set of 386 SLC sequences, including 10 new members, to identify previously unknown relationships among human SLC members and to guide future attempts directed at characterizing their functions. We also com-

pared the human SLC sequences to the sequences of atomic structures in the PDB, to evaluate the current structural coverage of the human SLC superfamily, which includes a number of recently determined SLC structures. Finally, we discuss the applicability of structure-based discovery to select human SLC members and its clinical significance.

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at <http://www.nature.com/cpt>

ACKNOWLEDGMENTS

We are grateful to David Smith (University of Michigan), Robert Stroud (UCSF), Christof Grever (Binghamton University), and Seok-Yong Lee (Duke University) for discussions about SLC transporters. We also thank Ursula Pieper for technical assistance. The project was supported by grants from National Institutes of Health (R01 GM54762, U54 GM074929, P01 GM71790, and GM61390 to A. Sali.). K.M.G. and S.W.Y. acknowledge the funding from an NIH-NIGMS grant (GM61390).

CONFLICT OF INTEREST

The authors declared no conflict of interest.

© 2013 ASCPT

- Schlessinger, A. *et al.* Comparison of human solute carriers. *Protein Sci.* **19**, 412–428 (2010).
- Hediger, M.A., Romero, M.F., Peng, J.-B., Rolfs, A., Takanaga, H. & Bruford, E.A. The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteins. *Pflugers Arch.* **447**, 465–468 (2004).
- Schlessinger, A. *et al.* Structure-based discovery of prescription drugs that interact with the norepinephrine transporter, NET. *Proc. Natl. Acad. Sci. USA* **108**, 15810–15815 (2011).
- Geier, E.G. *et al.* Structure-based ligand discovery for the Large-neutral Amino Acid Transporter 1, LAT-1. *Proc. Natl. Acad. Sci. USA* **110**, 5480–5485 (2013).
- Gruswitz, F. *et al.* Function of human Rh based on structure of RhCG at 2.1 Å. *Proc. Natl. Acad. Sci. USA* **107**, 9638–9643 (2010).
- Sali, A. & Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
- Guan, L. & Kaback, H.R. Lessons from lactose permease. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 67–91 (2006).
- Pieper, U. *et al.* Coordinating the impact of structural genomics on the human alpha-helical transmembrane proteome. *Nat. Struct. Mol. Biol.* **20**, 135–138 (2013).
- Schlessinger, A., Khuri, N., Giacomini, K.M. & Sali, A. Molecular modeling and ligand docking for solute carrier (SLC) transporters. *Curr. Top. Med. Chem.* **13**, 843–856 (2013).
- Johnson, Z.L., Cheong, C.G. & Lee, S.Y. Crystal structure of a concentrative nucleoside transporter from *Vibrio cholerae* at 2.4 Å. *Nature* **483**, 489–493 (2012).

International Transporter Consortium Commentary on Clinically Important Transporter Polymorphisms

KM Giacomini¹, PV Balimane², SK Cho³, M Eadon⁴, T Edeki⁵, KM Hillgren⁶, S-M Huang⁷, Y Sugiyama⁸, D Weitz⁹, Y Wen¹⁰, CQ Xia¹¹, SW Yee¹, H Zimdahl¹² and M Niemi¹³; on behalf of the International Transporter Consortium

This Commentary focuses on genetic polymorphisms in membrane transporters. We present two polymorphisms for which there is a compelling body of literature supporting their clinical relevance: OATP1B1 (c.521T>C, p.V174A, rs4149056) and BCRP (c.421C>A, p.Q141K, rs2231142). The clinical evidence demonstrating their role in variation in pharmacokinetics and pharmacodynamics is described along with their allele frequencies in ethnic populations. Recommendations for incorporating studies of transporter polymorphisms in drug development are provided, along with the regulatory implications.

Membrane transporters play a critical role in drug absorption, distribution, elimination, response, and toxicity. A recent white paper from the International Transporter Consortium (ITC) focused on transporter-mediated drug–drug interactions (DDIs), which result in varied plasma and tissue concentrations of drugs and corresponding safety and efficacy issues. Seven transporters with considerable evidence from the literature for their involvement in clinical DDIs were described initially¹ and have been included in recent US Food and Drug Administration (FDA)

and European Medicines Agency (EMA) guidelines, which include additional transporters.^{2,3} The focus of the white paper and the guidances from the FDA and EMA is on DDIs.

In this Commentary, we discuss transporter polymorphisms in drug development, which have not previously been addressed.¹ Similar to DDIs, genetic polymorphisms in transporters may cause variability in drug concentrations and associated toxicities or response. Our focus here is on polymorphisms in two transporters, OATP1B1 (*SLCO1B1*

¹Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, California, USA; ²Bristol-Myers Squibb Company, Princeton, New Jersey, USA; ³Department of Pharmacology, Yonsei University College of Medicine, Seoul, South Korea; ⁴Section of Nephrology, Department of Medicine, The University of Chicago, Chicago, Illinois, USA; ⁵AstraZeneca Pharmaceuticals, Wilmington, Delaware, USA; ⁶Drug Disposition, Lilly Research Laboratories, Indianapolis, Indiana, USA; ⁷Office of Clinical Pharmacology, Office of Translational Sciences, Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, Maryland, USA; ⁸RIKEN Innovation Center, Research Cluster for Innovation, RIKEN (The Institute of Physical and Chemical Research), Yokohama City, Kanagawa, Japan; ⁹Research and Development Drug Disposition, Sanofi-Aventis Deutschland, Frankfurt, Germany; ¹⁰Section of Hematology/Oncology, The University of Chicago, Chicago, Illinois, USA; ¹¹Millennium Pharmaceuticals, Inc., Cambridge, Massachusetts, USA; ¹²Boehringer-Ingelheim Pharma GmbH & Co KG, Biberach, Germany; ¹³Department of Clinical Pharmacology, University of Helsinki, Helsinki, Finland. Correspondence: KM Giacomini (kathy.giacomini@ucsf.edu) or M Niemi (mikko.niemi@helsinki.fi)

doi:10.1038/clpt.2013.12