# DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

„ Analysis of SLC families according to substrate similarity”

verfasst von / submitted by

Catrin Gabrail

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Magistra der Pharmazie (Mag. Pharm.)

Wien, 2021 / Vienna, 2021

| | |
|---|---|
| Studienkennzahl lt. Studienblatt / The degree programme code as it appears on the student record sheet: | A 449 |
| Studienrichtung lt. Studienblatt / degree programme code as it appears on the student recod sheet: | Diplomstudium Pharmazie |
| Betreut von / Supervisor: | Univ.-Prof. Mag. Dr. Gerhard Ecker |
| Mitbetreut von / Co-Supervisor: | Dipl.-Ing. (FH) Dr. Daniela Digles |

Danksagung

An dieser Stelle möchte ich mich bei all denjenigen bedanken, die durch ihre Unterstützung das Gelingen dieser Arbeit erst möglich gemacht haben.

Zuerst gebührt mein Dank Herrn Univ. Prof. Mag. Dr. Gerhard Ecker, der mir ermöglicht hat eine Diplomarbeit bei ihm zu machen.

Vor allem möchte ich mich bei meiner Mitbetreuerin Dipl. Ing. Dr. Daniela Digles bedanken, für die hilfreichen Anregungen und die konstruktive Kritik bei der Erstellung dieser Arbeit. Sie stand mir von Anfang an helfend zur Seite und hat sich jedes Mal genügend Zeit genommen, wenn ich nicht weiterwusste.

Ebenfalls möchte ich mich bei der Pharmacoinformatics Research Group bedanken, an die ich mich wenden konnte, wenn ich etwas nicht wusste.

Abschließend geht mein ganz besonderer Dank an meine Mama und meinen Onkel Maged, die mir immer wieder Zuversicht gegeben haben, auch in schwierigen Phasen nicht aufzugeben und mein Ziel konsequent weiterzuverfolgen.

Abstract:

The SLC, Solute Carriers, are transmembrane proteins responsible for the transport of solutes into and out of the cell. The transport of exogenous and endogenous substances occurs through passive and secondary active transport. There are 70 different SLC families with about 446 different human transporter genes that play an important role in both health and disease of the human organism. SLC families are classified according to their biological function.

The goal of my work was to create an automated workflow, using the KNIME Analytics Platform, to collect and later analyze data from four databases. Based on the data, an SLC substrate list was generated to analyze the SLC families according to the similarity of the substrates. Data were obtained from the Metrabase, IUPHAR/BPS, UCSF-FDA, and BioParadigms databases. A name to structure search was performed using KNIME, which obtained various structural information from PubChem, Chemical Identifier Resolver, and Chemical Translation Service. The data were compared and evaluated.

## Zusammenfassung:

Die SLC, solute carrier, sind Transmembranproteine, die für den Transport von gelösten Stoffen in und aus der Zelle verantwortlich sind. Der Transport von exogenen und endogenen Substanzen erfolgt durch passiven und sekundären aktiven Transport. Dabei handelt es sich um 70 verschiedene SLC-Familien mit etwa 446 unterschiedlichen menschlichen Transporter Genen, die sowohl für die Gesundheit als auch verschiedenen Krankheiten des menschlichen Organismus eine wichtige Rolle spielen. Die SLC-Familien werden nach ihrer biologischen Funktion klassifiziert.

Das Ziel meiner Arbeit war es, einen automatisierten Arbeitsablauf mit Hilfe der KNIME Analytics Platform zu erstellen, um Daten aus vier Datenbanken zu sammeln und später zu analysieren. Anhand der Daten wurde eine SLC-Substratliste erstellt, um die SLC-Familien nach der Ähnlichkeit der Substrate zu analysieren. Die Daten stammen aus den Datenbanken Metrabase, IUPHAR/BPS, UCSF-FDA und BioParadigms. Eine Struktursuche wurde mit KNIME durchgeführt, die verschiedene Strukturinformationen von PubChem, Chemical Identifier Resolver und Chemical Translation Service erhielt. Die Daten wurden miteinander verglichen und ausgewertet.

# Table of Content


1 Introduction ..... 8

2 Materials and Methods ..... 10

2.1 Databases ..... 10

2.1.1 Metabolism and Transport Database (Metrabase) ..... 10

2.1.2 UCSF-FDA ..... 12

2.1.3 BioParadigms ..... 13

2.1.4 IUPHAR/BPS Guide to PHARMACOLOGY (GtoPdb) ..... 14

2.1.5 Chemical Identifier Resolver ..... 15

2.1.6 Chemical Translation Service ..... 16

2.1.7 PubChem ..... 17

2.2 Tools ..... 18

2.2.1 KNIME Analytics Platform ..... 18

2.2.2 Workflow ..... 19

2.2.2.1 Collecting data from the databases ..... 20

2.2.2.2 Find standardized molecular structure ..... 33

2.2.2.3 Merge SLC families and substrates ..... 37

2.2.2.4 Visualization ..... 38

3 Results and Discussion ..... 40

4 Appendix ..... 46

Table list ..... 111

References ..... 114

# 1 Introduction

There are many different ways to transport endogenous and exogenous substances into and out of a cell, e.g., by active transport of hydrophilic substances across biomembranes such as ion pumps, or by transmembrane transporters for the passive transport of hydrophilic substances across biomembranes such as carrier proteins and ion channels. One example of carrier proteins is solute carriers, also abbreviated as SLC.

First of all, I would like to introduce you to the topic of solute carriers and explain to you which functions they have in the human body. Solute carrier proteins (SLCs) are transmembrane transporters that facilitate solute influx and efflux. SLCs are part of the passive and secondary active substrate transporters. For passive transport, no energy input is required; for secondary active transport, ion movement along an electrochemical gradient is necessary. They are divided into 70 SLC families with about 446 membrane proteins[1], and new transporter genes are constantly being discovered. The SLC transporter superfamily are classified according to its biological function rather than their sequence homology[2]. Nevertheless, members of the same SLC transporter family that bind similar substrates exhibit at least 20% protein sequence identity.[3]

SLC transporters are highly expressed in organs such as kidney, liver, brain and intestine, which are metabolically active. They affect the development of some diseases such as Parkinson's disease, Alzheimer's disease, epilepsy or schizophrenia, which are due to dysfunctions of the neurotransmitter transporters of the SLC1 or SLC6 family[4,5]. Therefore, they play an important role in drug discovery[6], as their dysfunction has implications for the development of several diseases.

There are several databases dealing with solute carriers, such as IUPHAR/BPS Guide to Pharmacology (IUPHAR/BPS)[7], SLC tables BioParadigms[8], DrugBank[9], ChEMBL[10], UCSF-FDA[11] and Metabolism and Transporter Database (Metrabase)[12]. In my diploma thesis I dealt with these four databases: Metrabase, IUPHAR/BPS, UCSF-FDA and BioParadigms.

My focus was to collect data from these four databases on solute carrier and their substrates and merge them to see if there is an overlap in the data, for example, if the same SLCs are mentioned but not with the same substrates. As certain differences appear between the databases, such as Metrabase which gives information about the structure, in the form of SMILES and InChIs and UCSF-FDA which does not reveal this information, I had to do a name to structure search to get this information. In addition, not all databases allowed access to the data in the same way, for example, in the BioParadigms database it was not possible to download the data, but it was necessary to copy the data manually and collect them externally in an Excel spreadsheet in order to work with them. On the other hand, Metrabase made it very easy to access the data, as it was possible to download the data for each SLC as a text file. Thus, data collection was easier for some databases and more difficult for others. My goal of the diploma thesis was to analyze the SLC families according to their substrate similarity. I worked with the data in the form of a workflow created with the freely available software program KNIME Analytics Platform[13].

Sarah Foth, a former pharmacy student and diploma student of Prof. Dr. Gerhard Ecker, also dealt with SLC and their substrates. The topic of her diploma thesis[14] was "An Analysis of SLC Transporter Substrates.". She also forms a workflow with the software program KNIME. In contrast to my work, she dealt with one substrate per SLC and worked with data from the BioParadigms SLCtable and IUPHAR/BPS databases.  Descriptors were calculated for the structures and a tree diagram was created to visualize them.

Another example of working with SLC was Meixner et al[1]. In contrast to my work, Meixner et al. searched for substrates manually, while I tried to get the information as automatically as possible.

In the following chapters I would like to present my work and give you an overview of the solute carrier families and their substrates.

# 2 Materials and Methods

## 2.1 Databases

### 2.1.1 Metabolism and Transport Database (Metrabase)

The Metabolism and Transport Database (Metrabase)[12] is concerned with cheminformatics and bioinformatics data. Metrabase combines various resources, such as ChEMBL, DrugBank and UniProt, as well as manually collected data from published literature. The data offers information about the interactions among proteins and chemical substances associated with their metabolic and transporter properties. The resource provides structural and physicochemical information of molecules interacting with transporters and metabolic enzymes. The biological interaction is classified as action type (e.g. substrate or inhibitor).

Metrabase is certified under a Creative Commons Attribution-ShareAlike 4.0 International License. However, the incorporated information keeps the licensing of the original data assets.

| Protein symbol | Gene symbol | Name | Activity count | Compound count | UniProt ID | HGNC ID |
|---|---|---|---|---|---|---|
| ASBT | SLC10A2 | solute carrier family 10 (sodium/bile acid cotransporter family) member 2 | 89 | 82 | Q12908 | 10906 |
| BCRP1 | ABCG2 | ATP-binding cassette sub-family G (WHITE) member 2 | 1,960 | 980 | Q9UNQ0 | 74 |
| GLUT1 | SLC2A1 | solute carrier family 2 (facilitated glucose transporter), member 1 | 110 | 83 | P11166 | 11005 |
| LAT1 | SLC7A5 | solute carrier family 7 (amino acid transporter light chain, L system), member 5 | 14 | 6 | Q01650 | 11063 |
| MCT1 | SLC16A1 | solute carrier family 16 member 1 (monocarboxylic acid transporter 1) | 50 | 38 | P53985 | 10922 |
| MDR1 | ABCB1 | ATP-binding cassette sub-family B (MDR/TAP) member 1 | 2,459 | 1,182 | P08183 | 40 |
| MRP1 | ABCC1 | ATP-binding cassette, sub-family C (CFTR/MRP), member 1 | 233 | 194 | P33527 | 51 |
| MRP2 | ABCC2 | ATP-binding cassette sub-family C (CFTR/MRP) member 2 | 999 | 519 | Q92887 | 53 |
| MRP3 | ABCC3 | ATP-binding cassette sub-family C (CFTR/MRP) member 3 | 322 | 133 | O15438 | 54 |
| MRP4 | ABCC4 | ATP-binding cassette, sub-family C (CFTR/MRP), member 4 | 71 | 66 | O15439 | 55 |
| OATP1A2 | SLCO1A2 | solute carrier organic anion transporter family, member 1A2 | 188 | 112 | P46721 | 10956 |
| OATP1B1 | SLCO1B1 | solute carrier organic anion transporter family member 1B1 | 1,105 | 444 | Q9Y6L6 | 10959 |
| OATP1B3 | SLCO1B3 | solute carrier organic anion transporter family member 1B3 | 629 | 344 | Q9NPD5 | 10961 |
| OATP2A1 | SLCO2A1 | solute carrier organic anion transporter family member 2A1 | 18 | 16 | Q92959 | 10955 |
| OATP2B1 | SLCO2B1 | solute carrier organic anion transporter family, member 2B1 | 758 | 365 | O94956 | 10962 |
| OATP3A1 | SLCO3A1 | solute carrier organic anion transporter family member 3A1 | 21 | 13 | Q9UIG8 | 10952 |
| OATP4A1 | SLCO4A1 | solute carrier organic anion transporter family member 4A1 | 12 | 9 | Q96BD0 | 10953 |
| OCT1 | SLC22A1 | solute carrier family 22 (organic cation transporter) member 1 | 1,057 | 506 | O15245 | 10963 |
| OSTA | SLC51A | solute carrier family 51, alpha subunit | 0 | 0 | Q86UW1 | 29955 |
| OSTA/OSTB | | solute carrier family 51 | 6 | 6 | | |
| OSTB | SLC51B | solute carrier family 51, beta subunit | 0 | 0 | Q86UW2 | 29956 |
| PEPT1 | SLC15A1 | solute carrier family 15 (oligopeptide transporter) member 1 | 1,042 | 475 | P46059 | 10920 |

*Figure 1 - Transporter protein list of the Metabolism and Transporter Database. The list includes data about solute carries and ATP-binding cassette sub-families (25.5.21).*

The database contains a protein list of transporter protein, such as solute carriers (SLCs) and ATP-binding cassette sub families (ABC-proteins), and metabolic proteins, such as Cytochrome P450s (CYPs). The website offers the possibility to select their substrates, non-substrates, repressors, inhibitors/non-inhibitors, stimulators, inducers, and non-inducers depending on the protein, as well as a page for the tissue expression data.  My thesis dealt with the solute carriers and their substrates.

To get an overview of the proteins in this database, they are presented in a table on their website. In this table the gene symbol, protein symbol, tissue, expression type, activity count, compound count, Uniprot ID and HGNC ID are given for each protein, as it is shown in

| Protein symbol | Gene symbol | Name | Activity count | Compound count | UniProt ID | HGNC ID |
|---|---|---|---|---|---|---|
| ASBT | SLC10A2 | solute carrier family 10 (sodium/bile acid cotransporter family) member 2 | 89 | 82 | Q12908 | 10906 |
| BCRP1 | ABCG2 | ATP-binding cassette sub-family G (WHITE) member 2 | 1,960 | 980 | Q9UNQ0 | 74 |
| GLUT1 | SLC2A1 | solute carrier family 2 (facilitated glucose transporter), member 1 | 110 | 83 | P11166 | 11005 |
| LAT1 | SLC7A5 | solute carrier family 7 (amino acid transporter light chain, L system), member 5 | 14 | 6 | Q01650 | 11063 |
| MCT1 | SLC16A1 | solute carrier family 16 member 1 (monocarboxylic acid transporter 1) | 50 | 38 | P53985 | 10922 |
| MDR1 | ABCB1 | ATP-binding cassette sub-family B (MDR/TAP) member 1 | 2,459 | 1,182 | P08183 | 40 |
| MRP1 | ABCC1 | ATP-binding cassette, sub-family C (CFTR/MRP), member 1 | 233 | 194 | P33527 | 51 |
| MRP2 | ABCC2 | ATP-binding cassette sub-family C (CFTR/MRP) member 2 | 999 | 519 | Q92887 | 53 |
| MRP3 | ABCC3 | ATP-binding cassette sub-family C (CFTR/MRP) member 3 | 322 | 133 | O15438 | 54 |
| MRP4 | ABCC4 | ATP-binding cassette, sub-family C (CFTR/MRP), member 4 | 71 | 66 | O15439 | 55 |
| OATP1A2 | SLCO1A2 | solute carrier organic anion transporter family, member 1A2 | 188 | 112 | P46721 | 10956 |
| OATP1B1 | SLCO1B1 | solute carrier organic anion transporter family member 1B1 | 1,105 | 444 | Q9Y6L6 | 10959 |
| OATP1B3 | SLCO1B3 | solute carrier organic anion transporter family member 1B3 | 629 | 344 | Q9NPD5 | 10961 |
| OATP2A1 | SLCO2A1 | solute carrier organic anion transporter family member 2A1 | 18 | 16 | Q92959 | 10955 |
| OATP2B1 | SLCO2B1 | solute carrier organic anion transporter family, member 2B1 | 758 | 365 | O94956 | 10962 |
| OATP3A1 | SLCO3A1 | solute carrier organic anion transporter family member 3A1 | 21 | 13 | Q9UIG8 | 10952 |
| OATP4A1 | SLCO4A1 | solute carrier organic anion transporter family member 4A1 | 12 | 9 | Q96BD0 | 10953 |
| OCT1 | SLC22A1 | solute carrier family 22 (organic cation transporter) member 1 | 1,057 | 506 | O15245 | 10963 |
| OSTA | SLC51A | solute carrier family 51, alpha subunit | 0 | 0 | Q86UW1 | 29955 |
| OSTA/OSTB | | solute carrier family 51 | 6 | 6 | | |
| OSTB | SLC51B | solute carrier family 51, beta subunit | 0 | 0 | Q86UW2 | 29956 |
| PEPT1 | SLC15A1 | solute carrier family 15 (oligopeptide transporter) member 1 | 1,042 | 475 | P46059 | 10920 |

. To access the data, I had the option of downloading a data file for each SLC in text format. The file contains the protein name, the compound ID, SMILES, InChI, reference ID, DOI, and assay description, as you can see in the Figure 2.

GLUT1-SLC2A1- Metrabase 26.11.20 - Editor

Datei Bearbeiten Format Ansicht Hilfe

```
Protein Action  Compound ID     Compound Name   SMILES  InChI   Reference ID    DOI     Assay Description
GLUT1   substrate       mbcd0029202     3-O-methyl-D-Glucose    CO[C@H]([C@@H](O)C=O)[C@H](O)[C@H](O)CO InChI=1S/C7H14O6/c1-13-7(5(11)3-9)6(12)4(10)2-8/h3-8,10-12H,2H2,1H3/t4-,5+,6-,7-/m1/s
GLUT1   substrate       mbcd0029202     3-O-methyl-D-Glucose    CO[C@H]([C@@H](O)C=O)[C@H](O)[C@H](O)CO InChI=1S/C7H14O6/c1-13-7(5(11)3-9)6(12)4(10)2-8/h3-8,10-12H,2H2,1H3/t4-,5+,6-,7-/m1/s
GLUT1   substrate       mbcd0029203     D-glucose       OC[C@@H](O)[C@@H](O)[C@H](O)[C@@H](O)C=O        InChI=1S/C6H12O6/c7-1-3(9)5(11)6(12)4(10)2-8/h1,3-6,8-12H,2H2/t3-,4+,5+,6+/m0/s1
GLUT1   substrate       mbcd0029203     D-glucose       OC[C@@H](O)[C@@H](O)[C@H](O)[C@@H](O)C=O        InChI=1S/C6H12O6/c7-1-3(9)5(11)6(12)4(10)2-8/h1,3-6,8-12H,2H2/t3-,4+,5+,6+/m0/s1
GLUT1   substrate       mbcd0029198     2-deoxy-d-glucose (acyclic)     OC[C@@H](O)[C@@H](O)[C@H](O)CC=O        InChI=1S/C6H12O5/c7-2-1-4(9)6(11)5(10)3-8/h2,4-6,8-11H,1,3H2/t4-,5-,6+/m1/s1
GLUT1   substrate       mbcd0029202     3-O-methyl-D-Glucose    CO[C@H]([C@@H](O)C=O)[C@H](O)[C@H](O)CO InChI=1S/C7H14O6/c1-13-7(5(11)3-9)6(12)4(10)2-8/h3-8,10-12H,2H2,1H3/t4-,5+,6-,7-/m1/s1
GLUT1   substrate       mbcd0029198     2-deoxy-d-glucose (acyclic)     OC[C@@H](O)[C@@H](O)[C@H](O)CC=O        InChI=1S/C6H12O5/c7-2-1-4(9)6(11)5(10)3-8/h2,4-6,8-11H,1,3H2/t4-,5-,6+/m1/s1
GLUT1   substrate       mbcd0000327     2-Deoxyglucose  OC[C@H]1O[C@H](O)C[C@@H](O)[C@@H]1O     InChI=1S/C6H12O5/c7-2-4-6(10)3(8)1-5(9)11-4/h3-10H,1-2H2/t3-,4-,5+,6+/m1/s1     mbrf0002593
GLUT1   substrate       mbcd0029213     Dehydroascorbic acid    OC[C@H](O)[C@H]1OC(=O)C(=O)C1=O InChI=1S/C6H6O6/c7-1-2(8)5-3(9)4(10)6(11)12-5/h2,5,7-8H,1H2/t2-,5+/m0/s1        mbrf0002593
GLUT1   substrate       mbcd0029203     D-glucose       OC[C@@H](O)[C@@H](O)[C@H](O)[C@@H](O)C=O        InChI=1S/C6H12O6/c7-1-3(9)5(11)6(12)4(10)2-8/h1,3-6,8-12H,2H2/t3-,4+,5+,6+/m0/s1
GLUT1   substrate       mbcd0029203     D-glucose       OC[C@@H](O)[C@@H](O)[C@H](O)[C@@H](O)C=O        InChI=1S/C6H12O6/c7-1-3(9)5(11)6(12)4(10)2-8/h1,3-6,8-12H,2H2/t3-,4+,5+,6+/m0/s1
```

*Figure 2 – Download as text file from the Metabolism and Transporter Database for the solute carrier SLC2A1.*

## 2.1.2 UCSF-FDA

The University of California, San Francisco-Food and Drug Administration (UCSF-FDA) [11]TransPortal is a public drug transporter database. It focuses on 31 drug transporters, which have a role in the disposition, toxicity, and efficacy, from the ATP-binding cassette (ABC) and solute carrier (SLC) transporter superfamilies.

For each transporter, literature was collected on its expression levels, subcellular localization, and direction of transporter in the liver, kidney, small intestine, placenta, and blood-brain barrier, as you can see in Figure 3. They also listed for each transporter in-vitro substrates and in-vitro inhibitors and summarized transport kinetic data like Km, Ki and IC50.

However, the UCSF-FDA transporter database does not include information about the structure of the substrates, like InChI, InChI Keys or Smiles and the possibility of downloading the information as a table is not given either.

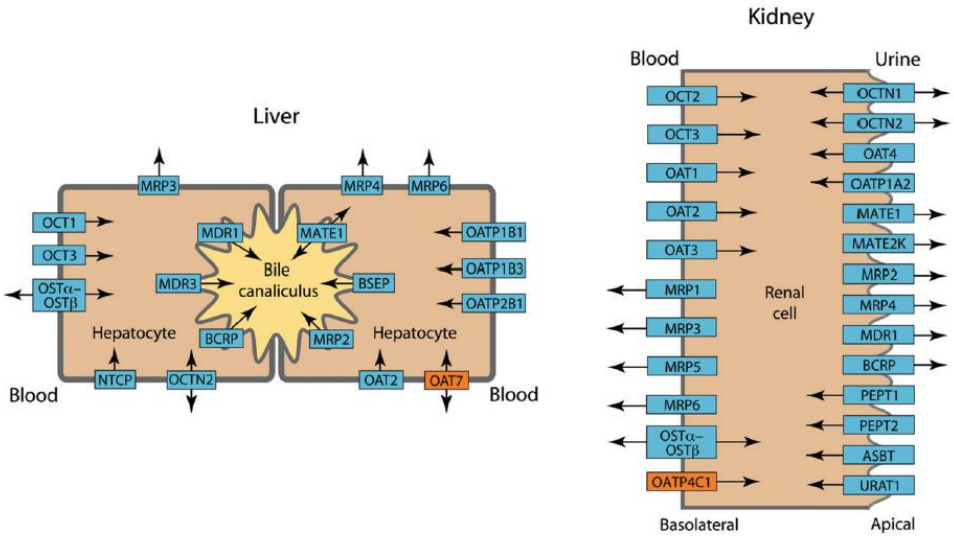



*Figure 3 – UCSF-FDA database: drug transporters in human liver and kidney – PMC Copyright Notice (public domain material) (25.5.21)*

## 2.1.3 BioParadigms

The BioParadigms database[2,15] is a scientific resource, whose aim is to allow the exchange of information between membrane biologists in academia and industry. It was launched by Prof. Matthias A. Hediger, University of Bern, Switzerland in May 2004.

The database contains the Genomic Transporter - SLC series[8], which consists of 65 SLC families with 458 different human transporter genes. The SLC members are grouped in tables by their families. The table contains SLC name, protein name, aliases, transport type, substrates, tissue, and cellular expression, an example is shown in Figure 4. Similar to the UCSF-FDA database, this transporter database does not include information about the structure of the substrates, as well as it does not provide the possibility of downloading the information as a table.

[-] **SLC1** **High-affinity glutamate and neutral amino acid transporter family**

**References**

*Original version of the SLC table:*

**Kanai Y, Clémençon B, Simonin A, Leuenberger M, Lochner M, Weisstanner M, Hediger MA.** The SLC1 high-affinity glutamate and neutral amino acid transporter family. Mol Aspects Med. 2013 Apr;34(2-3):108-20.

| SLC name | Protein name | Aliases | Transport type* | Substrates | Tissue and cellular expression |
|---|---|---|---|---|---|
| SLC1A1 | EAAC1, EAAT3 | System X-AG | C / Na+, H+, K+ | L-Glu, D/L-Asp | brain (neurons), intestine, kidney, liver, heart, placenta |
| SLC1A2 | GLT-1, EAAT2 | System X-AG | C / Na+, H+, K+ | L-Glu, D/L-Asp | brain (astrocytes, Bergmann glia, neurons), liver, pancreas |
| SLC1A3 | GLAST, EAAT1 | System X-AG | C / Na+, H+, K+ | L-Glu, D/L-Asp | brain (astrocytes, Bergmann glia), heart, skeletal muscle, placenta |
| SLC1A4 | ASCT1, SATT | System ASC | C / Na+, E / amino acids | L-Ala, L-Ser, L-Cys, L-Thr | widespread |
| SLC1A5 | ASCT2, AAAT | System ASC | C / Na+, E / amino acids | L-Ala, L-Ser, L-Cys, L-Thr, L-Gln, L-Asn | lung, skeletal muscle, large intestine, kidney, testis, adipose tissue |
| SLC1A6 | EAAT4 | System X-AG | C / Na+, H+, K+ | L-Glu, D/L-Asp | cerebellum (Purkinje cells) |
| SLC1A7 | EAAT5 | System X-AG | C / Na+, H+, K+ | L-Glu, D/L-Asp | retina (rod photoreceptors, bipolar cells) |

*Figure 4 – SLC list of the Bioparadigms Database. The list includes data about SLC1 – High-affinity glutamate and neutral amino acid transporter family (06.07.21).*

## 2.1.4 IUPHAR/BPS Guide to PHARMACOLOGY (GtoPdb)

The IUPHAR/BPS is a cooperation between the International Union of Basic and Clinical Pharmacology (IUPHAR) and the British Pharmacological Society (BPS).[7,16] This public database provides information about ligand-activity-target relationships from drug targets, prescription medicines and experimental drugs.

It contains a list of SLC superfamily of solute carriers. For a closer look and even more information, you can click on the SLC transporter name.

For each SLC transporter target ID, nomenclature, systematic nomenclature, common abbreviation, previous and unofficial names, genes, ensemble ID, UniProtKB AC, Bioparadigms SLC tables, endogenous substrates, substrates, inhibitors, labelled ligands and stoichiometry where known are given (Figure 5). [16]

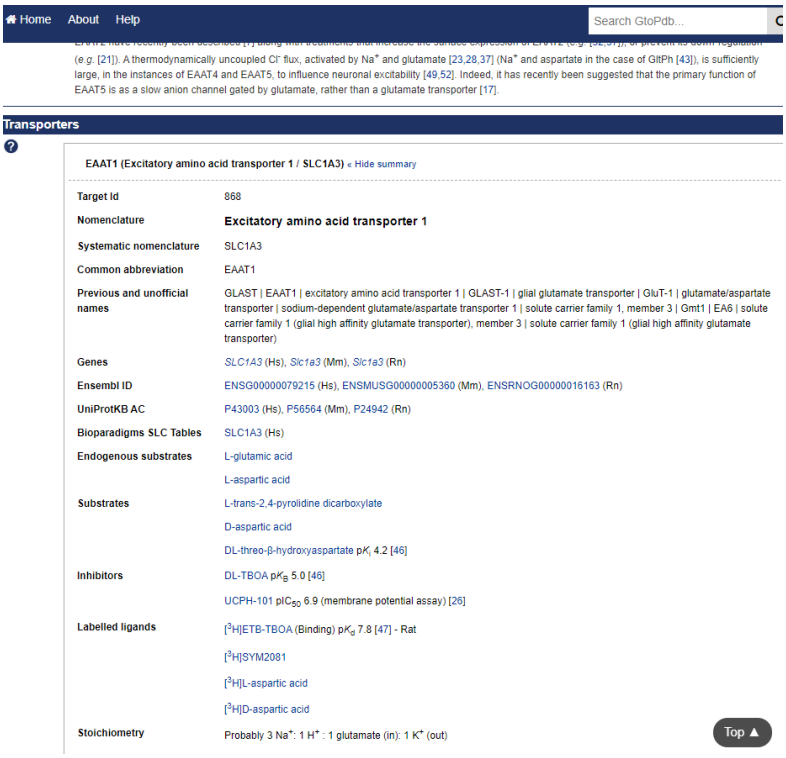



*Figure 5 – Example page of the IUPHAR/BPS database.*

The ligand information includes among others ligand ID, name, 2D structure, calculated physico-chemical properties, summary with classification, IUPAC name and database links, biological activity, references, structure, canonical and isomeric SMILES, InChI and InChIKey and similar ligands.[17]

### 2.1.5 Chemical Identifier Resolver

The Chemical Identifier Resolver is a freely accessible website. The website is powered by the Computer-Aided Drug Design group at the National Cancer Institute (NCI/CADD).[18]

Through the URL API https://cactus.nci.nih.gov/chemical/structure/"structure identifier"/"representation" it is possible to get access to the website.[19]

The Structure Identifier gives you the possibility to convert the name of a structure into different identifiers, such as InChI, InChIKey, SMILES. In addition, descriptors can be calculated for the molecules, an example is given in Figure 6.

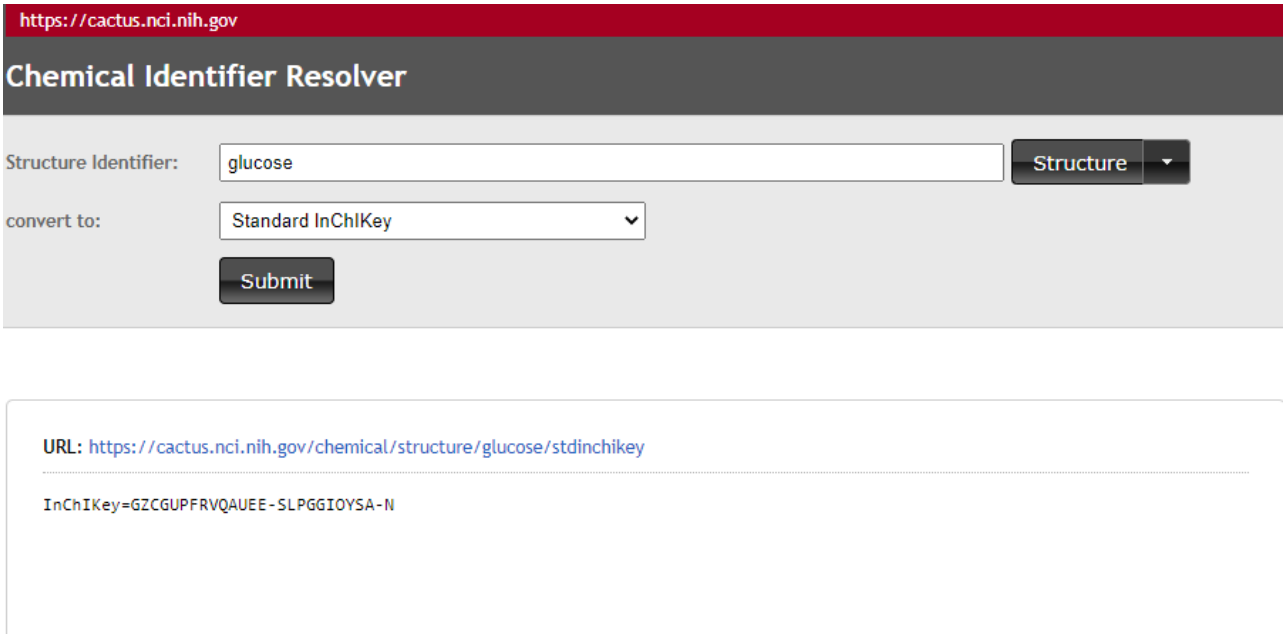



*Figure 6 – Search Interface of the Chemical Identifier Resolver.*

## 2.1.6 Chemical Translation Service

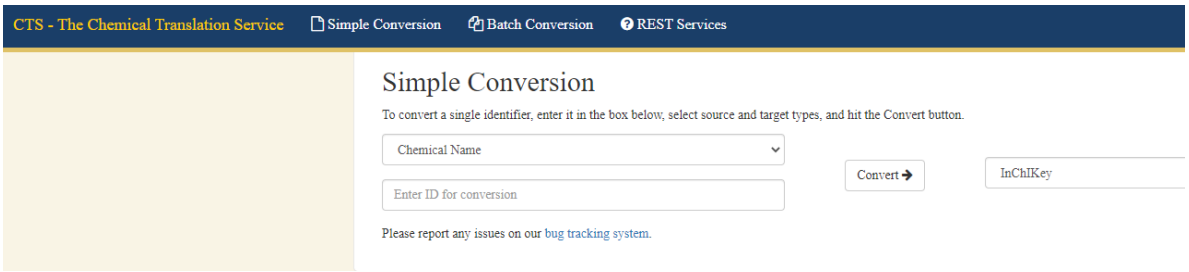



*Figure 7 – Search Interface of the Chemical Translation Service.*

The Chemical Translation Service is like the Chemical Identifier Resolver in respect of the conversion. Here it is also possible to use the name of the structure to convert it into various properties, for example name to InChI, InChIKey or SMILES and the other way round (Figure 7). The second similar aspect is the access the website through several rest services (Figure 8).[20]

The database is powered by the Fiehn laboratory at UC Davis.

Get a list of scored InChIKeys

| Title | Score |
|---|---|
| URL | `/rest/score/{from}/{query}/{algorithm}` |
| Method | GET |
| URL Params | `from=[string]`<br>The source ID type<br>`query=[string]`<br>The ID to convert<br>`algorithm=[string]`<br>The scoring algorithm to use, either 'biological' or 'popularity' |
| Sample Call | `curl https://cts.fiehnlab.ucdavis.edu/rest/score/Chemical%20Name/threonine/biological` |
| Sample Response | {<br>"searchTerm": "threonine",<br>"from": "chemical name",<br>"result":<br>[<br>{<br>"InChIKey": "AYFVYJQAPQTCCC-GBXIJSLDSA-N",<br>"score": 1<br>},<br>{<br>"InChIKey": "AYFVYJQAPQTCCC-STHAYSLISA-N",<br>"score": 0.86<br>},<br>{<br>"InChIKey": "AYFVYJQAPQTCCC-UHFFFAOYSA-N",<br>"score": 0<br>}<br>]<br>} |

*Figure 8 – Chemical Translation Service_ Restweb Service for InChIkeys.*

### 2.1.7 PubChem

PubChem[21] is a collection of scientific data from different databases (681 sources), such as DrugBank, ChEMBL. The chemistry database is powered by National Institutes of health (NIH).

PubChem includes a variety of different molecules, and should therefore be well suited to provide structures for SLC substrates, such as small molecules, peptides, etc. It provides information about the molecules, such as chemical structures, chemical and physical properties, biological activities, toxicity data.[22]

## 2.2 Tools

### 2.2.1 KNIME Analytics Platform

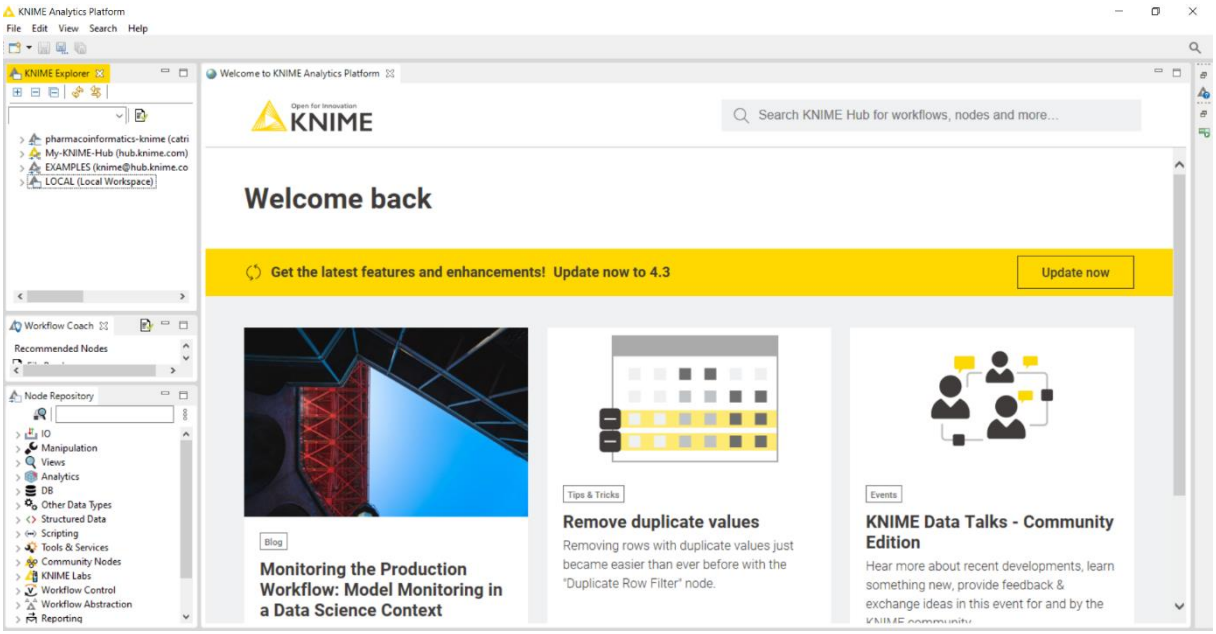



*Figure 9 - KNIME start screen page.*

KNIME Analytics Platform[23] is a free accessible software for creating data science from the KNIME AG. They have different locations like in Zurich, Berlin, Constance and Austin (USA).

It offers different tools, for example to build workflows to automatize processes in data analysis, to choose from nodes or to use publicly available workflows from the KNIME Hub[13].

The different nodes provide you with various options, such as merge files, visualize, delete duplicates, read files, save files and much more. If a green dot appears at the node, the step works, but if a yellow or red dot appears, the settings are not correct, and you need to change them.

In the following chapter I would like to introduce you to my workflow and take a closer look at it.

## 2.2.2 Workflow

The following figure (Figure 10) represents the KNIME workflow created for my diploma thesis. The aim of this workflow was combining different databases to obtain the reported substrates for each SLC and cluster them according to their structural similarity. The workflow to achieve this aim is divided into four parts: (1) Collecting data from the databases, (2) find a standardized molecular structure for a given substrate name, (3) merge SLC families and substrates, and (4) visualize the data. Each of these steps is described in more detail in the following sections.

I worked with the KNIME Analytics Platform Version 4.2.5.

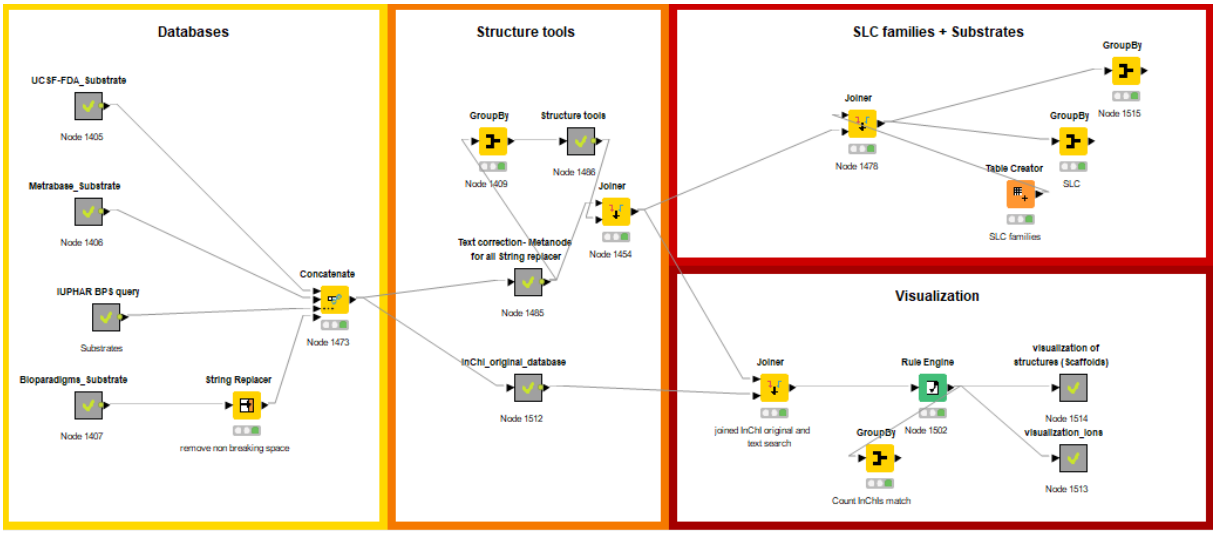



*Figure 10 – KNIME_workflow_Analysis of SLC families according to substrates similarity.*

### 2.2.2.1Collecting data from the databases

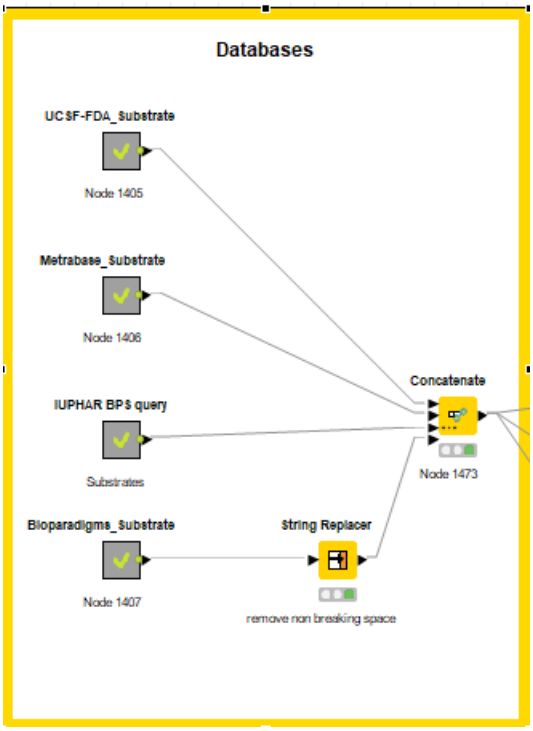



*Figure 11 – KNIME workflow_Data collection from UCSF-FDA, Metrabase, IUPHAR/BPS and Bioparadigms.*

I used four different databases to collect data about solute carrier families and their substrates: UCSF-FDA, Metrabase, IUPHAR/BPS and Bioparadigms.

The Metabolism and Transporter database allowes to download the data as a file, the UCSF-FDA database and the IUPHAR/BPS database do not have this possibility but are accessible via REST API, which stands for Representational State Transfer - Application Programming Interface and is a communication method to a webservice. In contrast to the two databases the Bioparadigms, do not have both access options and had to be manually entered into an excel spreadsheet.

#### 2.2.2.1.1 Metabolism and Transporter Database

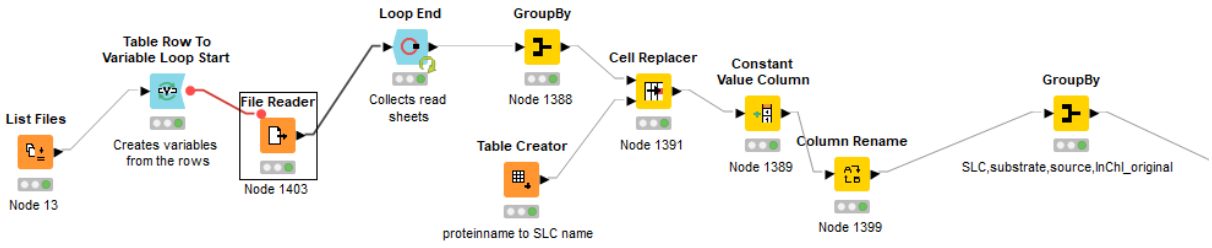



*Figure 12 – KNIME workflow collecting data from the Metabolism and Transporter Database.*

The Metabolism and Transport Database (Metrabase) offers the capability to download the substrates as text format for each SLC family. After manually

downloading the files, I added the entire folder to the List File node and merged all files via Table RowTo Variable Loop Strat node and terminated the loop through the Loop End node. Since Metrabase provides the SLC names as protein symbols rather than gene symbols, I had to use a Cell Replacer node and a Table Creator node to end up with a consistent SLC name. In the appendix you will find the replacement list (Table 9).

At the end of the workflow, I removed all duplicates through the GroupBy node (Figure 13) and received 14 different SLCs (SLC10A2, SLC2A1, SLC7A5, SLC16A1, SLCO1A2, SLCO1B1, SLCO1B3, SLCO2A1, SLCO2B1, SLCO3A1, SLCO4A1, SLC22A1, SLC51A-SLC51B and SLC15A1) and their substrates.

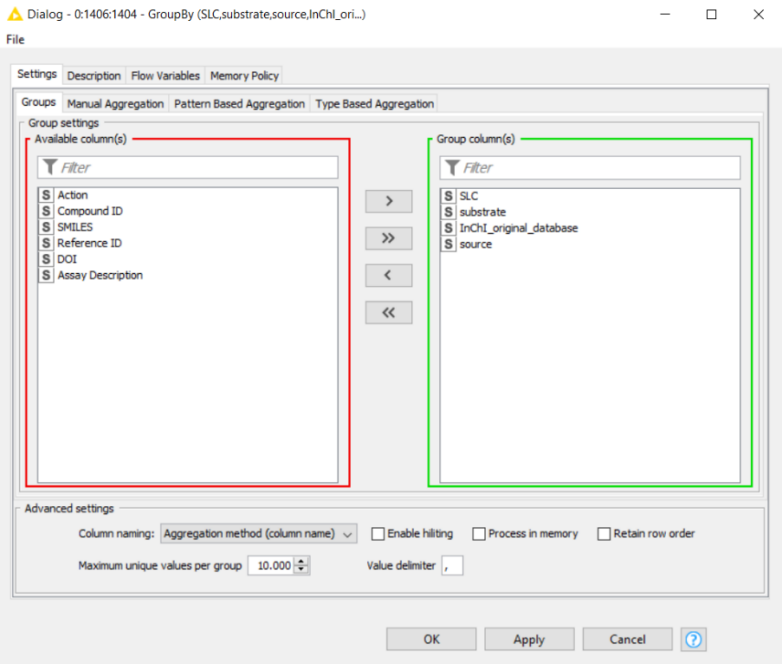



*Figure 13 – Configuring the GroupBy node for the workflow of the Metabolism and Transporter database data collection.*

***Table 1- Results of the Metabolism and Transporter database substrate collection.***

| SLC | Unique count*(substrate) |
|---|---|
| SLC10A2 | 54 |
| SLC15A1 | 247 |
| SLC16A1 | 8 |

| | |
|---|---|
| SLC22A1 | 166 |
| SLC2A1 | 5 |
| SLC51A-SLC51B | 4 |
| SLC7A5 | 2 |
| SLCO1A2 | 56 |
| SLCO1B1 | 97 |
| SLCO1B3 | 59 |
| SLCO2A1 | 5 |
| SLCO2B1 | 48 |
| SLCO3A1 | 5 |
| SLCO4A1 | 7 |

#### 2.2.2.1.2 **UCSF-FDA**

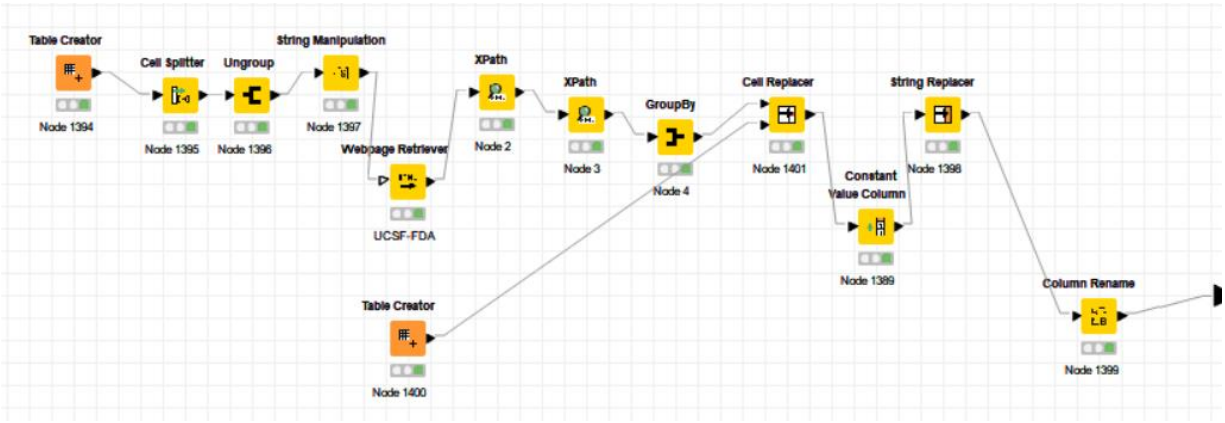



*Figure 14 – KNIME workflow collecting data from the UCSF-FDA database.*

The **UCSF-FDA database** had the ability to access the website through the URL, which was accessed via the Webpage Retriever node. In the String Manipulation node, I defined the expression, which was "join("http://transportal.compbio.ucsf.edu/transporters/",$UCSF-FDA SLC_SplitResultList$)", to get the URL for each SLC in the database. I received a list of URLs, one of them was for example http://transportal.compbio.ucsf.edu/transporters/SLC10A2. With the Xpath node I defined the information and the interface I needed from the web page. This is

because the UCSF- DFA database contains information not only about substrates, but also about inhibitors and more.

Since some substrate names were not consistent, I had to work with a Cell Replacer Node and a Table Creator Node here as well and replace some substrate names to avoid later problems by finding the structure, for example Bisglucuronosyl bilirubin to Bilirubin diglucuronide or 17beta-Glucuronosyl estradiol to Estradiol-17beta-glucuronide. In the appendix you will find the replacement list (Table 10).

By using the GroupBy node (Figure 15) I removed the duplicates and added through the Constant Value column node an extra column for the source. With this database I received the substrates for 20 SLCs. In ***Table 2*** you can see the results of the substrates collection.

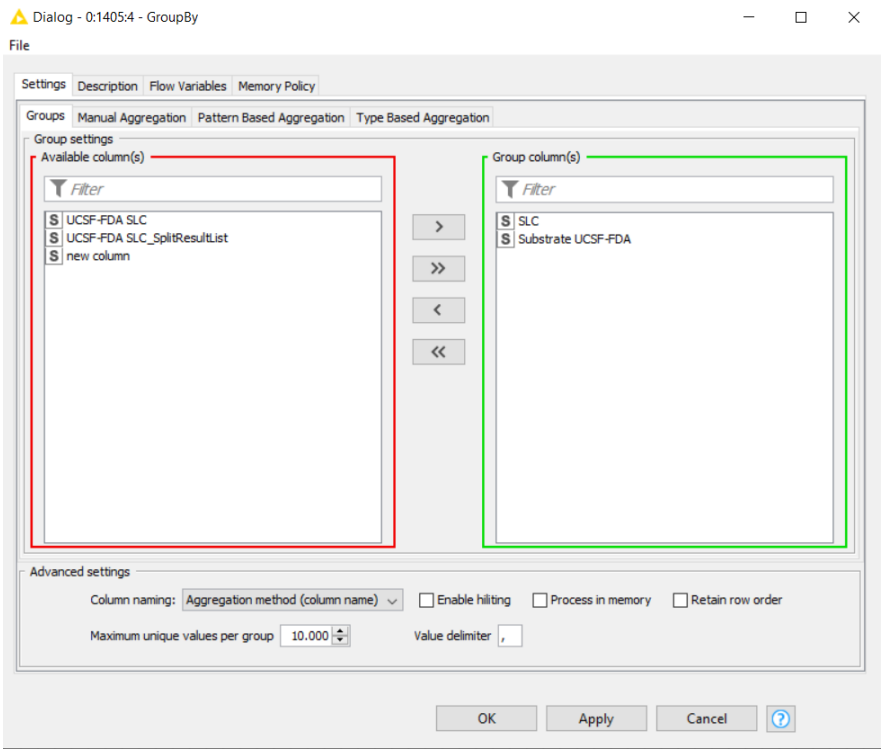



*Figure 15 – Configuring the GroupBy node for the workflow of the USCF-FDA data collection*

***Table 2 – Results of the UCSF-FDA database substrate collection.***

| SLC | substrates |
|---|---|
| SLC10A1 | 7 |
| SLC10A2 | 5 |
| SLC15A1 | 9 |
| SLC15A2 | 6 |
| SLC22A1 | 13 |
| SLC22A11 | 11 |
| SLC22A12 | 2 |
| SLC22A2 | 20 |
| SLC22A3 | 6 |
| SLC22A4 | 5 |
| SLC22A5 | 6 |
| SLC22A6 | 18 |
| SLC22A7 | 11 |
| SLC22A8 | 23 |
| SLC47A1 | 11 |
| SLC47A2 | 10 |
| SLCO1A2 | 13 |
| SLCO1B1 | 23 |
| SLCO1B3 | 26 |
| SLCO2B1 | 12 |

### 2.2.2.1.3 **Bioparadigms**

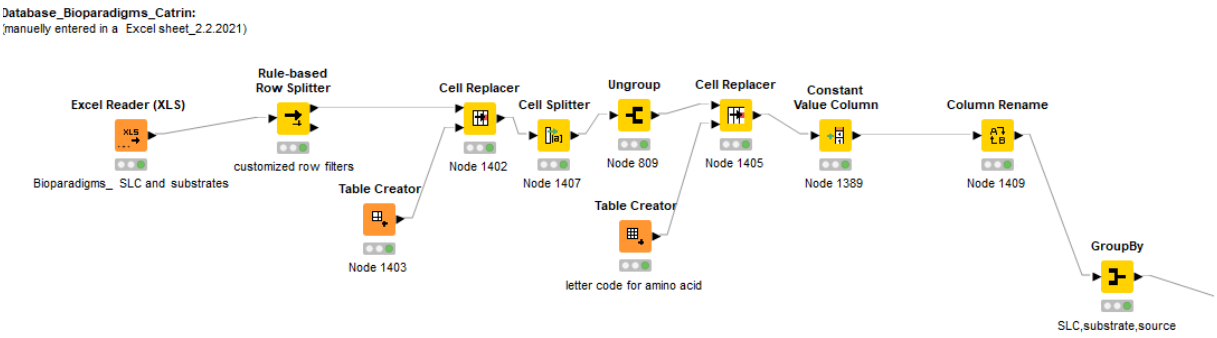



*Figure 16 – KNIME workflow collecting data from the Bioparadigms database.*

The **Bioparadigms** database did not give the option to access the website via KNIME, by using the API or downloading the information as a file. Here I had to manually enter the solute carriers and their substrates into an excel sheet (Table 12). By inserting the Excel spreadsheet into the Excel Reader node, KNIME was able to retrieve the file. The table was divided into a first output and a second output by the Rule-Based Row Splitter node. The expression for the Rule-based Row Splitter is shown in the appendix (**Table 11**). The work was continued with the first output table. In the next step, I used the Table Creator node and the Cell Replacer node to replace some rows, as an illustration chloride cicarbonate with chloride bicarbonate or urate (fructose, glucose) with urate, fructose, glucose (Table 13). The Cell Splitter node had the task of splitting each column into parts so that each SLC-substrate has its own row and then the duplicates were removed by the GroupBy node.

Furthermore, I created a table to replace the amino acids one letter code with the amino acid's names (Table 14). This database was much larger than the other two therefore I received here for 292 solute carriers substrates, which can be seen well in **Table 3**.

.

***Table 3 – Results of the Bioparadigms database substrate collection***

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| SLC25A38 | 1 | SLC22A16 | 2 | SLC8A2 | 2 | SLC5A3 | 2 |
| SLC25A37 | 1 | SLC22A13 | 2 | SLC8A1 | 2 | SLC5A2 | 1 |
| SLC25A36 | 1 | SLC22A12 | 2 | SLC7A9 | 2 | SLC5A12 | 1 |
| SLC25A33 | 1 | SLC22A11 | 1 | SLC16A8 | 1 | SLC5A11 | 2 |
| SLC25A32 | 1 | SLC22A1 | 1 | SLC7A8 | 4 | SLC5A10 | 3 |
| SLC25A31 | 2 | SLC12A2 | 3 | SLC7A7 | 2 | SLC5A1 | 4 |
| SLC25A3 | 1 | SLC20A2 | 2 | SLC7A6 | 2 | SLC59A1 | 1 |
| SLC25A29 | 2 | SLC20A1 | 2 | SLC7A5 | 5 | SLC58A2 | 4 |
| SLC25A28 | 1 | SLC1A7 | 2 | SLC7A3 | 1 | SLC58A1 | 1 |
| SLC25A26 | 2 | SLC1A6 | 2 | SLC7A2 | 1 | SLC16A10 | 3 |
| SLC12A6 | 2 | SLC1A5 | 6 | SLC7A13 | 2 | SLC57A4 | 3 |
| SLC25A24 | 5 | SLC1A4 | 4 | SLC7A11 | 2 | SLC57A3 | 5 |
| SLC25A23 | 5 | SLC1A3 | 2 | SLC7A10 | 1 | SLC57A2 | 1 |
| SLC25A22 | 1 | SLC1A2 | 2 | SLC7A1 | 1 | SLC57A1 | 4 |
| SLC25A21 | 2 | SLC1A1 | 2 | SLC16A7 | 3 | SLC56A1 | 1 |
| SLC25A20 | 2 | SLC19A3 | 1 | SLC6A9 | 1 | SLC55A1 | 3 |
| SLC25A2 | 5 | SLC11A2 | 8 | SLC6A8 | 1 | SLC54A3 | 1 |
| SLC25A19 | 3 | SLC19A2 | 1 | SLC6A7 | 1 | SLC54A2 | 1 |
| SLC25A18 | 1 | SLC19A1 | 2 | SLC6A6 | 1 | SLC54A1 | 1 |
| SLC25A17 | 8 | SLC18A3 | 1 | SLC6A5 | 1 | SLC53A1 | 1 |
| SLC25A15 | 4 | SLC18A2 | 5 | SLC6A4 | 1 | SLC16A1 | 3 |
| SLC12A5 | 2 | SLC18A1 | 4 | SLC6A3 | 1 | SLC51B | 1 |
| SLC25A13 | 2 | SLC17A9 | 1 | SLC6A20 | 3 | SLC51A | 1 |
| SLC25A12 | 2 | SLC17A8 | 1 | SLC6A2 | 1 | SLC50A1 | 1 |
| SLC25A11 | 2 | SLC17A7 | 1 | SLC6A19 | 1 | SLC4A8 | 1 |
| SLC25A10 | 5 | SLC17A6 | 1 | SLC16A5 | 3 | SLC4A7 | 1 |
| SLC25A1 | 4 | SLC17A5 | 2 | SLC6A18 | 1 | SLC4A5 | 2 |
| SLC24A5 | 3 | SLC11A1 | 2 | SLC6A17 | 1 | SLC4A4 | 2 |
| SLC24A4 | 3 | SLCO2B1 | 2 | SLC6A15 | 2 | SLC4A3 | 1 |
| SLC24A3 | 3 | SLCO2A1 | 2 | SLC6A14 | 2 | SLC4A2 | 1 |
| SLC24A2 | 3 | SLC17A3 | 1 | SLC6A13 | 1 | SLC4A11 | 2 |
| SLC24A1 | 3 | SLCO1C1 | 3 | SLC6A12 | 2 | SLC15A4 | 3 |
| SLC12A4 | 2 | SLCO1B3 | 2 | SLC6A11 | 1 | SLC10A6 | 3 |
| SLC23A2 | 1 | SLCO1B1 | 2 | SLC6A1 | 1 | SLC4A10 | 1 |
| SLC23A1 | 1 | SLCO1A2 | 2 | SLC65A2 | 1 | SLC4A1 | 1 |
| SLC22B4 | 1 | SLC9C1 | 2 | SLC65A1 | 1 | SLC49A2 | 1 |
| SLC22B1 | 4 | SLC9B2 | 2 | SLC16A3 | 2 | SLC49A1 | 1 |
| SLC22A9 | 1 | SLC9A9 | 3 | SLC64A1 | 3 | SLC48A1 | 1 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| SLC22A8 | 1 | SLC9A8 | 3 | SLC63A2 | 6 | SLC47A2 | 6 |
| SLC22A7 | 1 | SLC9A7 | 5 | SLC62A1 | 1 | SLC47A1 | 8 |
| SLC22A6 | 1 | SLC9A6 | 3 | SLC61A1 | 1 | SLC46A3 | 1 |
| SLC22A5 | 3 | SLC17A1 | 3 | SLC60A2 | 2 | SLC46A1 | 3 |
| SLC22A4 | 3 | SLC9A5 | 4 | SLC5A9 | 3 | SLC45A4 | 3 |
| SLC12A3 | 2 | SLC9A4 | 4 | SLC5A8 | 1 | SLC15A3 | 3 |
| SLC22A32 | 5 | SLC9A3 | 4 | SLC5A7 | 1 | SLC45A3 | 3 |
| SLC22A3 | 1 | SLC9A2 | 4 | SLC5A6 | 3 | SLC45A2 | 3 |
| SLC22A20 | 1 | SLC9A1 | 4 | SLC5A5 | 5 | SLC26A3 | 3 |
| SLC22A2 | 1 | SLC8B1 | 3 | SLC16A2 | 4 | SLC26A2 | 3 |
| SLC44A1 | 1 | SLC33A1 | 1 | SLC45A1 | 3 | SLC12A9 | 1 |
| SLC43A2 | 2 | SLC13A3 | 5 | SLC44A2 | 1 | SLC26A11 | 4 |
| SLC43A1 | 2 | SLC32A1 | 1 | SLC34A2 | 1 | SLC26A1 | 3 |
| SLC42A3 | 2 | SLC31A2 | 2 | SLC34A1 | 1 | SLC25A9 | 1 |
| SLC42A2 | 4 | SLC31A1 | 2 | SLC10A1 | 1 | SLC25A8 | 1 |
| SLC42A1 | 2 | SLC2A9 | 3 | SLC13A5 | 3 | SLC25A7 | 1 |
| SLC15A2 | 4 | SLC2A8 | 3 | SLC38A3 | 4 | SLC25A6 | 2 |
| SLC41A2 | 6 | SLC2A7 | 2 | SLC38A2 | 9 | SLC25A5 | 2 |
| SLC41A1 | 8 | SLC2A6 | 1 | SLC38A1 | 6 | SLC25A42 | 6 |
| SLC40A1 | 1 | SLC2A5 | 1 | SLC36A4 | 2 | SLC25A41 | 1 |
| SLC39A8 | 3 | SLC2A4 | 2 | SLC36A2 | 4 | SLC25A4 | 2 |
| SLC39A7 | 2 | SLC2A3 | 4 | SLC36A1 | 4 | SLC12A7 | 2 |
| SLC39A6 | 1 | SLC13A2 | 3 | SLC35D2 | 1 | SLC10A2 | 1 |
| SLC39A5 | 1 | SLC2A2 | 5 | SLC35D1 | 2 | | |
| SLC39A4 | 1 | SLC2A13 | 1 | SLC35C2 | 1 | | |
| SLC39A3 | 1 | SLC2A12 | 1 | SLC35C1 | 1 | | |
| SLC39A2 | 1 | SLC2A11 | 2 | SLC13A4 | 3 | | |
| SLC15A1 | 4 | SLC2A10 | 2 | SLC35B4 | 2 | | |
| SLC39A14 | 4 | SLC2A1 | 4 | SLC35B3 | 1 | | |
| SLC39A13 | 1 | SLC27A6 | 2 | SLC35B2 | 1 | | |
| SLC39A12 | 1 | SLC27A5 | 2 | SLC35A3 | 1 | | |
| SLC39A10 | 1 | SLC27A4 | 2 | SLC35A2 | 2 | | |
| SLC39A1 | 1 | SLC27A3 | 2 | SLC35A1 | 1 | | |
| SLC38A9 | 7 | SLC13A1 | 3 | SLC34A3 | 1 | | |
| SLC38A8 | 4 | SLC27A2 | 2 | SLC26A7 | 5 | | |
| SLC38A7 | 5 | SLC27A1 | 2 | SLC26A6 | 5 | | |
| SLC38A5 | 4 | SLC26A9 | 4 | SLC26A5 | 4 | | |
| SLC38A4 | 6 | SLC26A8 | 3 | SLC26A4 | 3 | | |


*Figure 17 – Removing the non-breaking space by a String Replacer node.*

When I manually entered the data into an Excel spreadsheet, a non-breaking space was copied, which was not visible at the beginning. Only by the structure search, I noticed this, because an error appeared in the search. The reason for this was that the non-breaking space was recognized as a foreign symbol. In order to solve this problem and perform a successful structure search, I had to insert a String Replacer node (Figure 17) and configure it with a specific code so that this space would not appear as a symbol. For the non-breaking space, I used the code “\u00A0”, which I got with the help of the KNIME Forum (Figure 18).

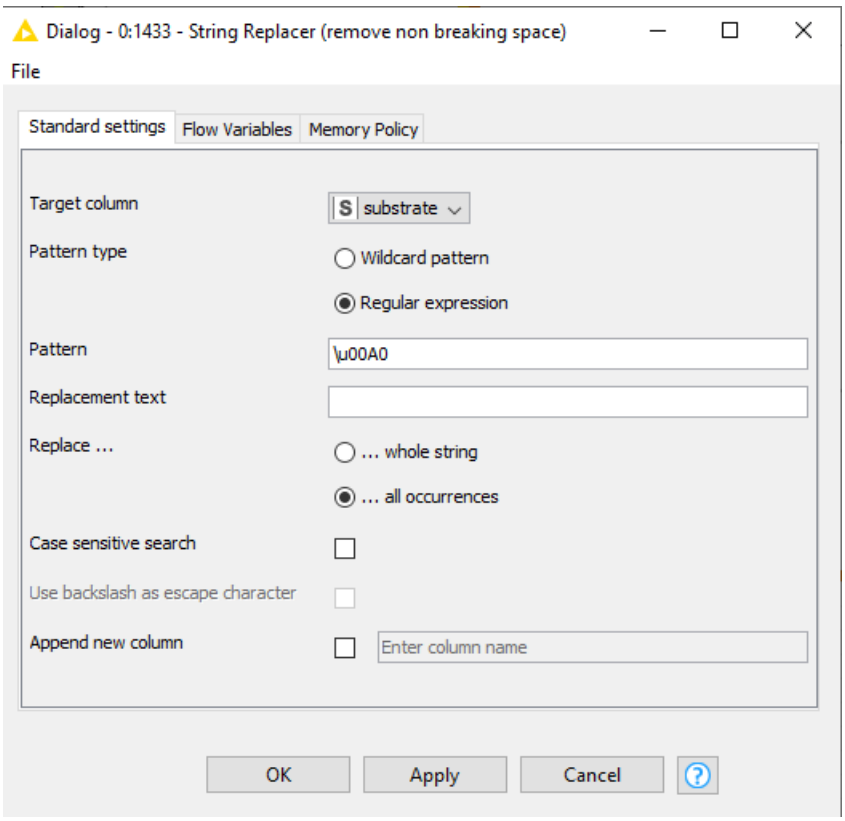



*Figure 18- Configuring the String Replacer node to remove the non-breaking space.*

### 2.2.2.1.4 **IUPHAR/BPS**

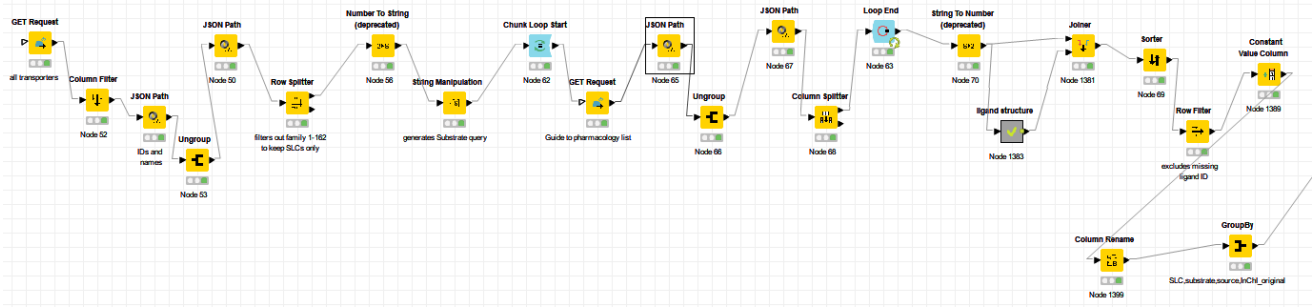



*Figure 19 - KNIME workflow collecting data from the IUPHAR/BPS database created by Daniela Digles, a member of the Pharmacoinformatic Research Group.*

For the diploma thesis, I used the IUPHAR query from the SLC_substrates_inhibitors workflow, by Daniela Digles. Through the URL *"http://www.guidetopharmacology.org/services/targets?type=TRANSPORTER*" and the Get Request node, access to the data was enabled. In the String Manipulation node, the expression was defined with the API, which was *"join("http://www.guidetopharmacology.org/services/targets/" , $targetId$ , "/substrates")",* to get the substrates to the SLCs. The String Manipulation node was followed by the Chunk loop node with the Get Request to make the query.

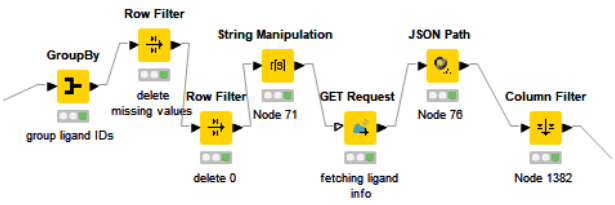



*Figure 20 – The workflow for the ligand structure search of the IUPHAR-BPS query.*

In order to retrieve the structural information of the substrate, another string manipulation node was used with the URL

*"join("http://www.guidetopharmacology.org/services/ligands/" , $ligand_ID$, "/structure")"* in combination with a Get Request node. The structural information was merged with the SLC-substrate search through a Joiner node, after organizing ascending targetID with the Sorter node and excluding rows with missing values with the Row Filter node.By this workflow I received 784 substrates for 204 solute carriers, which are evident in **Table 4**.

***Table 4 – Results of the GtoPbd database substrate collection***

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| SLC10A1 | 4 | SLC19A3 | 1 | SLC25A21 | 2 | SLC32A1 | 4 |
| SLC10A6 | 4 | SLC1A1 | 6 | SLC25A22 | 1 | SLC33A1 | 1 |
| SLC11A1 | 2 | SLC1A2 | 5 | SLC25A26 | 1 | SLC35A1 | 1 |
| SLC11A2 | 5 | SLC1A3 | 5 | SLC25A42 | 1 | SLC35A2 | 2 |
| SLC12A8 | 4 | SLC1A6 | 5 | SLC26A1 | 2 | SLC35A3 | 1 |
| SLC13A1 | 3 | SLC1A7 | 5 | SLC26A11 | 1 | SLC35B2 | 1 |
| SLC13A2 | 2 | SLC20A1 | 2 | SLC26A2 | 1 | SLC35B3 | 1 |
| SLC13A3 | 2 | SLC20A2 | 1 | SLC26A3 | 1 | SLC35B4 | 2 |
| SLC13A4 | 1 | SLC22A1 | 9 | SLC26A4 | 5 | SLC35C1 | 1 |
| SLC13A5 | 2 | SLC22A10 | 1 | SLC26A5 | 2 | SLC35D2 | 1 |
| SLC14A1 | 6 | SLC22A11 | 4 | SLC26A6 | 7 | SLC36A1 | 23 |
| SLC14A2 | 1 | SLC22A12 | 2 | SLC26A8 | 3 | SLC36A2 | 9 |
| SLC15A1 | 14 | SLC22A16 | 1 | SLC28A1 | 8 | SLC36A4 | 2 |
| SLC15A2 | 11 | SLC22A2 | 9 | SLC28A2 | 9 | SLC37A1 | 2 |
| SLC15A3 | 8 | SLC22A3 | 7 | SLC28A3 | 15 | SLC37A2 | 1 |
| SLC15A4 | 9 | SLC22A4 | 5 | SLC29A1 | 23 | SLC37A4 | 1 |
| SLC16A1 | 4 | SLC22A5 | 6 | SLC29A2 | 18 | SLC38A1 | 1 |
| SLC16A10 | 4 | SLC22A6 | 4 | SLC29A3 | 15 | SLC38A2 | 1 |
| SLC16A2 | 2 | SLC22A7 | 3 | SLC29A4 | 9 | SLC38A3 | 1 |
| SLC16A3 | 2 | SLC22A8 | 6 | SLC2A1 | 3 | SLC38A4 | 1 |
| SLC16A7 | 2 | SLC22B1 | 1 | SLC2A10 | 2 | SLC38A5 | 1 |
| SLC16A8 | 1 | SLC23A1 | 1 | SLC2A11 | 2 | SLC39A14 | 3 |
| SLC17A1 | 6 | SLC23A2 | 1 | SLC2A12 | 1 | SLC39A8 | 1 |
| SLC17A5 | 6 | SLC23A4 | 1 | SLC2A13 | 4 | SLC40A1 | 1 |
| SLC17A9 | 3 | SLC25A1 | 3 | SLC2A2 | 2 | SLC41A1 | 8 |
| SLC18A1 | 10 | SLC25A10 | 5 | SLC2A3 | 1 | SLC41A2 | 6 |
| SLC18A2 | 10 | SLC25A11 | 2 | SLC2A4 | 2 | SLC42A1 | 3 |
| SLC18A3 | 6 | SLC25A12 | 3 | SLC2A5 | 2 | SLC42A3 | 1 |
| SLC18B1 | 2 | SLC25A13 | 3 | SLC2A7 | 2 | SLC43A1 | 8 |
| SLC19A1 | 9 | SLC25A15 | 4 | SLC2A8 | 1 | SLC43A2 | 7 |
| SLC19A2 | 1 | SLC25A16 | 1 | SLC2A9 | 2 | SLC44A1 | 1 |
| SLC25A19 | 4 | SLC25A17 | 3 | SLC31A1 | 2 | SLC44A2 | 1 |

| | | | | | |
|---|---|---|---|---|---|
| SLC25A2 | 10 | SLC25A18 | 1 | SLC31A2 | 2 |
| SLC46A1 | 5 | SLC5A4 | 6 | SLCO1C1 | 4 |
| SLC47A1 | 8 | SLC5A5 | 5 | SLCO2A1 | 3 |
| SLC47A2 | 9 | SLC5A6 | 4 | SLCO2B1 | 12 |
| SLC49A1 | 1 | SLC5A7 | 2 | SLCO3A1 | 4 |
| SLC49A2 | 1 | SLC5A8 | 19 | SLCO4A1 | 5 |
| SLC4A1 | 2 | SLC5A9 | 3 | SLCO4C1 | 6 |
| SLC4A10 | 1 | SLC60A2 | 1 | SLC5A2 | 2 |
| SLC4A11 | 2 | SLC61A1 | 1 | SLC45A1 | 2 |
| SLC4A2 | 2 | SLC62A1 | 1 | SLCO1B3 | 21 |
| SLC4A3 | 2 | SLC63A2 | 3 | | |
| SLC4A4 | 1 | SLC64A1 | 2 | | |
| SLC4A5 | 1 | SLC65A1 | 1 | | |
| SLC4A7 | 1 | SLC65A2 | 1 | | |
| SLC4A8 | 2 | SLC6A1 | 3 | | |
| SLC52A1 | 1 | SLC6A11 | 4 | | |
| SLC52A2 | 1 | SLC6A12 | 2 | | |
| SLC52A3 | 1 | SLC6A13 | 4 | | |
| SLC53A1 | 1 | SLC6A14 | 5 | | |
| SLC54A1 | 1 | SLC6A2 | 6 | | |
| SLC54A2 | 1 | SLC6A20 | 1 | | |
| SLC54A3 | 1 | SLC6A3 | 6 | | |
| SLC55A1 | 1 | SLC6A4 | 3 | | |
| SLC57A1 | 2 | SLC6A5 | 1 | | |
| SLC57A2 | 1 | SLC6A6 | 3 | | |
| SLC57A3 | 1 | SLC6A7 | 1 | | |
| SLC57A4 | 1 | SLC6A8 | 1 | | |
| SLC58A1 | 1 | SLC6A9 | 2 | | |
| SLC58A2 | 1 | SLC7A1 | 4 | | |
| SLC59A1 | 1 | SLC7A2 | 4 | | |
| SLC5A1 | 3 | SLC7A3 | 3 | | |
| SLC5A10 | 2 | SLCO1A2 | 17 | | |
| SLC5A12 | 3 | SLCO1B1 | 18 | | |

### 2.2.2.1.5 Concatenation of the databases

To avoid later problems when merging the four databases, I used the Column Rename node to have a default naming of the columns and added through the Constant Value column node an extra column for the source. I performed these steps at the end of each data collection workflow to achieve a consistent representation of the data.

After getting the substrates to the SLCs from the four databases, I concatenated them with a Concatenate node to do later the structure search for all of them and have them in one workflow. Since I already had the structural information of the substrates from the Metrabase and IUPHAR/BPS databases, I named their column "InChI_original_database" to distinguish which substrates I already had the InChI code for and which I got through the text search.

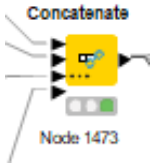



*Figure 21 – Concatenate node*

For example, as you can see in the Figure 22, which is an excerpt from the concatenated table, I already had the InChI code for amiloride from the Metrabase database, but not from the UCSF-FDA database.

| Row ID | SLC | substrate | source | InChI_original_database |
|---|---|---|---|---|
| Row54 | SLC22A2 | Amantadine | UCSF-FDA | ? |
| Row55 | SLC22A2 | Amiloride | UCSF-FDA | ? |
| Row385 | SLC22A1 | Amiloride | Metrabase | InChI=1S/C6H8ClN7O/c7-2-4(9)13-3(8)1(12-2)5(15)14-6(10)11/h(... |
| Row88_dup | SLC15A1 | Aminolevulinic acid | Metrabase | InChI=1S/C5H9NO3/c6-3-4(7)1-2-5(8)9/h1-3,6H2,(H,8,9) |

*Figure 22 – Expert of the concatenated table, which shows Amiloride from two different databases.*

## 2.2.2.2Find standardized molecular structure

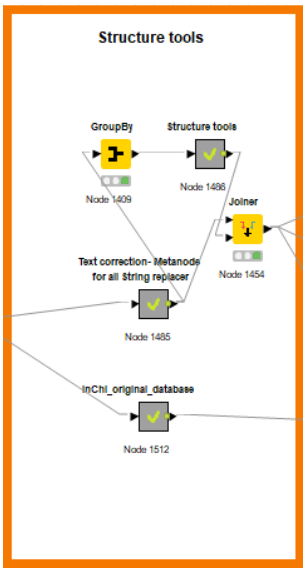



*Figure 23 – workflow overview of finding a standardized molecular structure for a given substrate name.*

In the next section, I will present the method of assigning standardized molecular structures to SLC substrate names.

After linking the data from all four databases, as described in the previous chapter, I had to make a text correction to remove some symbols. For this process, I used five String Replacer nodes that were configured differently to remove a symbol at each step. For example, the substrate name was “Fe<sup>2+</sup>” at the beginning and was corrected to “Fe2+” after the String Replacer node. The table (**Table 5**) illustrate this point clearly. After I finished the text correction, I connected a GroupBy node to the "Text correction metanode" to filter for the group columns "Substrate" and "InChI_original_database" and remove all duplicates.

***Table 5 – Configuration for the five Stringe Replacer nodes for the text correction.***

| Pattern | Replacement text | E.g. |
|---|---|---|
| <su?> | | “Fe<sup>2+</sup>” was changed to “Fe2+</sup>” |
| </su?> | | “Fe2+</sup>” was changed to “Fe2+” |
| β | beta | “β-alanine” was changed to “beta-alanine”. |
| α | alpha | “α-ketoglutaric acid” was changed to “alpha-ketoglutaric acid”. |
| γ | gamma | “γ-hydroxybutyric acid” was changed to “gamma-hydroxybutyric acid”. |

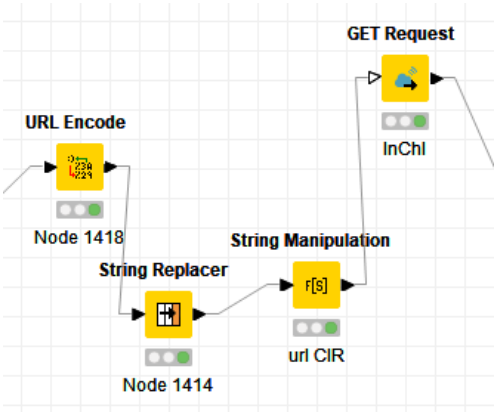



*Figure 24 – structure tools query*

Furthermore, I adapted for the structure search the CTS query, the CIR query, and the PubChem query from the “SLC_ substrates_structure_tools_Sarahf” workflow, by Sarah Foth[23], who also wrote her diploma thesis in our group and worked on SLC and its substrates. At the beginning of each query, I inserted a URL encode node, which unfortunately replaced each space with a "+" sign, resulting in the URL not being found in the web server, e.g., 17beta-Glucuronosyl estradiol was changed to 17beta-Glucuronosyl+estradiol. By using a String replacer node with a configuration, the problem could be solved (Figure 25).

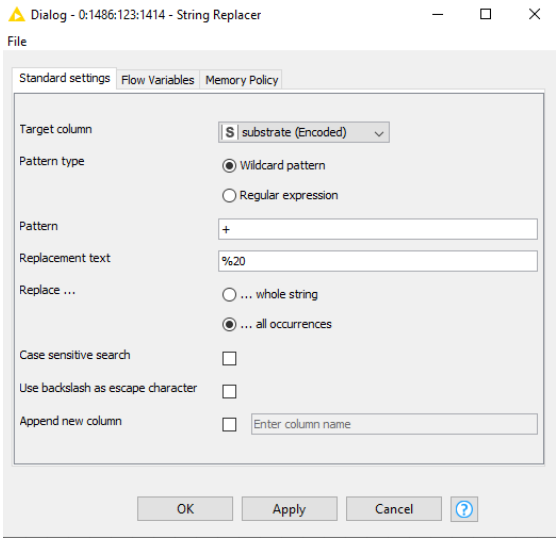



*Figure 25 – Removing the “+”-sign with a String Replacer node by using the code “%20”.*

The workflow for CTS query and CIR query were similar in structure, and each differed by the URL. For the CIR query the expression

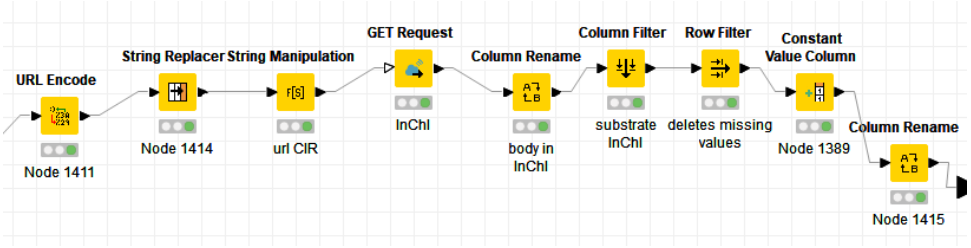



*Figure 26 – CIR query for the structure search.*

“join("https://cactus.nci.nih.gov/chemical/structure/" ,$substrate (Encoded)$, "/stdinchi")” and for the CTS query the expression

“join("https://cts.fiehnlab.ucdavis.edu/rest/convert/Chemical%20Name/InChI%20Code/", $substrate (Encoded)$)” was used. The workflow for the PubChem query (Figure 27) is very similar to the CIR and CTS query. Here the URL “join("https://pubchem.ncbi.nlm.nih.gov/rest/pug/compound/name/" ,$substrate (Encoded)$, "/property/IUPACName,InChI,InChIKey,IsomericSMILES,CanonicalSMILES/JSON")” was used.

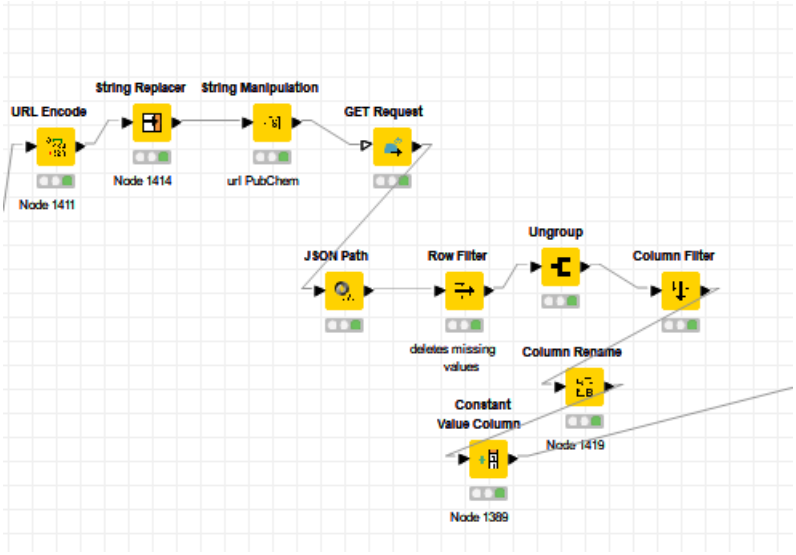



*Figure 27 – The workflow of the PubChem query for the structure search.*

After the structure search was performed, the data from all three retrievals was concatenated through a Concatenate node and then an RDKit From InChI node was appended, which was important for the next step. Then a

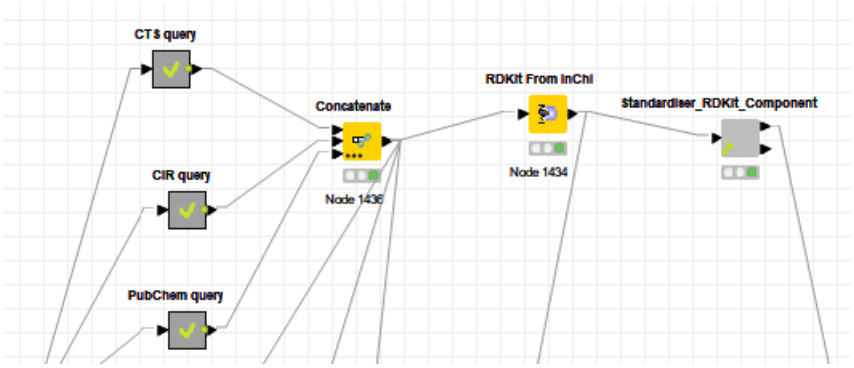



*Figure 28 – Knime Workflow_structure text search*

Standardiser_RDKit_Component metanode was added after the RDKit From InChI node. The Standardiser_RDKit_Component metanode was created by Jennifer Hemmerich[24], who was a member of the Pharmacoinformatics Research Group. The goal of this metanode is to standardize the structures. In this way, duplicate

structures from different databases can be detected, which look like different structures, although they are not. Therefore, the standardizer converts the RDKits to the same representation so that the duplicate structures are detected.

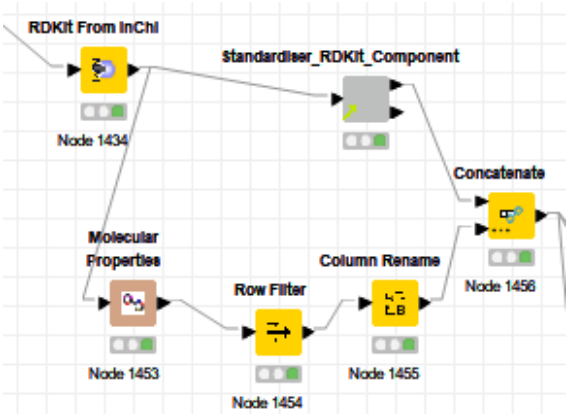



*Figure 29 – Problem solving for the removed salts and ions by inserting a Molecular properties node.*

Nevertheless, the structural standardizer removes the salts and ions, so after the RDKit From InChI node, another branch was appended with a Molecular Properties node (Figure 29,Figure 30) to filter all the removed salts and ions wit the No. of Heavy Atoms “1” and then connect them to the standardizer's results via a Concatenate node. After the structure search was performed, the Structure Tools meta node and the Text Correction meta node were connected by a Joiner node.

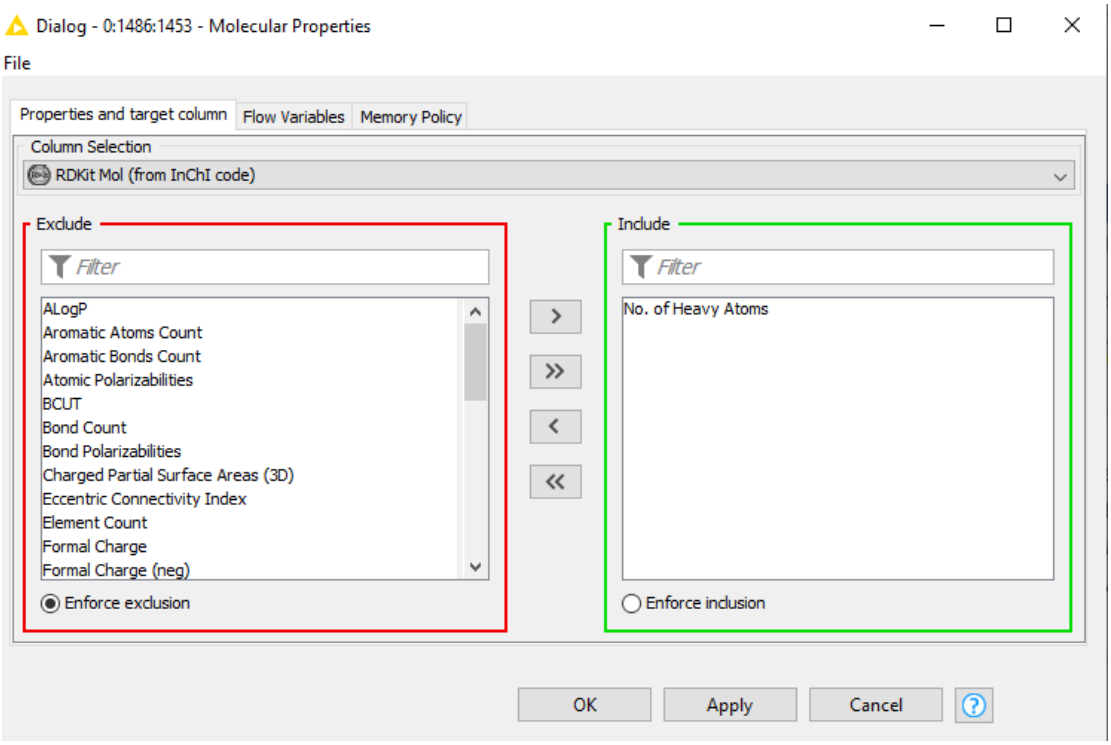



*Figure 30 – Configuring the Molecular Properties node to calculate the Number of the Heavy Atoms.*

In addition to the standardized InChI codes I got from the Standardiser Metanode, I also had the original InChI codes from the Metabolism and Transporter database and IUPHAR. Therefore, to distinguish between the original InChIs and the InChIs obtained from the text search, I created a meta node "InChI_original_database" (Figure 31), which I then connected to the text search outputs. To avoid later hidden duplicates, it was also necessary to use a standardiser_RDKit_Component here, as with the text search. The column for the InChI code from the databases Metrabase and IUPHAR were named InChI_original_database and from the text search were named InChICode_standardised.

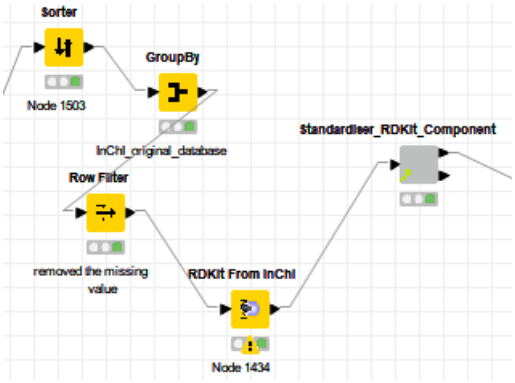
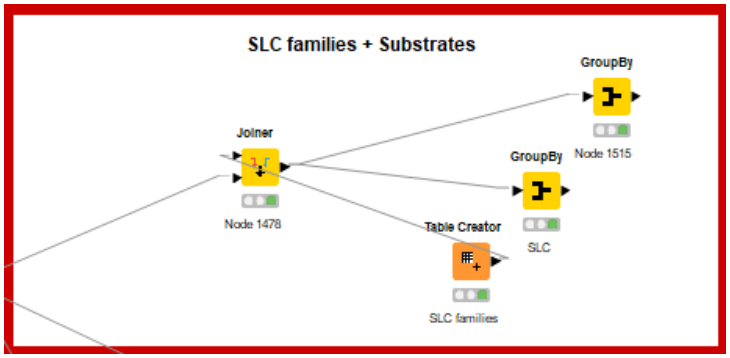



*Figure 31 – The workflow for the metanode InChI_original_database*

### 2.2.2.3 Merge SLC families and substrates




In the third part, I merged the SLC family information to the SLC names. Therefore, I used a Table creator node. In the appendix you can find the table that was used for this (**Table 15**).

*Figure 32 - Merging the substrates to the SLC names.*

### 2.2.2.4 Visualization

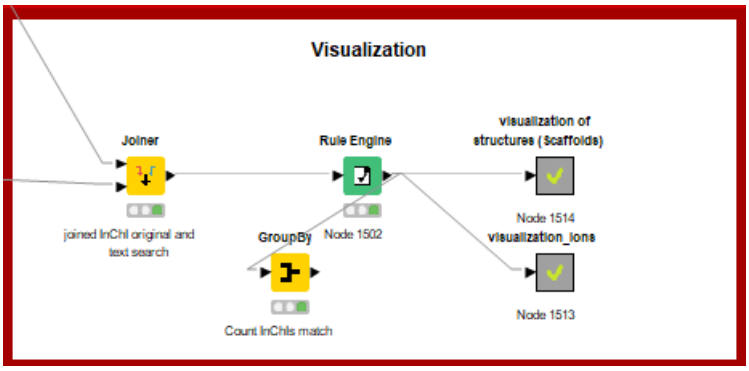



*Figure 33 – KNIME workflow_ visualization the data.*

In the last part of the workflow, I was concerned with visualizing the data and presenting them. The original InChIs and the InChIs from the text search were joined by a Joiner node, for which I used the "full outer join " as the join mode. A Rule Engine node was attached to the Joiner node and configured with rules, which are shown in the appendix (**Table 16**). The aim of using the Rule Engine node was to distinguish the InChIs from the text search and from the originals according to certain rules, in order to see which InChIs were found, and which were not.

As a result (**Fehler! Verweisquelle konnte nicht gefunden werden.**), 270 substrates had neither an InChI via structure search nor an original InChI, for 738 substrates there was no original InChI code, for 98 substrates no InChIs were found via text search, of 914 substrates the originals and the text search InChIs are not identical and for 455 substrates the originals and the text search InChIs are identical. Because of this, I worked with the original InChIs, which have also been standardized by the Standardiser Metanode. For the substrate, which I did not have the original InChIs, I used the standardized InChIs. I used the Column Merge node to merge the two columns (original and standardized) and to choose the one of them that is non-missing. For the configuration of the Column Merge node, I used "InChI_original_database" as the primary column and "InChICode_standardised" as the secondary column.

| Row ID | S InChIs ... | I Count*... |
|---|---|---|
| Row0 | EMPTY_both | 270 |
| Row1 | EMPTY_original | 738 |
| Row2 | EMPTY_text... | 98 |
| Row3 | NO | 914 |
| Row4 | YES | 455 |

*Figure 34 – Count InChIs match*

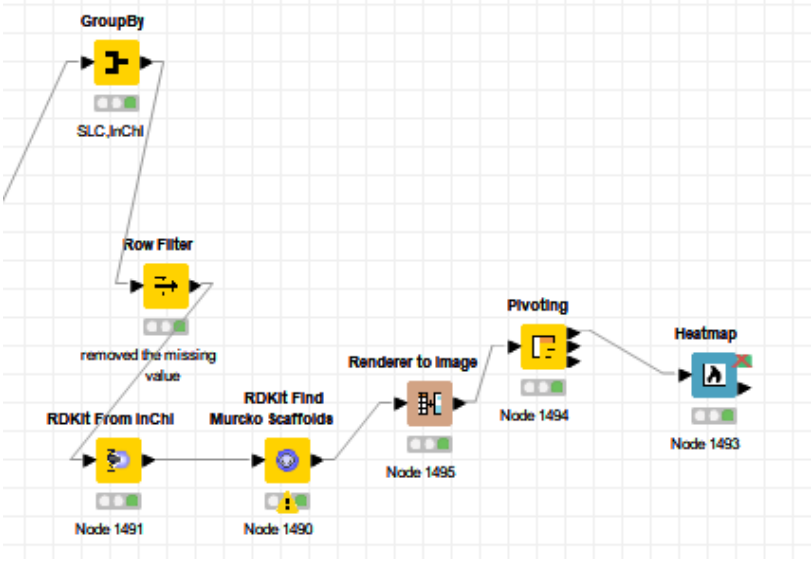



*Figure 35 – Knime Workflow_Visualisation of structures (Scaffolds)*

Once the structure search was complete and I had obtained the InChIs for the substrate, I wanted to visualize the data, for which I used this workflow (Figure 35). The GroupBy node grouped by SLCs and InChICode_standardised then a Row Filter node was added to remove the missing values. Through the RDKit From the InChI node, the InChIs were converted to RDKit, which was then necessary for the RDKiT Find Murcko Scaffolds node. The partitioning of the substrates by Murcko scaffolds was important to identify the similarity of the substrates transported by the SLCs, which was the goal of the diploma thesis. The Renderer to Image node appended a new column for SVG images (Figure 36). After these steps, a Pivoting node grouped by "RDKit (from InChI code) (Murcko)" and "Murcko Scaffold" and a Heatmap node were added.

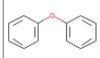
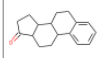
| Row ID | SLC | InChICode_standardised | Unique ... | RDKit Mol (... | RDKit Mol (... | Murcko Sc... |
|---|---|---|---|---|---|---|
| Row1 | SLC10A1 | InChI=1S/C15H11I4NO4/c16-8-4-7(5-9(17)13(8)21)24-14-10(18)1-6(2-11(14)19)3-12(20)15(22)23/h1-2,4-5,12,21H,3,20H2,(H,22,23) | T4 |  |  |  |
| Row2 | SLC10A1 | InChI=1S/C18H22O5S/c1-18-9-8-14-13-5-3-12(23-24(20,21)22)10-11(13)2-4-15(14)16(18)6-7-17(18)19/h3,5,10,14-16H,2,4,6-9H2,... | Estrone 3-s... |  |  |  |

*Figure 36 – Extract from the Renderer to Image node results.*

# 3 Results and Discussion

In summary, I obtained substrates for 317 different solute carriers from the four databases - Metabolism and Transporter Database, UCSF-FDA, BioParadigms and GtoPdb. Some solute carriers occur in more than one database, such as SLCO1B3, which was mentioned in three of the four databases, while others, such as SLC25A24 or SLC26A5, occur only in the BioParadigms database.

Of the four databases, most of the data came from the GtoPdb and BioParadigms databases. Although GtoPdb and BioParadigms covered all 317 SLCs, they did not contain the information about every substrate they transport. Thus, it would seem that the two databases would be sufficient for this work, but that is not the case. As this example shows. The BioParadigms database shows that SLC10A1 transports only bile acids. In contrast, the Metrabase database shows many other substrates that SLC10A1 transports, such as Cholate, Estrone-3-sulfate, Glycoursodeoxycholate, Pitavastatin, Rosuvastatin, Taurocholate, and Tauroursodeoxycholate. This suggests that not every database has dealt with the SLC in detail in the granularity.

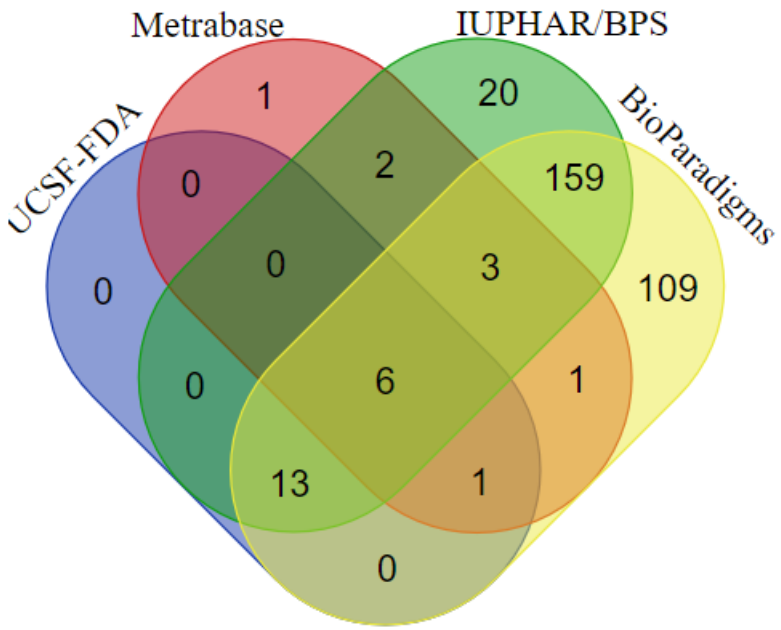



*Figure 37 - Data overlap of the databases.*

## *Table 6 – Data overlap of the databases*

| Names | total | SLC |
|---|---|---|
| BioParadigms IUPHAR/BPS Metrabase UCSF-FDA | 4 | SLC22A1 SLC15A1 SLCO1B1 SLCO1A2 |
| BioParadigms Metrabase UCSF-FDA | 1 | SLC10A2 |
| IUPHAR/BPS Metrabase UCSF-FDA | 2 | SLCO1B3 SLCO2B1 |
| BioParadigms IUPHAR/BPS UCSF-FDA | 13 | SLC22A8 SLC22A11 SLC22A4 SLC15A2 SLC10A1 SLC22A2 SLC22A12 SLC22A7 SLC22A6 SLC47A1 SLC47A2 SLC22A3 SLC22A5 |
| BioParadigms IUPHAR/BPS Metrabase | 2 | SLC16A1 SLC2A1 |
| BioParadigms Metrabase | 1 | SLC7A5 |
| IUPHAR/BPS Metrabase | 3 | SLCO4A1 SLCO2A1 SLCO3A1 |
| BioParadigms IUPHAR/BPS | 158 | SLC18A1 SLC6A2 SLC4A1 SLC22B1 SLC19A3 SLC35C1 SLC25A10 SLC2A8 SLC25A22 SLC13A5 SLC54A1 SLC39A8 SLC17A1 SLC25A18 SLC17A9 SLC1A3 SLC1A6 SLC23A1 SLC6A11 SLC4A5 SLC25A13 SLC57A4 SLC41A1 SLC11A2 SLC25A2 SLC2A5 SLC26A1 SLC26A5 SLC33A1 SLC6A4 SLC13A2 SLC38A5 SLC2A10 SLC55A1 SLC4A8 SLC35A3 SLC42A1 SLC38A1 SLC13A1 SLC7A1 SLC1A2 SLC2A7 SLC6A12 SLC5A12 SLC5A7 SLC31A1 SLC39A14 SLC25A26 SLC2A2 SLC57A3 SLC23A2 SLC36A1 SLC5A6 SLC38A2 SLC6A9 SLC44A2 SLC5A1 SLC2A12 SLC61A1 SLC36A4 SLC18A3 SLC5A5 SLC6A1 SLC25A12 SLC44A1 SLC1A7 SLC58A2 SLC2A13 SLC16A7 SLC5A2 SLC26A4 SLC16A8 SLC10A6 SLC6A13 SLC65A1 SLC26A8 SLC16A2 SLC25A1 SLC25A17 SLC6A5 SLC4A2 SLC1A1 SLC15A4 SLC6A7 SLC49A1 SLC35B2 SLC5A9 SLC18A2 SLC38A3 SLC26A3 SLC7A3 SLC2A3 SLC40A1 SLC59A1 SLC35D2 SLC13A3 SLC6A6 SLC43A1 SLC26A11 SLC54A3 SLC35A2 SLC35B3 SLC19A1 SLC6A20 SLC4A7 SLC25A11 SLC2A9 SLC20A1 SLC4A3 SLC31A2 SLC65A2 SLC53A1 SLC13A4 SLC64A1 SLC45A1 SLC42A3 SLC20A2 SLC58A1 SLC16A3 SLC63A2 SLC6A8 SLC4A11 SLC57A2 SLC38A4 SLC25A21 SLC4A4 SLC6A14 SLC54A2 SLC35A1 SLC5A10 SLC11A1 SLC43A2 SLC57A1 SLC22A16 SLC2A4 SLC17A5 SLC49A2 SLC41A2 SLC26A2 SLC36A2 SLC25A19 SLC5A8 SLC15A3 SLC62A1 SLC60A2 SLC16A10 SLC5A4 SLC6A3 SLC19A2 SLC25A15 SLC46A1 SLC26A6 SLC2A11 SLC35B4 SLC25A42 SLC4A10 SLC7A2 SLC32A1 |
| Metrabase | 1 | SLC51A-SLC51B |
| BioParadigms | 109 | SLC39A3 SLC5A11 SLC7A7 SLC7A6 SLC25A3 SLC27A1 SLC9A7 SLC12A7 SLC39A4 SLC27A6 SLC7A13 SLC34A3 SLC25A9 SLC25A37 SLC39A10 SLC22A32 SLC45A3 SLC51B SLC9A5 SLC12A3 SLC22B4 SLC39A7 SLC12A4 SLC38A9 SLC50A1 SLC25A28 SLC22A18 SLC38A8 SLC34A2 SLC46A3 SLC25A29 SLC8A1 SLC8A2 SLC9A8 SLC17A6 SLC45A2 SLC25A4 SLC27A3 SLC27A2 SLC9A2 SLC9A6 SLC25A32 SLC25A5 SLC24A3 SLC6A15 SLC1A4 SLC56A1 SLC12A9 SLC2A6 SLC16A5 SLC6A19 SLC8A3 SLC42A2 SLC8B1 SLC25A7 SLC7A9 SLC24A5 SLC48A1 SLC39A13 SLC22A13 SLC25A38 SLC12A6 SLC25A31 SLC24A2 SLC39A1 SLC7A11 SLC9A9 SLC9C1 SLC17A7 SLC27A4 SLC7A10 SLC5A3 SLC17A3 SLC35C2 SLC26A7 SLC25A41 SLC9A1 SLC25A23 SLC25A6 SLC24A4 SLC38A7 SLC25A20 SLC22A9 SLC17A8 SLC25A33 SLC1A5 SLC9B2 SLC7A8 SLC45A4 SLC26A9 SLC6A18 SLC27A5 SLC22A20 SLC25A8 SLC24A1 SLC9A3 SLC39A5 SLC9A4 SLC51A SLC35D1 SLC12A2 SLC6A17 SLC25A24 SLC39A2 SLC25A36 SLC12A5 SLC34A1 SLC39A12 SLC39A6 |
| IUPHAR/BPS | 21 | SLC37A4 SLC52A1 SLC28A3 SLC29A2 SLC29A4 SLC12A8 SLC28A2 SLC29A3 SLC23A4 SLC37A2 SLC25A16 SLC52A3 SLC14A1 SLC37A1 SLC22A10 SLC14A2 SLC28A1 SLC52A2 SLC18B1 SLC29A1 SLCO1C1 |

Figure 37 and **Table 6** show the overlap of the 4 databases covering the SLCs without substrate. It can be seen that IUPHAR/BPS and BioParadigms cover the SLCs. Metrabase shows one SLC group that is unique to its database, SLC51A-SLC51B, but appearances are deceiving as SLC51A and SLC51B appear separately in the other databases. The only 4 SLCs that appear in all four databases are SLC22A1, SLC15A1, SLCO1B1 and SLCO1A2. 109 are mentioned only in the BioParadigms database and 21 only in IUPHAR/BPS and do not occur in any other database of the other three.

Regarding, structure searches using the Chemical Identifier Resolver, Chemical Translation Service, and PubChem databases were successful. Out of 836 different substrates, it was not possible to get a name to structure search for 98 substrates, such as (R)-Xamoterol or Acidocillin. For 261 of 836 substrates, the InChI codes were found in all three queries.

The following two tables give you an insight into the SLCs and their substrates.

**Table 7** lists the 10 SLCs with the most substrates, e.g., SLC15A1, to which 263 different substrates bind. SLC15A1[25] belongs to the Solute Carrier Family 15, also known as oligopeptide transporters, and plays an important role in the uptake and digestion of dietary proteins and it enables the absorption of peptidomimetic drugs, such as ACE inhibitors like Enalapril and Zofenopril. The oligopeptide transporter is most abundantly expressed in small intestine, duodenum and gall bladder. The complete table with the listing of the substates to the corresponding SLCs can be found in the appendix (**Table 17**). **Table 8**, on the other hand, shows the top 10 SLCs with the different Murcko Scaffolds. Here there are differences in the order of the SLC. It can be observed that SLC22A1 is in first place with 82 different Murcko Scaffolds and not SLC15A1, which is placed third. SLC22A1[26], also called organic cation transporter, belongs to the solute carrier family 22 and is responsible for the excretion of endogenous organic cations as well as drugs. The transporter is a plasma integral membrane protein consisting of twelve transmembrane domains and is expressed in the liver.

Another significant aspect of the second table, is that two SLCs appear that are not included in the first table. These are SLC22A6, an Organic Anion Transporter, to

which 23 substrates bind, corresponding to 13 different Murcko scaffolds, and SLC29A2, an Equilibrative Nitrobenzylmercaptopurine Riboside-Insensitive Nucleoside Transporter, to which 23 substrates bind, corresponding to 13 different Murcko scaffolds. In the appendix you will find a more detailed list of the SLC and their Murcko Scaffolds (**Table 18**).

**Table 7 – Top 10 SLCs with the most substrates**

| | SLC | Unique substrates count |
|---|---|---|
| **1** | SLC15A1 | 263 |
| **2** | SLC22A1 | 183 |
| **3** | SLCO1B1 | 130 |
| **4** | SLCO1B3 | 96 |
| **5** | SLCO1A2 | 83 |
| **6** | SLCO2B1 | 65 |
| **7** | SLC10A2 | 58 |
| **8** | SLC22A2 | 30 |
| **9** | SLC22A8 | 30 |
| **10** | SLC36A1 | 25 |

***Table 8 – Top 10 SLCs with the most different Murcko Scaffolds***

| | SLC | Unique count Murcko Scaffolds |
|---|---|---|
| **1** | SLC22A1 | 82 |
| **2** | SLCO1B1 | 61 |
| **3** | SLC15A1 | 59 |
| **4** | SLCO1B3 | 49 |
| **5** | SLCO2B1 | 39 |
| **6** | SLCO1A2 | 35 |
| **7** | SLC22A8 | 22 |
| **8** | SLC22A2 | 16 |
| **9** | SLC22A6 | 13 |
| **10** | SLC29A1 | 13 |

The two most common Murcko scaffolds, such as the tetrahydropyran ring, which is a Murcko scaffold for D-glucose, is transported by 28 different SLCs, e.g., SLC15A1, SLC15A2, SLC16A1, and SLC16A10, making it one of the most common Murcko scaffolds, or the Murcko scaffold for ATP, which is transported by 9 different SLCs, such as SLC17A9, SLC25A17, or SLC25A23. On the other hand, the Murcko scaffold for testosterone is only transported by SLCO1B3 or the antidiabetic drug sitagliptin is also only transported by one SLC, like SLC22A1.

The figure (Figure 38) below shows a small section of the visualization of the data, from which, for example, the same substrate transported by different SLCs. The x-axis represents the Murcko Scaffolds, and the y-axis represents the different solute carriers. The blue squares indicate from which SLC the substrate is transported, where a black area appears, the substrate is not transported.


*Figure 38 – Extract of the data-visualization as a heatmap. The x-axis represents the Murcko Scaffolds, and the y-axis represents the different solute carriers. The blue squares indicate from which SLC the substrate is transported, where a black area appears, the substrate is not transported.*

Türkova et al.[27] showed in their work that they also worked with SLCs, except that they did not deal with each family but focused on only one, namely OATPs, organic anion-transporting polypeptides belonging to the SLCO (SLC21) superfamily of solute carriers. The focus was on these three SLCs, SLCO1B1 (OATP1B1), SLCO1B3 (OATP1B1), and SLCO2B1 (OATP2B1), which are expressed at the basolateral membrane of hepatocytes and are responsible for the uptake of endogenous substances such as bilirubin and bile salts into the liver cell. Therefore, they play an important role in the function and physiology of the liver. The goal of their work was to create a semi-automated KNIME workflow to collect data from various databases, as was the case in my work, with the difference that inhibitors were collected rather than substrates. Another difference from my work was that they were not working with the BioParadigms database, but with ChEMBL and DrugBank in addition to UCSF-FDA, Metrabase, and IUPHAR/BPS. It would be very interesting to add other databases such as ChEMBL or DrugBank to see if this would lead to a larger data collection or if there would be further overlap with the SLCs and their substrates.

# 4 Appendix

**Table 9 – Replacement list**

| Protein name | Gene name |
|---|---|
| ASBT | SLC10A2 |
| GLUT1 | SLC2A1 |
| LAT1 | SLC7A5 |
| MCT1 | SLC16A1 |
| OATP1A2 | SLCO1A2 |
| OATP1B1 | SLCO1B1 |
| OATP1B3 | SLCO1B3 |
| OATP2A1 | SLCO2A1 |
| OATP2B1 | SLCO2B1 |
| OATP3A1 | SLCO3A1 |
| OATP4A1 | SLCO4A1 |
| OCT1 | SLC22A1 |
| OSTA/OSTB | SLC51A-SLC51B |
| PEPT1 | SLC15A1 |

**Table 10 – Replacement list**

| Input (Lookup) | Output (Replacement) |
|---|---|
| 4-(4-dimethylamino)styryl-N-methylpyri-dinium (ASP+) | 4-(4-dimethylamino)styryl-N-methylpyridinium |
| Amantinin | Amanitin |
| Bisglucuronosyl bilirubin | Bilirubin diglucuronide |
| Monoglucuronosyl bilirubin | Bilirubin monoglucuronide |
| Bromsulphthalein (BSP) | Bromsulphthalein |

| | |
|---|---|
| Cholyl-glycylamido-fluorescein (CGamF) | Cholyl-glycylamido-fluorescein |
| 17beta-Glucuronosyl estradiol | Estradiol-17beta-glucuronide |

***Table 11 – Expression for the Rule-based Row Splitter for the Bioparadigms workflow***

| |
|---|
| $Substrates$ LIKE "*SLC*" => FALSE |
| $Substrates$ LIKE "unknown" => FALSE |
| $Substrates$ LIKE "not determined" => FALSE |
| $Substrates$ LIKE "inconclusive" => FALSE |
| //$Substrates$ LIKE "*falsely identified*" => FALSE |
| MISSING $Substrates$ => FALSE |
| TRUE => TRUE |

**Table 12 – Bioparadigms_SLC and substrate list (02.02.2021)**

| SLC name | Substrates |
|---|---|
| SLC1A1 | L-Glu, D/L-Asp |
| SLC1A2 | L-Glu, D/L-Asp |
| SLC1A3 | L-Glu, D/L-Asp |
| SLC1A4 | L-Ala, L-Ser, L-Cys, L-Thr |
| SLC1A5 | L-Ala, L-Ser, L-Cys, L-Thr, L-Gln, L-Asn |
| SLC1A6 | L-Glu, D/L-Asp |
| SLC1A7 | L-Glu, D/L-Asp |
| SLC2A1 | glucose, galactose, mannose, glucosamine |
| SLC2A2 | glucose, galactose, fructose, mannose, glucosamine |
| SLC2A3 | glucose, galactose, mannose, xylose |
| SLC2A3P1 | ? |
| SLC2A3P2 | ? |
| SLC2A3P4 | ? |
| SLC2A4 | glucose, glucosamine |

| | |
|---|---|
| SLC2A5 | fructose |
| SLC2A6 | glucose |
| SLC2A7 | glucose, fructose |
| SLC2A8 | glucose, fructose, galactose |
| SLC2A9 | urate (glucose, fructose) |
| SLC2A10 | glucose, galactose |
| SLC2A11 | glucose, fructose |
| SLC2A12 | glucose |
| SLC2A13 | myo-inositol |
| SLC2A14 | ? |
| SLC2AXP1 | ? |
| SLC3A1 | system b0,+, heterodimerizes with light subunit SLC7A9 |
| SLC3A2 | systems L, y+L, xc- and asc with light subunits SLC7A5-8 and SLC7A10-11 |
| SLC4A1 | chloride cicarbonate |
| SLC4A2 | chloride bicarbonate |
| SLC4A3 | chloride bicarbonate |
| SLC4A4 | sodium bicarbonate (and/or carbonate) |
| SLC4A5 | sodium bicarbonate (and/or carbonate) |
| SLC4A7 | chloride bicarbonate |
| SLC4A8 | sodium bicarbonate chloride |
| SLC4A9 | inconclusive |
| SLC4A10 | sodium bicarbonate chloride |
| SLC4A11 | sodium, borate |
| SLC5A1 | glucose and galactose (urea and water) |
| SLC5A2 | glucose |
| SLC5A3 | myoinositol (glucose) |
| SLC5A4 | Na+ (H+) |
| SLC5A5 | I- (ClO4-, SCN-, NO3-, Br-) |
| SLC5A6 | biotin, lipoate panthothenate, I- |
| SLC5A7 | choline |
| SLC5A8 | short chain fatty acids |
| SLC5A9 | mannose, fructose, glucose |

| | |
|---|---|
| SLC5A10 | mannose, fructose, glucose |
| SLC5A11 | myoinositol, chiro-inositol |
| SLC5A12 | short chain fatty acids |
| SLC6A1 | GABA |
| SLC6A2 | norepinephrine |
| SLC6A3 | dopamine |
| SLC6A4 | serotonin |
| SLC6A5 | glycine |
| SLC6A6 | taurine |
| SLC6A7 | proline |
| SLC6A8 | creatine |
| SLC6A9 | glycine |
| SLC6A10 | ? |
| SLC6A10P | ? |
| SLC6A11 | GABA |
| SLC6A12 | betaine, GABA |
| SLC6A13 | GABA |
| SLC6A14 | neutral, cationic amino acids |
| SLC6A15 | large, neutral amino acids |
| SLC6A16 | unknown |
| SLC6A17 | neutral amino acids |
| SLC6A18 | neutral amino acids |
| SLC6A19 | neutral amino acids |
| SLC6A20 | proline, pipecolate, sarcosine |
| SLC6A21 | ? |
| SLC7A1 | cationic L-amino acids |
| SLC7A2 | cationic L-amino acids |
| SLC7A3 | cationic L-amino acids |
| SLC7A4 | ? |
| SLC7A5 | large neutral L-amino acids, T3, T4, L-DOPA, BCH |
| SLC7A5P1 | ? |
| SLC7A6 | cationic amino acids (Na+ indep.), large neutral L-amino acids (Na+ dep.) |

| | |
|---|---|
| SLC7A7 | cationic amino acids (Na+ indep.), large neutral L-amino acids (Na+ dep.) |
| SLC7A8 | neutral L-amino acids, T3, T4, BCH |
| SLC7A9 | cationic amino acids, large neutral amino acids |
| SLC7A10 | small neutral amino acids |
| SLC7A11 | cystine (anionic form), L-glutamate |
| SLC7A13 | L-aspartate, L-glutamate |
| SLC7A14 | ? |
| SLC7A15P | ? |
| SLC8A1 | Na+, Ca2+ |
| SLC8A2 | Na+, Ca2+ |
| SLC8A3 | Na+, Ca2+ |
| SLC8B1 | Na+, Li+, Ca2+ |
| SLC9A1 | Na+, Li+, H+, NH4+ |
| SLC9A2 | Na+, Li+, H+, NH4+ |
| SLC9A3 | Na+, Li+, H+, NH4+ |
| SLC9A3P1 | ? |
| SLC9A3P2 | ? |
| SLC9A3P3 | ? |
| SLC9A3P4 | ? |
| SLC9A4 | Na+, Li+ (?), H+, NH4+ |
| SLC9A5 | Na+, Li+, H+, NH4+ (?) |
| SLC9A6 | Na+, K+, H+ |
| SLC9A7 | Na+, K+, Li+, H+, NH4+ (?) |
| SLC9A7P1 | ? |
| SLC9A8 | Na+, K+, H+ |
| SLC9A9 | Na+, K+, H+ |
| SLC9B1 | ? |
| SLC9B2 | Na+, Li+ |
| SLC9C1 | Na+, H+ |
| SLC9C2 | ? |
| SLC10A1 | bile acids |
| SLC10A2 | bile acids |

| | |
|---|---|
| SLC10A3 | ? |
| SLC10A4 | ? |
| SLC10A5 | ? |
| SLC10A6 | estrone-3-sulfate, dehydroepiandrosterone sulfate, pregnenolone sulfate |
| SLC10A7 | ? |
| SLC11A1 | Mn2+, Fe2+, other divalent metal ions |
| SLC11A2 | Fe2+, Cd2+, Co2+, Cu1+, Mn2+, Ni2+, Pb2+, Zn2+ |
| SLC12A2 | Na+, K+, Cl- |
| SLC12A3 | Na+, Cl- |
| SLC12A4 | K+, Cl- |
| SLC12A5 | K+, Cl- |
| SLC12A6 | K+, Cl- |
| SLC12A7 | K+, Cl- |
| SLC12A8 | unknown |
| SLC12A9 | polyamines? |
| SLC13A1 | sulfate, selenate, thiosulfate |
| SLC13A2 | succinate, citrate, ?-ketoglutarate |
| SLC13A3 | succinate, citrate, ?-ketoglutarate, NALA, glutarate and its derivatives |
| SLC13A4 | sulfate, oxyanions selenium and chromium |
| SLC13A5 | citrate, succinate, pyruvate |
| SLC14A1_UTB1 | ? |
| SLC14A1_UTB2 | ? |
| SLC14A2_UTA1 | ? |
| SLC14A2_UTA2 | ? |
| SLC14A2_UTA3 | ? |

| | |
|---|---|
| SLC14A2_UT-A4 | ? |
| SLC14A2_UT-A5 | ? |
| SLC14A2_UT-A6 | ? |
| SLC15A1 | di- and tri-peptides, protons, beta-lactam antibiotics |
| SLC15A2 | di- and tri-peptides, protons, beta-lactam antibiotics |
| SLC15A3 | di- and tri-peptides, protons, histidine |
| SLC15A4 | di- and tri-peptides, protons, histidine |
| SLC16A1 | lactate, pyruvate, ketone bodies |
| SLC16A2 | T2, rT3, T3, T4 |
| SLC16A3 | lactate, ketone bodies |
| SLC16A4 | ? |
| SLC16A5 | bumetanide, probenecid, nateglinide ? |
| SLC16A6 | ? |
| SLC16A7 | pyruvate, lactate, ketone bodies |
| SLC16A8 | lactate |
| SLC16A9 | ? |
| SLC16A10 | aromatic amino acids, T3, T4 |
| SLC16A11 | ? |
| SLC16A12 | ? |
| SLC16A13 | ? |
| SLC16A14 | ? |
| SLC17A1 | organic anions, phosphate, chloride |
| SLC17A2 | unknown |
| SLC17A3 | organic anions |
| SLC17A4 | unknown |
| SLC17A5 | sialic acid, other acidic sugars |
| SLC17A6 | glutamate |
| SLC17A7 | glutamate |
| SLC17A8 | glutamate |
| SLC17A9 | purine nucleotides |

| | |
|---|---|
| SLC18A1 | 5-HT, DA, NE, epinephrine |
| SLC18A2 | 5-HT, DA, NE, epinephrine, histamine |
| SLC18A3 | acetylcholine |
| SLC18B1 | unknown |
| SLC19A1 | reduced folates, antifolates |
| SLC19A2 | thiamine |
| SLC19A3 | thiamine |
| SLC20A1 | inorganic phosphate (monovalent) |
| SLC20A2 | inorganic phosphate (monovalent) |
| SLCO2A1 | prostaglandins (C/lactate) |
| SLCO4A1 | ? |
| SLCO5A1 | ? |
| SLCO3A1 | ? |
| SLCO6A1 | ? |
| SLCO1A2 | bile salts, organic anions and cations |
| SLCO2B1 | E-3-S, DHEAS |
| SLCO1B1 | bile salts, organic anions |
| SLCO1B3 | bile salts, organic anions |
| SLCO4C1 | ? |
| SLCO1C1 | T4, T3, rT3 |
| SLC22A1 | organic cations |
| SLC22A2 | organic cations |
| SLC22A3 | organic cations |
| SLC22A4 | ergothioneine, zwitterions, organic cations |
| SLC22A5 | zwitterions (L-carnitine), organic cations |
| SLC22A6 | organic anions |
| SLC22A7 | organic anions |
| SLC22A8 | organic anions |
| SLC22A9 | organic anions |
| SLC22A10 | not determined |
| SLC22A11 | organic anions |
| SLC22A12 | urate, organic anions |
| SLC22A13 | urate, organic anions |

| | |
|---|---|
| SLC22A14 | not determined |
| SLC22A15 | not determined |
| SLC22A16 | L-carnitine, noncharged compounds |
| SLC22A17 | not determined |
| SLC22A18 | probably organic anions |
| SLC22A20 | probably organic anions |
| SLC22A23 | ? |
| SLC22A24 | ? |
| SLC22A25 | ? |
| SLC22A31 | ? |
| SLC22A32 | p-aminohippuric acid (PAH), indomethacin, diclofenac, mefenamic acid, etodolac [PMID: 18638446] |
| SLC22B1 | Galactose [PMID: 25326386]; also binds levetiracetam [PMID: 15210974], selectracetam [PMID: 18183537], brivaracetam [PMID: 17785672]. |
| SLC22B2 | ? |
| SLC22B3 | ? |
| SLC22B4 | Nicotinate [PMID: 21953179]; binds nucleotides in the TM region [PMID: 19390693] |
| SLC22B5 | ? |
| SLC23A1 | L-ascorbic acid |
| SLC23A2 | L-ascorbic acid |
| SLC23A3 | ? |
| SLC23A4P | ? |
| SLC24A1 | Na+, Ca2+, K+ |
| SLC24A2 | Na+, Ca2+, K+ |
| SLC24A3 | Na+, Ca2+, K+ |
| SLC24A4 | Na+, Ca2+, K+ |
| SLC24A5 | Na+, Ca2+, K+ |
| SLC25A1 | citrate, isocitrate, malate, PEP |
| SLC25A2 | ornithine, citrulline, lysine, arginine, histidine |
| SLC25A3 | phosphate |
| SLC25A4 | ADP, ATP |

| | |
|---|---|
| SLC25A5 | ADP, ATP |
| SLC25A5P1 | ? |
| SLC25A6 | ADP, ATP |
| SLC25A6P1 | ? |
| SLC25A7 | H+ |
| SLC25A8 | H+ |
| SLC25A9 | H+ |
| SLC25A10 | malate, phosphate, succinate, sulphate, thiosulphate |
| SLC25A11 | 2-oxoglutarate, malate |
| SLC25A12 | aspartate, glutamate |
| SLC25A13 | aspartate, glutamate |
| SLC25A14 | ? |
| SLC25A15 | ornithine, citrulline, lysine, arginine |
| SLC25A15P1 | ? |
| SLC25A16 | ? |
| SLC25A17 | CoA, FAD, NAD+, AMP, ADP, PAP, dPCoA, FMN |
| SLC25A18 | glutamate |
| SLC25A19 | thiamine pyrophosphate, thiamine monophosphate, (deoxy)nucleotides |
| SLC25A20 | carnitine, acylcarnitine |
| SLC25A20P1 | ? |
| SLC25A21 | oxoadipate, oxoglutarate |
| SLC25A22 | glutamate |
| SLC25A23 | ATP-Mg2+, ATP, ADP, AMP, Pi |
| SLC25A24 | ATP-Mg2+, ATP, ADP, AMP, Pi |
| SLC25A25 | ? |
| SLC25A26 | S-adenosyl-methionine, S-adenosyl-homocysteine |
| SLC25A27 | ? |
| SLC25A28 | Fe2+ |
| SLC25A29 | ornithine, acylcarnitine |
| SLC25A30 | ? |
| SLC25A31 | ADP, ATP |
| SLC25A32 | folate |

| | |
|---|---|
| SLC25A33 | UTP |
| SLC25A34 | ? |
| SLC25A35 | ? |
| SLC25A36 | pyrimidine nucleotides |
| SLC25A37 | Fe2+ |
| SLC25A38 | glycine ? |
| SLC25A39 | ? |
| SLC25A40 | ? |
| SLC25A41 | ATP-Mg / Pi |
| SLC25A42 | CoA, ADP, ATP, adenosine 3',5'-diphosphate, dPCoA |
| SLC25A43 | ? |
| SLC25A44 | ? |
| SLC25A45 | ? |
| SLC25A46 | ? |
| SLC25A47 | ? |
| SLC25A48 | ? |
| SLC25A49 | ? |
| SLC25A50 | ? |
| SLC25A51 | ? |
| SLC25A51P1 | ? |
| SLC25A51P2 | ? |
| SLC25A51P3 | ? |
| SLC25A52 | ? |
| SLC25A53 | ? |
| SLC26A1 | SO42-, oxalate, glyoxylate |
| SLC26A2 | SO42-, oxalate, Cl- |
| SLC26A3 | Cl-, HCO3-, oxalate |
| SLC26A4 | I-, Cl-, HCO3- |
| SLC26A5 | Cl-, formate, oxalate, SO42- |
| SLC26A6 | Cl-, HCO3-, oxalate, OH-, formate |
| SLC26A7 | Cl-, HCO3-, OH-, SO42-, Ch: Cl- |
| SLC26A8 | Cl-, HCO3-, OH- |
| SLC26A9 | Cl-, HCO3, Ch: Cl-, HCO3- |

| | |
|---|---|
| SLC26A10 | ? |
| SLC26A11 | Cl-, HCO3-, SO42-, oxalate (?), Ch: Cl- |
| SLC27A1 | LCFA, VLCFA |
| SLC27A2 | LCFA, VLCFA |
| SLC27A3 | LCFA, VLCFA |
| SLC27A4 | LCFA, VLCFA |
| SLC27A5 | LCFA, bile acids |
| SLC27A6 | LCFA, VLCFA |
| SLC28A1 | ? |
| SLC28A2 | ? |
| SLC28A3 | ? |
| SLC29A1 | ? |
| SLC29A2 | ? |
| SLC29A3 | ? |
| SLC29A4 | ? |
| SLC30A1 | ? |
| SLC30A2 | ? |
| SLC30A3 | ? |
| SLC30A4 | ? |
| SLC30A5 | ? |
| SLC30A6 | ? |
| SLC30A7 | ? |
| SLC30A8 | ? |
| SLC30A9 | ? |
| SLC30A10 | ? |
| SLC31A1 | copper (I), cisplatin |
| SLC31A1P1 | ? |
| SLC31A2 | copper, cisplatin |
| SLC32A1 | GABA / glycine |
| SLC33A1 | acetyl-CoA |
| SLC33A2 | ? |
| SLC34A1 | inorganic phosphate (divalent) |
| SLC34A2 | inorganic phosphate (divalent) |

| | |
|---|---|
| SLC34A3 | inorganic phosphate (divalent) |
| SLC35A1 | CMP-sialic acid |
| SLC35A2 | UDP-galactose, UDP-N-acetylgalactosamine |
| SLC35A3 | UDP-N-acetylglucosamine |
| SLC35A4 | ? |
| SLC35A5 | ? |
| SLC35B1 | ? |
| SLC35B2 | PAPS |
| SLC35B3 | PAPS |
| SLC35B4 | UDP-xylose, UDP-N-acetylglucosamine |
| SLC35C1 | GDP-fucose |
| SLC35C2 | GDP-fucose (?) |
| SLC35D1 | UDP-glucuronic acid, UDP-N-acetylgalactosamine |
| SLC35D2 | UDP-N-acetylglucosamine |
| SLC35D3 | ? |
| SLC35D4 | ? |
| SLC35E1 | ? |
| SLC35E2 | ? |
| SLC35E3 | ? |
| SLC35E4 | ? |
| SLC35F1 | ? |
| SLC35F2 | ? |
| SLC35F3 | ? |
| SLC35F4 | ? |
| SLC35F5 | ? |
| SLC35F6 | ? |
| SLC35G1 | ? |
| SLC35G2 | ? |
| SLC35G3 | ? |
| SLC35G4 | ? |
| SLC35G5 | ? |
| SLC35G6 | ? |
| SLC36A1 | GABA, P, G, beta-alanine |

| | |
|---|---|
| SLC36A2 | P, G, A, hydroxyproline |
| SLC36A3 | ? |
| SLC36A4 | P, tryptophan |
| SLC37A1 | ? |
| SLC37A2 | ? |
| SLC37A3 | ? |
| SLC37A4 | ? |
| SLC38A1 | Q, A, N, C, H, S |
| SLC38A2 | A, N, C, Q, G, H, M, P, S |
| SLC38A3 | Q, H, A, N |
| SLC38A4 | A, N, C, G, S, T |
| SLC38A5 | Q, N, H, S |
| SLC38A6 | ? |
| SLC38A7 | Q, H, S ,A, N |
| SLC38A8 | Q, H, A, N [PMID: 25451601] |
| SLC38A9 | R, Q, H, P, K, E, L |
| SLC38A10 | ? |
| SLC38A11 | ? |
| SLC39A1 | Zn |
| SLC39A2 | Zn |
| SLC39A3 | Zn, not specific |
| SLC39A4 | Zn |
| SLC39A5 | Zn |
| SLC39A6 | Zn |
| SLC39A7 | Zn, Mn |
| SLC39A8 | Zn, Cd, Mn |
| SLC39A9 | ? |
| SLC39A10 | Zn |
| SLC39A11 | ? |
| SLC39A12 | Zn |
| SLC39A13 | Zn |
| SLC39A14 | Zn, Fe, Mn, Cd |
| SLC40A1 | ferrous iron |

| | |
|---|---|
| SLC41A1 | Mg2+ (Sr2+, Zn2+, Cu2+, Fe2+, Co2+, Ba2+, Cd2+) |
| SLC41A2 | Mg2+ (Ba2+, Ni2+, Co2+, Fe2+, Mn2+) |
| SLC41A3 | ? |
| SLC42A1 | NH4+, NH3 |
| SLC42A2 | NH4+, NH3, methyl amine, methyl ammonium |
| SLC42A3 | NH4+, NH3 |
| SLC43A1 | L-BCAAs, amino alcohols |
| SLC43A2 | L-BCAAs, amino alcohols |
| SLC43A3 | ? |
| SLC44A1 | choline |
| SLC44A2 | choline |
| SLC44A3 | ? |
| SLC44A4 | ? |
| SLC44A5 | ? |
| SLC45A1 | glucose, galactose, sucrose |
| SLC45A2 | sucrose, glucose, fructose [PMID: 25164149] |
| SLC45A3 | sucrose, glucose, fructose [PMID: 25164149] |
| SLC45A4 | sucrose, glucose, fructose [PMID: 25164149] |
| SLC46A1 | Reduced folates, folic acid, antifolates |
| SLC46A2 | unknown |
| SLC46A3 | Uncertain; lysosomal export of maytansine conjugates |
| SLC47A1 | tetraethylammonium (TEA), 1-methyl-4-phenylpyridinium (MPP), cimetidine, metformin, guanidine, procainamide, cephalexin, cephradine |
| SLC47A2 | TEA, MPP, cimetidine, metformin, guanidine, procainamide |
| SLC48A1 | heme |
| SLC49A1 | heme |
| SLC49A2 | heme |
| SLC49A3 | unknown |
| SLC49A4 | unknown |
| SLC50A1 | glucose |
| SLC51A | bile acids |
| SLC51B | steroids |

| | |
|---|---|
| SLC52A1 | ? |
| SLC52A2 | ? |
| SLC52A3 | ? |
| SLC53A1 | Phosphate [PMID: 23791524] |
| SLC54A1 | Pyruvate [PMID: 22628554] [PMID: 22628558] |
| SLC54A2 | Pyruvate [PMID: 22628558] |
| SLC54A3 | Pyruvate [PMID: 27317664] |
| SLC55A1 | Ca2+, K+, H+ (K+: [PMID: 20197279] [PMID: 15138253], [PMID: 17925330], yeast protein: [PMID: 17541427]), arguments against Ca2+ transport [PMID: 20197279]. |
| SLC55A2 | ? |
| SLC55A3 | ? |
| SLC56A1 | Probably falsely identified as a tricarboxylate carrier in [PMID: 8132491] as discussed in [PMID: 11274051]. Likely transports pyridoxin or another heme precursor or ALAS2 cofactor [PMID: 11274051] [PMID: 12670026]. |
| SLC56A2 | ? |
| SLC56A3 | ? |
| SLC56A4 | ? |
| SLC56A5 | ? |
| SLC57A1 | Mg2+ [PMID: 17166836], also Sr2+, Fe2+ and Co2+ to a lesser extent [PMID: 18667602] |
| SLC57A2 | Mg2+ [PMID: 18667602] |
| SLC57A3 | Mg2+, Sr2+, Ba2+, Fe2+, Cu2+ [PMID: 18667602] |
| SLC57A4 | Mg2+, Sr2+, Ba2+ [PMID: 18667602] |
| SLC57A5 | ? |
| SLC57A6 | ? |
| SLC58A1 | Mg2+ [PMID: 15804357] |
| SLC58A2 | Mg2+, Fe2+, Cu2+, Mn2+ [PMID: 18667602] [PMID: 19940067] |
| SLC59A1 | LPC (lysophosphatidylcholine) form of DHA (docosahexaenoic acid) [PMID: 24828044] |
| SLC59A2 | ? |
| SLC60A1 | ? |

| | |
|---|---|
| SLC60A2 | α-Me-glucose, D-glucose [PMID: 12590146], inhibited by phlorizin [PMID: 12590146], also inhibited by 2-DG, slightly by fructose and phloretin [PMID: 12590146]; urea [PMID: 26423860] |
| SLC61A1 | Molybdate [PMID: 21464289] |
| SLC62A1 | Pyrophosphate [PMID: 10894769] |
| SLC63A1 | ? |
| SLC63A2 | Phosphorylated sphingolipids [PMID: 19074308]; phosphorylated Fingolimod (FTY720-P) [PMID: 21084291]; sphingosine-1-phosphate (S1P), dihydrosphingosine-1-phosphate (DH-S1P), phyto-S1P, C17-S1P [PMID: 21084291] |
| SLC63A3 | ? |
| SLC64A1 | Ca2+, H+ [PMID: 23569283]; or Mn2+ [PMID: 27008884] [PMID: 28270545] |
| SLC65A1 | Cholesterol [PMID: 17989073] [PMID: 17989072] [PMID: 27410046] |
| SLC65A2 | Cholesterol [PMID: 14976318] |

***Table 13 – Replacement list***

| Input (Lookup) | Output (Replacement) |
|---|---|
| chloride cicarbonate | chloride bicarbonate |
| urate (glucose, fructose) | urate,glucose,fructose |
| sodium bicarbonate (and/or carbonate) | sodium bicarbonate, carbonate |
| glucose and galactose (urea and water) | glucose, galactose, urea, water |
| myoinositol (glucose) | myoinositol, glucose |
| Na+ (H+) | Na+, H+ |
| I- (ClO4-, SCN-, NO3-, Br-) | I-, ClO4-, SCN-, NO3-, Br- |
| cystine (anionic form), L-glutamate | cystine anionic form, L-glutamate |
| Na+, Li+ (?), H+, NH4+ | Na+, Li+, H+, NH4+ |
| Na+, Li+, H+, NH4+ (?) | Na+, Li+, H+, NH4+ |

| | |
|---|---|
| Na+, K+, Li+, H+, NH4+ (?) | Na+, K+, Li+, H+, NH4+ |
| polyamines? | polyamines |
| Mn2+, Fe2+, other divalent metal ions | Mn2+, Fe2+ |
| succinate, citrate, ?-ketoglutarate | succinate, citrate, ketoglutarate |
| succinate, citrate, ?-ketoglutarate, NALA, glutarate and its derivatives | succinate, citrate, ketoglutarate, NALA, glutarate |
| bumetanide, probenecid, nateglinide ? | bumetanide, probenecid, nateglinide |
| inorganic phosphate (monovalent) | inorganic phosphate, monovalent |
| prostaglandins (C/lactate) | prostaglandins, lactate |
| zwitterions (L-carnitine), organic cations | zwitterions,L-carnitine, organic cations |
| p-aminohippuric acid (PAH), indomethacin, diclofenac, mefenamic acid, etodolac [PMID: 18638446] | p-aminohippuric acid, indomethacin, diclofenac, mefenamic acid, etodolac |
| Galactose [PMID: 25326386]; also binds levetiracetam [PMID: 15210974], selectracetam [PMID: 18183537], brivaracetam [PMID: 17785672]. | Galactose, levetiracetam, selectracetam, brivaracetam |
| Nicotinate [PMID: 21953179]; binds nucleotides in the TM region [PMID: 19390693] | Nicotinate |
| thiamine pyrophosphate, thiamine monophosphate, (deoxy)nucleotides | thiamine pyrophosphate, thiamine monophosphate, deoxynucleotides |
| Cl-, HCO3-, SO42-, oxalate (?), Ch: Cl- | Cl-, HCO3-, SO42-, oxalate, Cl- |
| copper (I), cisplatin | copper, cisplatin |
| inorganic phosphate (divalent) | inorganic phosphate |
| GDP-fucose (?) | GDP-fucose |
| Q, H, A, N [PMID: 25451601] | Q, H, A, N |
| Mg2+ (Sr2+, Zn2+, Cu2+, Fe2+, Co2+, Ba2+, Cd2+) | Mg2+, Sr2+, Zn2+, Cu2+, Fe2+, Co2+, Ba2+, Cd2+ |
| sucrose, glucose, fructose [PMID: 25164149] | sucrose, glucose, fructose |

| | |
|---|---|
| tetraethylammonium (TEA), 1-methyl-4-phenylpyridinium (MPP), cimetidine, metformin, guanidine, procainamide, cephalexin, cephradine | tetraethylammonium, 1-methyl-4-phenylpyridinium, cimetidine, metformin, guanidine, procainamide, cephalexin, cephradine |
| Phosphate [PMID: 23791524] | Phosphate |
| Pyruvate [PMID: 22628554] [PMID: 22628558] | Pyruvate |
| Pyruvate [PMID: 22628558] | Pyruvate |
| Pyruvate [PMID: 27317664] | Pyruvate |
| Ca2+, K+, H+ (K+: [PMID: 20197279] [PMID: 15138253], [PMID: 17925330], yeast protein: [PMID: 17541427]), arguments against Ca2+ transport [PMID: 20197279]. | Ca2+, K+, H+ , K+ |
| Mg2+ [PMID: 17166836], also Sr2+, Fe2+ and Co2+ to a lesser extent [PMID: 18667602] | Mg2+, Sr2+, Fe2+, Co2+ |
| Mg2+, Sr2+, Ba2+, Fe2+, Cu2+ [PMID: 18667602] | Mg2+, Sr2+, Ba2+, Fe2+, Cu2+ |
| Mg2+, Sr2+, Ba2+ [PMID: 18667602] | Mg2+, Sr2+, Ba2+ |
| Mg2+ [PMID: 15804357] | Mg2+ |
| Mg2+, Fe2+, Cu2+, Mn2+ [PMID: 18667602] [PMID: 19940067] | Mg2+, Fe2+, Cu2+, Mn2+ |
| α-Me-glucose, D-glucose [PMID: 12590146], inhibited by phlorizin [PMID: 12590146], also inhibited by 2-DG, slightly by fructose and phloretin [PMID: 12590146]; urea [PMID: 26423860] | α-Me-glucose, D-glucose |
| Molybdate [PMID: 21464289] | Molybdate |
| Pyrophosphate [PMID: 10894769] | Pyrophosphate |
| Phosphorylated sphingolipids [PMID: 19074308]; phosphorylated Fingolimod (FTY720-P) [PMID: 21084291]; sphingosine-1-phosphate (S1P), dihydrosphingosine-1-phosphate (DH-S1P), phyto-S1P, C17-S1P [PMID: 21084291] | Phosphorylated sphingolipids, phosphorylated Fingolimod,sphingosine-1-phosphate , dihydrosphingosine-1-phosphate , phyto-S1P, C17-S1P |
| Ca2+, H+ [PMID: 23569283]; or Mn2+ [PMID: 27008884] [PMID: 28270545] | Ca2+, H+ , Mn2+ |

| | |
|---|---|
| Cholesterol [PMID: 14976318] | Cholesterol |
| Cholesterol [PMID: 17989073] [PMID: 17989072] [PMID: 27410046] | Cholesterol |
| cationic amino acids (Na+ indep.), large neutral L-amino acids (Na+ dep.) | cationic amino acids, large neutral L-amino acids |
| glycine ? | glycine |
| Mg2+ (Ba2+, Ni2+, Co2+, Fe2+, Mn2+) | Mg2+, Ba2+, Ni2+, Co2+, Fe2+, Mn2+ |
| sulfate, oxyanions selenium and chromium | sulfate, oxyanions selenium, chromium |
| di- and tri-peptides, protons, beta-lactam antibiotics | di-peptides, tri-peptides, protons, beta-lactam antibiotics |
| di- and tri-peptides | di-peptides, tri-peptides |
| Zn, not specific | Zn |
| Mg2+ [PMID: 18667602] | Mg2+ |
| LPC (lysophosphatidylcholine) form of DHA (docosahexaenoic acid) [PMID: 24828044] | lysophosphatidylcholine |
| Probably falsely identified as a tricarboxylate carrier in [PMID: 8132491] as discussed in [PMID: 11274051]. Likely transports pyridoxin or another heme precursor or ALAS2 cofactor [PMID: 11274051] [PMID: 12670026]. | pyridoxin |

***Table 14- Replacement list***

| One letter code | Amino acid name |
|---|---|
| A | Alanine |
| R | Arginine |
| N | Asparagine |
| D | Aspartic acid |
| C | Cysteine |
| Q | Glutamine |
| E | Glutamic acid |

| | |
|---|---|
| G | Glycine |
| H | Histidine |
| I | Isoleucine |
| L | Leucine |
| K | Lysine |
| M | Methionine |
| F | Phenylalanine |
| P | Proline |
| S | Serine |
| T | Threonine |
| W | Tryptophan |
| Y | Tyrosine |
| V | Valine |
| B | Asparagine |
| Z | Glutamine |
| J | Leucine or isoleucine |

***Table 15 – SLC families***

| HGNC Symbol | UniProt Accession | family | SLC name |
|---|---|---|---|
| CLN3 | Q13286 | CLN3 | ? |
| MFSD6 | Q6ZSS7 | MFSD# | ? |
| MFSD6L | Q8IWD5 | MFSD# | ? |
| MFSD4B | Q5TF39 | MFSD# | ? |
| MFSD8 | Q8NHS3 | MFSD# | ? |
| MFSD13A | Q14CX5 | MFSD# | ? |
| MFSD14B | Q5SR56 | MFSD# | ? |
| MFSD14A | Q96MC6 | MFSD# | ? |
| MFSD12 | Q6NUT3 | MFSD# | ? |
| MFSD9 | Q8NBP5 | MFSD# | ? |
| MFSD1 | Q9H3U5 | MFSD# | ? |
| MFSD11 | O43934 | MFSD# | ? |
| OCA2 | Q04671 | OCA2 | ? |
| SLC1A2 | P43004 | SLC1 | SLC1A2 |
| SLC1A6 | P48664 | SLC1 | SLC1A6 |
| SLC1A7 | O00341 | SLC1 | SLC1A7 |
| SLC1A3 | P43003 | SLC1 | SLC1A3 |
| SLC1A5 | Q15758 | SLC1 | SLC1A5 |
| SLC1A4 | P43007 | SLC1 | SLC1A4 |
| SLC1A1 | P43005 | SLC1 | SLC1A1 |

| | | | |
|---|---|---|---|
| SLC10A3 | P09131 | SLC10 | SLC10A3 |
| SLC10A5 | Q5PT55 | SLC10 | SLC10A5 |
| SLC10A4 | Q96EP9 | SLC10 | SLC10A4 |
| SLC10A6 | Q3KNW5 | SLC10 | SLC10A6 |
| SLC10A7 | Q0GE19 | SLC10 | SLC10A7 |
| SLC10A1 | Q14973 | SLC10 | SLC10A1 |
| SLC10A2 | Q12908 | SLC10 | SLC10A2 |
| SLC11A2 | P49281 | SLC11 | SLC11A2 |
| SLC11A1 | P49279 | SLC11 | SLC11A1 |
| SLC12A2 | P55011 | SLC12 | SLC12A2 |
| SLC12A6 | Q9UHW9 | SLC12 | SLC12A6 |
| SLC12A5 | Q9H2X9 | SLC12 | SLC12A5 |
| SLC12A1 | Q13621 | SLC12 | SLC12A1 |
| SLC12A4 | Q9UP95 | SLC12 | SLC12A4 |
| SLC12A7 | Q9Y666 | SLC12 | SLC12A7 |
| SLC12A3 | P55017 | SLC12 | SLC12A3 |
| SLC12A9 | Q9BXP2 | SLC12 | SLC12A9 |
| SLC12A8 | A0AV02 | SLC12 | SLC12A8 |
| SLC13A4 | Q9UKG4 | SLC13 | SLC13A4 |
| SLC13A3 | Q8WWT9 | SLC13 | SLC13A3 |
| SLC13A1 | Q9BZW2 | SLC13 | SLC13A1 |
| SLC13A2 | Q13183 | SLC13 | SLC13A2 |
| SLC13A5 | Q86YT5 | SLC13 | SLC13A5 |
| SLC14A2 | Q15849 | SLC14 | SLC14A2 |
| SLC14A1 | Q13336 | SLC14 | SLC14A1 |
| SLC15A2 | Q16348 | SLC15 | SLC15A2 |
| SLC15A1 | P46059 | SLC15 | SLC15A1 |
| SLC15A3 | Q8IY34 | SLC15 | SLC15A3 |
| SLC15A5 | A6NIM6 | SLC15 | SLC15A5 |
| SLC15A4 | Q8N697 | SLC15 | SLC15A4 |
| SLC16A2 | P36021 | SLC16 | SLC16A2 |
| SLC16A6 | O15403 | SLC16 | SLC16A6 |
| SLC16A10 | Q8TF71 | SLC16 | SLC16A10 |
| SLC16A14 | Q7RTX9 | SLC16 | SLC16A14 |
| SLC16A9 | Q7RTY1 | SLC16 | SLC16A9 |
| SLC16A5 | O15375 | SLC16 | SLC16A5 |
| SLC16A8 | O95907 | SLC16 | SLC16A8 |
| SLC16A1 | P53985 | SLC16 | SLC16A1 |
| SLC16A4 | O15374 | SLC16 | SLC16A4 |
| SLC16A12 | Q6ZSM3 | SLC16 | SLC16A12 |
| SLC16A7 | O60669 | SLC16 | SLC16A7 |
| SLC16A11 | Q8NCK7 | SLC16 | SLC16A11 |
| SLC16A3 | O15427 | SLC16 | SLC16A3 |
| SLC16A13 | Q7RTY0 | SLC16 | SLC16A13 |
| SLC17A8 | Q8NDX2 | SLC17 | SLC17A8 |
| SLC17A6 | Q9P2U8 | SLC17 | SLC17A6 |
| SLC17A7 | Q9P2U7 | SLC17 | SLC17A7 |
| SLC17A4 | Q9Y2C5 | SLC17 | SLC17A4 |
| SLC17A5 | Q9NRA2 | SLC17 | SLC17A5 |

| | | | |
|---|---|---|---|
| SLC17A1 | Q14916 | SLC17 | SLC17A1 |
| SLC17A2 | O00624 | SLC17 | SLC17A2 |
| SLC17A9 | Q9BYT1 | SLC17 | SLC17A9 |
| SLC17A3 | O00476 | SLC17 | SLC17A3 |
| SLC18A3 | Q16572 | SLC18 | SLC18A3 |
| SLC18A1 | P54219 | SLC18 | SLC18A1 |
| SLC18A2 | Q05940 | SLC18 | SLC18A2 |
| SLC18B1 | Q6NT16 | SLC18 | SLC18B1 |
| SLC19A1 | P41440 | SLC19 | SLC19A1 |
| SLC19A2 | O60779 | SLC19 | SLC19A2 |
| SLC19A3 | Q9BZV2 | SLC19 | SLC19A3 |
| SLC2A13 | Q96QE2 | SLC2 | SLC2A13 |
| SLC2A12 | Q8TD20 | SLC2 | SLC2A12 |
| SLC2A10 | O95528 | SLC2 | SLC2A10 |
| SLC2A9 | Q9NRM0 | SLC2 | SLC2A9 |
| SLC2A2 | P11168 | SLC2 | SLC2A2 |
| SLC2A14 | Q8TDB8 | SLC2 | SLC2A14 |
| SLC2A7 | Q6PXP3 | SLC2 | SLC2A7 |
| SLC2A4 | P14672 | SLC2 | SLC2A4 |
| SLC2A6 | Q9UGQ3 | SLC2 | SLC2A6 |
| SLC2A5 | P22732 | SLC2 | SLC2A5 |
| SLC2A3 | P11169 | SLC2 | SLC2A3 |
| SLC2A11 | Q9BYW1 | SLC2 | SLC2A11 |
| SLC2A1 | P11166 | SLC2 | SLC2A1 |
| SLC2A8 | Q9NY64 | SLC2 | SLC2A8 |
| SLC20A1 | Q8WUM9 | SLC20 | SLC20A1 |
| SLC20A2 | Q08357 | SLC20 | SLC20A2 |
| SV2A | Q7L0J3 | SLC22 | SLC22B1 |
| SV2C | Q496J9 | SLC22 | SLC22B3 |
| SLC22A23 | A1A5C7 | SLC22 | SLC22A23 |
| SV2B | Q7L1I2 | SLC22 | SLC22B2 |
| SLC22A14 | Q9Y267 | SLC22 | SLC22A14 |
| SLC22A16 | Q86VW1 | SLC22 | SLC22A16 |
| SLC22A6 | Q4U2R8 | SLC22 | SLC22A6 |
| SLC22A31 | A6NKX4 | SLC22 | SLC22A31 |
| SLC22A5 | O76082 | SLC22 | SLC22A5 |
| SLC22A3 | O75751 | SLC22 | SLC22A3 |
| SLC22A2 | O15244 | SLC22 | SLC22A2 |
| SLC22A1 | O15245 | SLC22 | SLC22A1 |
| SLC22A9 | Q8IVM8 | SLC22 | SLC22A9 |
| SLC22A12 | Q96S37 | SLC22 | SLC22A12 |
| SLC22A4 | Q9H015 | SLC22 | SLC22A4 |
| SLC22A13 | Q9Y226 | SLC22 | SLC22A13 |
| SLC22A11 | Q9NSA0 | SLC22 | SLC22A11 |
| SLC22A7 | Q9Y694 | SLC22 | SLC22A7 |
| SVOP | Q8N4V2 | SLC22 | SLC22B4 |
| SLC22A15 | Q8IZD6 | SLC22 | SLC22A15 |
| SLC22A25 | Q6T423 | SLC22 | SLC22A25 |
| SLC22A8 | Q8TCC7 | SLC22 | SLC22A8 |

| | | | |
|---|---|---|---|
| SLC22A10 | Q63ZE4 | SLC22 | SLC22A10 |
| SLC22A17 | Q8WUG5 | SLC22 | SLC22A17 |
| SVOPL | Q8N434 | SLC22 | SLC22B5 |
| MFSD10 | Q14728 | SLC22 | SLC22A32 |
| SLC22A18 | Q96BI1 | SLC22 | SLC22A18 |
| SLC22A24 | Q8N4F4 | SLC22 | SLC22A24 |
| SLC23A2 | Q9UGH3 | SLC23 | SLC23A2 |
| SLC23A3 | Q6PIS1 | SLC23 | SLC23A3 |
| SLC23A1 | Q9UHI7 | SLC23 | SLC23A1 |
| SLC24A1 | O60721 | SLC24 | SLC24A1 |
| SLC24A2 | Q9UI40 | SLC24 | SLC24A2 |
| SLC24A3 | Q9HC58 | SLC24 | SLC24A3 |
| SLC24A4 | Q8NFF2 | SLC24 | SLC24A4 |
| SLC24A5 | Q71RS6 | SLC24 | SLC24A5 |
| SLC25A12 | O75746 | SLC25 | SLC25A12 |
| SLC25A13 | Q9UJS0 | SLC25 | SLC25A13 |
| SLC25A24 | Q6NUK1 | SLC25 | SLC25A24 |
| SLC25A25 | Q6KCM7 | SLC25 | SLC25A25 |
| SLC25A23 | Q9BV35 | SLC25 | SLC25A23 |
| SLC25A46 | Q96AG3 | SLC25 | SLC25A46 |
| MTCH1 | Q9NZJ7 | SLC25 | SLC25A49 |
| SLC25A41 | Q8N5S1 | SLC25 | SLC25A41 |
| SLC25A28 | Q96A46 | SLC25 | SLC25A28 |
| SLC25A3 | Q00325 | SLC25 | SLC25A3 |
| SLC25A39 | Q9BZJ4 | SLC25 | SLC25A39 |
| SLC25A43 | Q8WUT9 | SLC25 | SLC25A43 |
| SLC25A37 | Q9NYZ2 | SLC25 | SLC25A37 |
| SLC25A40 | Q8TBP6 | SLC25 | SLC25A40 |
| SLC25A16 | P16260 | SLC25 | SLC25A16 |
| SLC25A14 | O95258 | SLC25 | SLC25A14 |
| SLC25A22 | Q9H936 | SLC25 | SLC25A22 |
| SLC25A27 | O95847 | SLC25 | SLC25A27 |
| SLC25A33 | Q9BSK2 | SLC25 | SLC25A33 |
| SLC25A19 | Q9HC21 | SLC25 | SLC25A19 |
| SLC25A42 | Q86VD7 | SLC25 | SLC25A42 |
| SLC25A18 | Q9H1K4 | SLC25 | SLC25A18 |
| SLC25A31 | Q9H0C2 | SLC25 | SLC25A31 |
| SLC25A32 | Q9H2D1 | SLC25 | SLC25A32 |
| SLC25A11 | Q02978 | SLC25 | SLC25A11 |
| SLC25A44 | Q96H78 | SLC25 | SLC25A44 |
| UCP3 | P55916 | SLC25 | SLC25A9 |
| SLC25A1 | P53007 | SLC25 | SLC25A1 |
| SLC25A36 | Q96CQ1 | SLC25 | SLC25A36 |
| SLC25A48 | Q6ZT89 | SLC25 | SLC25A48 |
| UCP2 | P55851 | SLC25 | SLC25A8 |
| SLC25A47 | Q6Q0C1 | SLC25 | SLC25A47 |
| SLC25A17 | O43808 | SLC25 | SLC25A17 |
| SLC25A53 | Q5H9E4 | SLC25 | SLC25A53 |
| UCP1 | P25874 | SLC25 | SLC25A7 |

| | | | |
|---|---|---|---|
| SLC25A34 | Q6PIV7 | SLC25 | SLC25A34 |
| SLC25A38 | Q96DW6 | SLC25 | SLC25A38 |
| MTCH2 | Q9Y6C9 | SLC25 | SLC25A50 |
| SLC25A29 | Q8N8R3 | SLC25 | SLC25A29 |
| SLC25A2 | Q9BXI2 | SLC25 | SLC25A2 |
| SLC25A15 | Q9Y619 | SLC25 | SLC25A15 |
| SLC25A20 | O43772 | SLC25 | SLC25A20 |
| SLC25A35 | Q3KQZ1 | SLC25 | SLC25A35 |
| SLC25A21 | Q9BQT8 | SLC25 | SLC25A21 |
| SLC25A4 | P12235 | SLC25 | SLC25A4 |
| SLC25A5 | P05141 | SLC25 | SLC25A5 |
| SLC25A6 | P12236 | SLC25 | SLC25A6 |
| SLC25A51 | Q9H1U9 | SLC25 | SLC25A51 |
| SLC25A52 | Q3SY17 | SLC25 | SLC25A52 |
| SLC25A30 | Q5SVS4 | SLC25 | SLC25A30 |
| SLC25A45 | Q8N413 | SLC25 | SLC25A45 |
| SLC25A10 | Q9UBX3 | SLC25 | SLC25A10 |
| SLC25A26 | Q70HW3 | SLC25 | SLC25A26 |
| SLC26A8 | Q96RN1 | SLC26 | SLC26A8 |
| SLC26A9 | Q7LBE3 | SLC26 | SLC26A9 |
| SLC26A4 | O43511 | SLC26 | SLC26A4 |
| SLC26A3 | P40879 | SLC26 | SLC26A3 |
| SLC26A6 | Q9BXS9 | SLC26 | SLC26A6 |
| SLC26A5 | P58743 | SLC26 | SLC26A5 |
| SLC26A2 | P50443 | SLC26 | SLC26A2 |
| SLC26A1 | Q9H2B4 | SLC26 | SLC26A1 |
| SLC26A7 | Q8TE54 | SLC26 | SLC26A7 |
| SLC26A11 | Q86WA9 | SLC26 | SLC26A11 |
| SLC26A10 | Q8NG04 | SLC26 | SLC26A10 |
| SLC27A3 | Q5K4L6 | SLC27 | SLC27A3 |
| SLC27A5 | Q9Y2P5 | SLC27 | SLC27A5 |
| SLC27A1 | Q6PCB7 | SLC27 | SLC27A1 |
| SLC27A4 | Q6P1M0 | SLC27 | SLC27A4 |
| SLC27A2 | O14975 | SLC27 | SLC27A2 |
| SLC27A6 | Q9Y2P4 | SLC27 | SLC27A6 |
| SLC28A3 | Q9HAS3 | SLC28 | SLC28A3 |
| SLC28A2 | O43868 | SLC28 | SLC28A2 |
| SLC28A1 | O00337 | SLC28 | SLC28A1 |
| SLC29A4 | Q7RTT9 | SLC29 | SLC29A4 |
| SLC29A3 | Q9BZD2 | SLC29 | SLC29A3 |
| SLC29A1 | Q99808 | SLC29 | SLC29A1 |
| SLC29A2 | Q14542 | SLC29 | SLC29A2 |
| SLC3A1 | Q07837 | SLC3 | SLC3A1 |
| SLC3A2 | P08195 | SLC3 | SLC3A2 |
| SLC30A5 | Q8TAD4 | SLC30 | SLC30A5 |
| SLC30A9 | Q6PML9 | SLC30 | SLC30A9 |
| SLC30A1 | Q9Y6M5 | SLC30 | SLC30A1 |
| SLC30A10 | Q6XR72 | SLC30 | SLC30A10 |
| SLC30A6 | Q6NXT4 | SLC30 | SLC30A6 |

| | | | |
|---|---|---|---|
| SLC30A4 | O14863 | SLC30 | SLC30A4 |
| SLC30A3 | Q99726 | SLC30 | SLC30A3 |
| SLC30A7 | Q8NEW0 | SLC30 | SLC30A7 |
| SLC30A8 | Q8IWU4 | SLC30 | SLC30A8 |
| SLC30A2 | Q9BRI3 | SLC30 | SLC30A2 |
| SLC31A1 | O15431 | SLC31 | SLC31A1 |
| SLC31A2 | O15432 | SLC31 | SLC31A2 |
| SLC32A1 | Q9H598 | SLC32 | SLC32A1 |
| SLC33A1 | O00400 | SLC33 | SLC33A1 |
| MFSD3 | Q96ES6 | SLC33 | SLC33A2 |
| SLC34A2 | O95436 | SLC34 | SLC34A2 |
| SLC34A1 | Q06495 | SLC34 | SLC34A1 |
| SLC34A3 | Q8N130 | SLC34 | SLC34A3 |
| SLC35F5 | Q8WV83 | SLC35 | SLC35F5 |
| SLC35F4 | A4IF30 | SLC35 | SLC35F4 |
| SLC35B2 | Q8TB61 | SLC35 | SLC35B2 |
| SLC35A5 | Q9BS91 | SLC35 | SLC35A5 |
| SLC35F3 | Q8IY50 | SLC35 | SLC35F3 |
| SLC35D3 | Q5M8T2 | SLC35 | SLC35D3 |
| SLC35G2 | Q8TBE7 | SLC35 | SLC35G2 |
| SLC35E1 | Q96K37 | SLC35 | SLC35E1 |
| SLC35F1 | Q5T1Q4 | SLC35 | SLC35F1 |
| SLC35E2B | P0CK96 | SLC35 | SLC35E2B |
| SLC35B3 | Q9H1N7 | SLC35 | SLC35B3 |
| SLC35A2 | P78381 | SLC35 | SLC35A2 |
| SLC35F2 | Q8IXU6 | SLC35 | SLC35F2 |
| SLC35F6 | Q8N357 | SLC35 | SLC35F6 |
| SLC35C2 | Q9NQQ7 | SLC35 | SLC35C2 |
| SLC35G1 | Q2M3R5 | SLC35 | SLC35G1 |
| SLC35C1 | Q96A29 | SLC35 | SLC35C1 |
| SLC35D1 | Q9NTN3 | SLC35 | SLC35D1 |
| SLC35E4 | Q6ICL7 | SLC35 | SLC35E4 |
| SLC35G3 | Q8N808 | SLC35 | SLC35G3 |
| SLC35G4 | P0C7Q5 | SLC35 | SLC35G4 |
| SLC35G5 | Q96KT7 | SLC35 | SLC35G5 |
| SLC35G6 | P0C7Q6 | SLC35 | SLC35G6 |
| SLC35A1 | P78382 | SLC35 | SLC35A1 |
| SLC35D2 | Q76EJ3 | SLC35 | SLC35D2 |
| SLC35B4 | Q969S0 | SLC35 | SLC35B4 |
| SLC35A3 | Q9Y2D2 | SLC35 | SLC35A3 |
| SLC35A4 | Q96G79 | SLC35 | SLC35A4 |
| SLC35B1 | P78383 | SLC35 | SLC35B1 |
| SLC35E3 | Q7Z769 | SLC35 | SLC35E3 |
| TMEM241 | Q24JQ0 | SLC35 | SLC35D4 |
| SLC35E2A | P0CK97 | SLC35 | SLC35E2 |
| SLC36A4 | Q6YBV0 | SLC36 | SLC36A4 |
| SLC36A2 | Q495M3 | SLC36 | SLC36A2 |
| SLC36A1 | Q7Z2H8 | SLC36 | SLC36A1 |
| SLC36A3 | Q495N2 | SLC36 | SLC36A3 |

| | | | |
|---|---|---|---|
| SLC37A1 | P57057 | SLC37 | SLC37A1 |
| SLC37A2 | Q8TED4 | SLC37 | SLC37A2 |
| SLC37A3 | Q8NCC5 | SLC37 | SLC37A3 |
| SLC37A4 | O43826 | SLC37 | SLC37A4 |
| SLC38A10 | Q9HBR0 | SLC38 | SLC38A10 |
| SLC38A9 | Q8NBW4 | SLC38 | SLC38A9 |
| SLC38A4 | Q969I6 | SLC38 | SLC38A4 |
| SLC38A2 | Q96QD8 | SLC38 | SLC38A2 |
| SLC38A3 | Q99624 | SLC38 | SLC38A3 |
| SLC38A1 | Q9H2H9 | SLC38 | SLC38A1 |
| SLC38A5 | Q8WUX1 | SLC38 | SLC38A5 |
| SLC38A7 | Q9NVC3 | SLC38 | SLC38A7 |
| SLC38A6 | Q8IZM9 | SLC38 | SLC38A6 |
| SLC38A8 | A6NNN8 | SLC38 | SLC38A8 |
| SLC38A11 | Q08AI6 | SLC38 | SLC38A11 |
| SLC39A10 | Q9ULF5 | SLC39 | SLC39A10 |
| SLC39A6 | Q13433 | SLC39 | SLC39A6 |
| SLC39A12 | Q504Y0 | SLC39 | SLC39A12 |
| SLC39A4 | Q6P5W5 | SLC39 | SLC39A4 |
| SLC39A5 | Q6ZMH5 | SLC39 | SLC39A5 |
| SLC39A14 | Q15043 | SLC39 | SLC39A14 |
| SLC39A7 | Q92504 | SLC39 | SLC39A7 |
| SLC39A8 | Q9C0K1 | SLC39 | SLC39A8 |
| SLC39A13 | Q96H72 | SLC39 | SLC39A13 |
| SLC39A11 | Q8N1S5 | SLC39 | SLC39A11 |
| SLC39A1 | Q9NY26 | SLC39 | SLC39A1 |
| SLC39A3 | Q9BRY0 | SLC39 | SLC39A3 |
| SLC39A2 | Q9NP94 | SLC39 | SLC39A2 |
| SLC39A9 | Q9NUM3 | SLC39 | SLC39A9 |
| SLC4A2 | P04920 | SLC4 | SLC4A2 |
| SLC4A3 | P48751 | SLC4 | SLC4A3 |
| SLC4A7 | Q9Y6M7 | SLC4 | SLC4A7 |
| SLC4A5 | Q9BY07 | SLC4 | SLC4A5 |
| SLC4A10 | Q6U841 | SLC4 | SLC4A10 |
| SLC4A8 | Q2Y0W8 | SLC4 | SLC4A8 |
| SLC4A4 | Q9Y6R1 | SLC4 | SLC4A4 |
| SLC4A9 | Q96Q91 | SLC4 | SLC4A9 |
| SLC4A1 | P02730 | SLC4 | SLC4A1 |
| SLC4A11 | Q8NBS3 | SLC4 | SLC4A11 |
| SLC40A1 | Q9NP59 | SLC40 | SLC40A1 |
| SLC41A2 | Q96JW4 | SLC41 | SLC41A2 |
| SLC41A1 | Q8IVJ1 | SLC41 | SLC41A1 |
| SLC41A3 | Q96GZ6 | SLC41 | SLC41A3 |
| RHCG | Q9UBD6 | SLC42 | SLC42A3 |
| RHBG | Q9H310 | SLC42 | SLC42A2 |
| RHAG | Q02094 | SLC42 | SLC42A1 |
| SLC43A2 | Q8N370 | SLC43 | SLC43A2 |
| SLC43A1 | O75387 | SLC43 | SLC43A1 |
| SLC43A3 | Q8NBI5 | SLC43 | SLC43A3 |

| | | | |
|---|---|---|---|
| SLC44A5 | Q8NCS7 | SLC44 | SLC44A5 |
| SLC44A4 | Q53GD3 | SLC44 | SLC44A4 |
| SLC44A2 | Q8IWA5 | SLC44 | SLC44A2 |
| SLC44A1 | Q8WWI5 | SLC44 | SLC44A1 |
| SLC44A3 | Q8N4M1 | SLC44 | SLC44A3 |
| SLC45A1 | Q9Y2W3 | SLC45 | SLC45A1 |
| SLC45A4 | Q5BKX6 | SLC45 | SLC45A4 |
| SLC45A3 | Q96JT2 | SLC45 | SLC45A3 |
| SLC45A2 | Q9UMX9 | SLC45 | SLC45A2 |
| SLC46A2 | Q9BY10 | SLC46 | SLC46A2 |
| SLC46A3 | Q7Z3Q1 | SLC46 | SLC46A3 |
| SLC46A1 | Q96NT5 | SLC46 | SLC46A1 |
| SLC47A2 | Q86VL8 | SLC47 | SLC47A2 |
| SLC47A1 | Q96FL8 | SLC47 | SLC47A1 |
| SLC48A1 | Q6P1K1 | SLC48 | SLC48A1 |
| SLC49A3 | Q6UXD7 | SLC49 | SLC49A3 |
| FLVCR1 | Q9Y5Y0 | SLC49 | SLC49A1 |
| FLVCR2 | Q9UPI3 | SLC49 | SLC49A2 |
| DIRC2 | Q96SL1 | SLC49 | SLC49A4 |
| SLC5A3 | P53794 | SLC5 | SLC5A3 |
| SLC5A9 | Q2M3M2 | SLC5 | SLC5A9 |
| SLC5A11 | Q8WWX8 | SLC5 | SLC5A11 |
| SLC5A2 | P31639 | SLC5 | SLC5A2 |
| SLC5A1 | P13866 | SLC5 | SLC5A1 |
| SLC5A4 | Q9NY91 | SLC5 | SLC5A4 |
| SLC5A5 | Q92911 | SLC5 | SLC5A5 |
| SLC5A6 | Q9Y289 | SLC5 | SLC5A6 |
| SLC5A12 | Q1EHB4 | SLC5 | SLC5A12 |
| SLC5A8 | Q8N695 | SLC5 | SLC5A8 |
| SLC5A10 | A0PJK1 | SLC5 | SLC5A10 |
| SLC5A7 | Q9GZV3 | SLC5 | SLC5A7 |
| SLC50A1 | Q9BRV3 | SLC50 | SLC50A1 |
| SLC51A | Q86UW1 | SLC51 | SLC51A |
| SLC51B | Q86UW2 | SLC51 | SLC51B |
| SLC52A3 | Q9NQ40 | SLC52 | SLC52A3 |
| SLC52A1 | Q9NWF4 | SLC52 | SLC52A1 |
| SLC52A2 | Q9HAB3 | SLC52 | SLC52A2 |
| XPR1 | Q9UBH6 | SLC53 | SLC53A1 |
| MPC1L | P0DKB6 | SLC54 | SLC54A3 |
| MPC2 | O95563 | SLC54 | SLC54A2 |
| MPC1 | Q9Y5U8 | SLC54 | SLC54A1 |
| LETM1 | O95202 | SLC55 | SLC55A1 |
| LETM2 | Q2VYF4 | SLC55 | SLC55A2 |
| LETMD1 | Q6P1Q0 | SLC55 | SLC55A3 |
| SFXN5 | Q8TD22 | SLC56 | SLC56A5 |
| SFXN4 | Q6P4A7 | SLC56 | SLC56A4 |
| SFXN1 | Q9H9B4 | SLC56 | SLC56A1 |
| SFXN2 | Q96NB2 | SLC56 | SLC56A2 |
| SFXN3 | Q9BWM7 | SLC56 | SLC56A3 |

| | | | |
|---|---|---|---|
| NIPAL4 | Q0D2K0 | SLC57 | SLC57A6 |
| NIPAL1 | Q6NVV3 | SLC57 | SLC57A3 |
| NIPAL3 | Q6P499 | SLC57 | SLC57A5 |
| NIPAL2 | Q9H841 | SLC57 | SLC57A4 |
| NIPA2 | Q8N8Q9 | SLC57 | SLC57A2 |
| NIPA1 | Q7RTP0 | SLC57 | SLC57A1 |
| TUSC3 | Q13454 | SLC58 | SLC58A2 |
| MAGT1 | Q9H0U3 | SLC58 | SLC58A1 |
| MFSD2A | Q8NA29 | SLC59 | SLC59A1 |
| MFSD2B | A6NFX1 | SLC59 | SLC59A2 |
| SLC6A5 | Q9Y345 | SLC6 | SLC6A5 |
| SLC6A16 | Q9GZN6 | SLC6 | SLC6A16 |
| SLC6A15 | Q9H2J7 | SLC6 | SLC6A15 |
| SLC6A17 | Q9H1V8 | SLC6 | SLC6A17 |
| SLC6A9 | P48067 | SLC6 | SLC6A9 |
| SLC6A14 | Q9UN76 | SLC6 | SLC6A14 |
| SLC6A7 | Q99884 | SLC6 | SLC6A7 |
| SLC6A8 | P48029 | SLC6 | SLC6A8 |
| SLC6A19 | Q695T7 | SLC6 | SLC6A19 |
| SLC6A11 | P48066 | SLC6 | SLC6A11 |
| SLC6A4 | P31645 | SLC6 | SLC6A4 |
| SLC6A18 | Q96N87 | SLC6 | SLC6A18 |
| SLC6A3 | Q01959 | SLC6 | SLC6A3 |
| SLC6A6 | P31641 | SLC6 | SLC6A6 |
| SLC6A2 | P23975 | SLC6 | SLC6A2 |
| SLC6A12 | P48065 | SLC6 | SLC6A12 |
| SLC6A13 | Q9NSD5 | SLC6 | SLC6A13 |
| SLC6A1 | P30531 | SLC6 | SLC6A1 |
| SLC6A20 | Q9NP91 | SLC6 | SLC6A20 |
| MFSD4A | Q8N468 | SLC60 | SLC60A1 |
| MFSD5 | Q6N075 | SLC61 | SLC61A1 |
| ANKH | Q9HCJ1 | SLC62 | SLC62A1 |
| SPNS2 | Q8IVW8 | SLC63 | SLC63A2 |
| SPNS1 | Q9H2V7 | SLC63 | SLC63A1 |
| SPNS3 | Q6ZMD2 | SLC63 | SLC63A3 |
| TMEM165 | Q9HC07 | SLC64 | SLC64A1 |
| NPC1L1 | Q9UHC9 | SLC65 | SLC65A2 |
| NPC1 | O15118 | SLC65 | SLC65A1 |
| SLC7A14 | Q8TBB6 | SLC7 | SLC7A14 |
| SLC7A2 | P52569 | SLC7 | SLC7A2 |
| SLC7A4 | O43246 | SLC7 | SLC7A4 |
| SLC7A1 | P30825 | SLC7 | SLC7A1 |
| SLC7A3 | Q8WY07 | SLC7 | SLC7A3 |
| SLC7A8 | Q9UHI5 | SLC7 | SLC7A8 |
| SLC7A10 | Q9NS82 | SLC7 | SLC7A10 |
| SLC7A6 | Q92536 | SLC7 | SLC7A6 |
| SLC7A7 | Q9UM01 | SLC7 | SLC7A7 |
| SLC7A5 | Q01650 | SLC7 | SLC7A5 |
| SLC7A11 | Q9UPY5 | SLC7 | SLC7A11 |

| | | | |
|---|---|---|---|
| SLC7A9 | P82251 | SLC7 | SLC7A9 |
| SLC7A13 | Q8TCU3 | SLC7 | SLC7A13 |
| SLC8A1 | P32418 | SLC8 | SLC8A1 |
| SLC8A3 | P57103 | SLC8 | SLC8A3 |
| SLC8A2 | Q9UPR5 | SLC8 | SLC8A2 |
| SLC8B1 | Q6J4K2 | SLC8 | SLC8B1 |
| SLC9C1 | Q4G0N8 | SLC9 | SLC9C1 |
| SLC9C2 | Q5TAH2 | SLC9 | SLC9C2 |
| SLC9A5 | Q14940 | SLC9 | SLC9A5 |
| SLC9A3 | P48764 | SLC9 | SLC9A3 |
| SLC9A1 | P19634 | SLC9 | SLC9A1 |
| SLC9A2 | Q9UBY0 | SLC9 | SLC9A2 |
| SLC9A4 | Q6AI14 | SLC9 | SLC9A4 |
| SLC9A7 | Q96T83 | SLC9 | SLC9A7 |
| SLC9A6 | Q92581 | SLC9 | SLC9A6 |
| SLC9A9 | Q8IVB4 | SLC9 | SLC9A9 |
| SLC9A8 | Q9Y2E8 | SLC9 | SLC9A8 |
| SLC9B2 | Q86UD5 | SLC9 | SLC9B2 |
| SLC9B1 | Q4ZJI4 | SLC9 | SLC9B1 |
| SLCO5A1 | Q9H2Y9 | SLCO | SLCO5A1 |
| SLCO4C1 | Q6ZQN7 | SLCO | SLCO4C1 |
| SLCO4A1 | Q96BD0 | SLCO | SLCO4A1 |
| SLCO6A1 | Q86UG4 | SLCO | SLCO6A1 |
| SLCO1C1 | Q9NYB5 | SLCO | SLCO1C1 |
| SLCO3A1 | Q9UIG8 | SLCO | SLCO3A1 |
| SLCO2B1 | O94956 | SLCO | SLCO2B1 |
| SLCO1B3 | Q9NPD5 | SLCO | SLCO1B3 |
| SLCO1B1 | Q9Y6L6 | SLCO | SLCO1B1 |
| SLCO1A2 | P46721 | SLCO | SLCO1A2 |
| SLCO2A1 | Q92959 | SLCO | SLCO2A1 |
| SLCO1B7 | G3V0H7 | SLCO | SLCO1B7 |

***Table 16 – Configuration of the Rule Engine node***

| |
|---|
| MISSING $InChICode_standardised$ AND MISSING $InChI_original_database$ => "EMPTY_both" |
| $InChI_original_database$ LIKE $InChICode_standardised$ => "YES" |
| MISSING $InChI_original_database$ => "EMPTY_original" |
| MISSING $InChICode_standardised$ => "EMPTY_textsearch" |
| TRUE => "NO" |

***Table 17 – Results SLC and their substrates***

| SLC | Unique count*(substrate) | Unique concatenate*(substrate) | Unique concatenate*(source) |
|---|---|---|---|
| SLC15A1 | 263 | 5-aminolevulinic acid, Amoxicillin, Bestatin, Cefadroxil, Cephalexin, Enalapril, Oseltamivir, Valacyclovir, (1S,2S,5R,6S)-2-(L-Alanylamino)bicyclo[3.1.0]hexane-2,6-dicarboxylic acid, (2R)-2-amino-N-[2-(2,5-dimethoxyphenyl)-2-hydroxyethyl]-3-sulfanylpropanamide, (2S)-2-amino-3-[3-(naphthalen-1-yl)-5-(phosphonomethyl)phenyl]propanoic acid, (2S)-2-amino-N-[2-(2,5-dimethoxyphenyl)-2-hydroxyethyl]-3-(4-hydroxyphenyl)propanamide, (2S)-3-hydroxy-2-{[(2S,4R)-4-hydroxypyrrolidin-2-yl]formamido}propanoic acid, (2S)-N-[(2S)-1-[(2S)-2-carbamoylpyrrolidin-1-yl]-1-oxo-3-(2-propyl-1H-imidazol-4-yl)propan-2-yl]-5-oxopyrrolidine-2-carboxamide, (3S)-3-(L-lysylamino)-4-oxobutanoic acid, (4S)-4-amino-4-{[2-(2,5-dimethoxyphenyl)-2-hydroxyethyl]carbamoyl}butanoic acid, 2-(1H-1,3-benzodiazol-1-ylmethoxy)ethyl (2S)-2-amino-3-methylbutanoate, 2-(9H-purin-9-ylmethoxy)ethyl (2S)-2-amino-3-methylbutanoate, 2-amino-N-[2-(2,5-dimethoxyphenyl)-2-hydroxyethyl]-3-methylbutanamide, 2-{1H-pyrrolo[2,3-c]pyridin-1-ylmethoxy}ethyl (2S)-2-amino-3-methylbutanoate, 4-(4-methoxyphenyl)-l-phenylalanyl sarcosine, 4-aminophenylacetic acid, 5-Amino-2-benzyl-8-[(diaminomethylene)amino]-4-oxooctanoic acid, Ala-Ala, Ala-Ala-Ala, Ala-Asn-Ser, Ala-Asp, Ala-Asp-Tyr, Ala-Ile-Asp, Ala-Lys, Ala-Phe, Ala-Phe-Pro, Ala-Pro, Ala-Pro-Gly, Ala-Pro-Leu, Ala-Trp, Ala-Tyr, Aminolevulinic acid, Ampicillin, Arg-Gly, Arg-His-Asp, Arg-Ile-Gln, Arg-Ile-Thr, Arg-Pro-Ser, Arg-Ser-Ser, Asn-Arg-Leu, Asn-Gly-His, Asp-Arg-Arg, Asp-Asp, Asp-Cys-Asp, Asp-Gly, Asp-Lys, Asp-Trp, Asp-Tyr-Thr, Asp-Val, Benazepril, Benzenehexanoic acid, delta- | UCSF-FDA, Metrabase, IUPHAR, Bioparadigms |

| | | | |
|---|---|---|---|
| | | amino-gamma-oxo-, (S)-, Captopril, Carnosine, Cefaclor, Cefamandole, Cefazoline, Cefdinir, Cefixime, Cefoxitin, Cefpirome, Cefpodoxime, Ceftibuten, Cefuroxime, Cephaloridine, Cephalothin, Cephradine, Cyclacillin, Cyclo(Gly-Phe), Cyclo(Gly-Tyr), Cyclo(Phe-Ser), Cyclo(Ser-Tyr), Cys-Asn-Met, Cys-Gly, Cys-Leu-Tyr, Dexamethasone, Enalaprilat, Eprosartan, Flucloxacillin, Formyl-Met-Leu-Phe, Fosinopril, Furosemide, Gln-Gln, Gln-Glu, Gln-Glu-Ile, Glu-Glu, Glu-Glu-Glu, Glu-Glu-Ser, Glu-Gly, Glu-Lys, Glu-Phe-Tyr, Glu-Ser-Met, Gly-Arg, Gly-Asp, Gly-Glu, Gly-Gly, Gly-Gly-Gly, Gly-Gly-Leu, Gly-His, Gly-L-Pro, Gly-Leu, Gly-Leu-Gly, Gly-Leu-Phe, Gly-Lys, Gly-Phe, Gly-Sar, Gly-Thr-Asn, Gly-Trp, Gly-Trp-Val, Gly-Tyr, Gly-Tyr-Trp, Glycyl-L-glutamine, His-Gly, His-His, His-Thr-Asn, His-Trp, Ile-Leu-Met, Ile-Trp, Ile-Val-Tyr, Isospaglumic acid, L-Cephalexin, L-Phe-L-Ala, L-Phe-L-Pro, L-Pro-L-Leu, L-Pro-L-Ser, L-Tyr-Gly-Gly, L-dopa-L-Phe, Leu-Ala-Arg, Leu-Gly-Gly, Leu-Leu, Leu-Leu-Asp, Levovirin, Lisinopril, Losartan, Lys-Ala, Lys-Glu, Lys-Gly, Lys-Lys, Lys-Pro, Lys-Pro-Val, Lys-Val, Met-Glu-Tyr, Met-Leu-Asn, Metoprolol, Midodrine, N-[(2S)-1-Phenyl-2-propanyl]-L-lysinamide, Phe-Ala-NH2, Phe-Gly, Phe-Gly-Val, Phe-Leu-Ala, Phe-Phe, Phenylalanyl Tyrosine, Piperacillin, Pro-Arg, Pro-Pro, QCPac, QSP, Ribavirin, Ser-Asn-Asn, Ser-Cys-Glu, Ser-Gln-Phe, Ser-Leu-Ala, Ser-Pro-Ile, Ser-Ser, Ser-Ser-Ser, Sulfasalazine, Temocapril, Temocaprilat, Thr-Lys-Tyr, Thr-Met-Phe, Trp-Ala, Trp-Glu-Asp, Trp-Gly, Trp-Gly-Tyr, Trp-Pro-Tyr, Trp-Trp-Trp, Trp-Val, Tyr-Ala, Tyr-Gly, Tyr-Phe, Tyr-Tyr, Val-Lys, Val-Ser-Thr, Val-Trp, Val-Tyr, Val-Val, Val-Val-Val, alafosfalin, cefmetazole, cefodizime, cefotaxime, cefsulodin, ceftriaxone, cefuroxime axetil, cephaloglycin, ceronapril, desglymidodrine, irbesartan, methyldopa-L-phenylalanine, muramyl dipeptide, penicillin G, valganciclovir, valsartan, zofenopril, His-Leu-lopinavir, | |

| | | | |
|---|---|---|---|
| | | cefadroxil, cyclacillin, fMet-Leu-Phe, mifamurtide, valacyclovir, 3-[[(3R)-3-amino-4-methyl-1-thioxopentyl]amino]-, (3S)-butanoic acid, 3-[[(3S)-3-amino-1-oxobutyl]amino]-5-methyl-, (3S)-hexanoic acid, Acidocillin, Ala-Orn, Ala-Pyrr, BipR-OH, D-Phe-Ala, D-Phe-L-Gln, FITC-Val-OCH3, FK089, Gly-Ser(rebapimide), Ile-thiazolidide, L-4,4'-biphenylalanyl-L-Proline, L-Ala-gamma-D-Glu-m-diaminopimelic acid, L-Dopa-L-Phenyalalanine, L-Phe-D-Pro, L-Val-L-Val ACV, L-Valine-Didanosine, Levovirin 5'-(L)-prolinylvalinate, Lys-FITC-OCH3, Lys-FITC-OH, PCQac, Phe-Sar, Pyrr-Ala, R-sulpiride, Ser(rebapimide)-Gly, Val-Cytarabine, Val-Floxuridine, Val-Gemcitabine, Val-Zidovudine, cyclo-trans-4-L-hydroxyprolyl-L-serine, levovirin 2',3'-(L)-bis-valinate, levovirin 5'-(L)-alaninate, levovirin 5'-(L)-isoleu-cinate, levovirin 5'-(L)-leucinate, levovirin 5'-(L)-sarcosinate, levovirin 5'-(L)-valinylprolinate, levovirin 50-(D)-valinate, levovirin 50-(L)-valinate, Glycylsarcosine, beta-Ala-Lys-AMCA, D-Ala-Lys-AMCA, dipeptides, protons, tripeptides, beta-lactam antibiotics, di-peptides, tripeptides | |
| SLC22A1 | 183 | Acyclovir, Metformin, Oxaliplatin, Pentamidine, Ranitidine, Tetraethylammonium, (2S,4S)-4-{2-[4-carbamoyl-5-(carbamoylamino)thiophen-2-yl]phenoxy}-N,N-dimethylpyrrolidine-2-carboxamide, (3-{2-[4-carbamoyl-5-(carbamoylamino)thiophen-2-yl]phenoxy}-2-hydroxypropyl)trimethylazanium, (S)-Xamoterol, 1-methyl-4-phenylpyridinium, 2-(carbamoylamino)-5-(1-methanesulfonyl-1,2,3,6-tetrahydropyridin-4-yl)thiophene-3-carboxamide, 2-(carbamoylamino)-5-(2,4-dihydroxyphenyl)thiophene-3-carboxamide, 2-(carbamoylamino)-5-(2-hydroxy-4-methoxyphenyl)thiophene-3-carboxamide, 2-(carbamoylamino)-5-(2-{2-[(propan-2-yl)amino]propoxy}phenyl)thiophene-3-carboxamide, 2-(carbamoylamino)-5-(2-{2-hydroxy-3-[(2-hydroxyethyl)amino]propoxy}phenyl)thio- | UCSF-FDA, Metrabase, IUPHAR, Bioparadigms |

| | | | |
|---|---|---|---|
| | | phene-3-carboxamide, 2-(carbamoylamino)-5-(2-{2-hydroxy-3-[(propan-2-yl)amino]propoxy}phenyl)thiophene-3-carboxamide, 2-(carbamoylamino)-5-(2-{3-[(2-ethoxyethyl)amino]-2-hydroxypropoxy}phenyl)thiophene-3-carboxamide, 2-(carbamoylamino)-5-(2-{[(3R)-1-(2-methoxyethyl)pyrrolidin-3-yl]oxy}phenyl)thiophene-3-carboxamide, 2-(carbamoylamino)-5-(2-{[(3S)-1-(2-methoxyethyl)pyrrolidin-3-yl]oxy}phenyl)thiophene-3-carboxamide, 2-(carbamoylamino)-5-(2-{[(3S,5S)-5-(fluoromethyl)pyrrolidin-3-yl]oxy}phenyl)thiophene-3-carboxamide, 2-(carbamoylamino)-5-(2-{[(3S,5S)-5-(hydroxymethyl)pyrrolidin-3-yl]oxy}phenyl)thiophene-3-carboxamide, 2-(carbamoylamino)-5-(2-{[(3S,5S)-5-(methoxymethyl)pyrrolidin-3-yl]oxy}phenyl)thiophene-3-carboxamide, 2-(carbamoylamino)-5-[1-(ethanesulfonyl)-1,2,3,6-tetrahydropyridin-4-yl]thiophene-3-carboxamide, 2-(carbamoylamino)-5-[2-(1-carbamoylethoxy)phenyl]thiophene-3-carboxamide, 2-(carbamoylamino)-5-[2-(piperidin-3-ylmethoxy)phenyl]thiophene-3-carboxamide, 2-(carbamoylamino)-5-[2-(piperidin-4-yloxy)phenyl]thiophene-3-carboxamide, 2-(carbamoylamino)-5-[3-(carbamoylmethoxy)phenyl]thiophene-3-carboxamide, 2-(carbamoylamino)-5-[3-(pyrrolidin-3-yloxy)phenyl]thiophene-3-carboxamide, 2-(carbamoylamino)-5-[4-(carbamoylmethoxy)phenyl]thiophene-3-carboxamide, 2-(carbamoylamino)-5-[4-(diethylcarbamoyl)-2-methoxyphenyl]thiophene-3-carboxamide, 2-(carbamoylamino)-5-[4-(ethanesulfonyl)phenyl]thiophene-3-carboxamide, 2-(carbamoylamino)-5-[4-(pyrrolidin-3-yloxy)phenyl]thiophene-3-carboxamide, 2-(carbamoylamino)-5-[5-cyano-2-(pyrrolidin-3-yloxy)phenyl]thiophene-3-carboxamide, 2-(carbamoylamino)-5-{2-[(1-methylpyrrolidin-3-yl)methoxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{2-[(2S)-2-hydroxy-3-[(2-hydroxyethyl)amino]propoxy]phenyl}thio- | |

| | | | |
|---|---|---|---|
| | | phene-3-carboxamide, 2-(carbamoylamino)-5-{2-[(2S)-2-hydroxy-3-[(propan-2-yl)amino]propoxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{2-[(2S)-3-(cyclopropylamino)-2-hydroxypropoxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{2-[(3R)-pyrrolidin-3-ylmethoxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{2-[(3R)-pyrrolidin-3-yloxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{2-[(3S)-pyrrolidin-3-ylmethoxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{2-[(3S)-pyrrolidin-3-yloxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{2-[2-(methylamino)ethoxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{2-[2-hydroxy-3-(methylamino)propoxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{2-[3-(cyclobutylamino)-2-hydroxypropoxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{2-[3-(cyclopropylamino)-2-hydroxy-2-methylpropoxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{2-[3-(cyclopropylamino)-2-hydroxypropoxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{2-[3-(ethylamino)-2-hydroxypropoxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{3-methyl-2-[(3R)-pyrrolidin-3-yloxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{3-methyl-2-[(3S)-pyrrolidin-3-yloxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{5-chloro-2-[(3S)-pyrrolidin-3-yloxy]phenyl}thiophene-3-carboxamide, 2-(carbamoylamino)-5-{5-fluoro-2-[(3S)-pyrrolidin-3-yloxy]phenyl}thiophene-3-carboxamide, 2-({2-[2-(methylamino)pyrimidin-4-yl]-1H-indol-5-yl}formamido)-3-[phenyl(pyridin-2-yl)amino]propanoic acid, 2-[2-(2-aminoethoxy)phenyl]-5-(carbamoylamino)-1,3-thiazole-4-carboxamide, 2-[2-(3-aminopropoxy)phenyl]-5-(carbamoylamino)-1,3-thiazole-4-carboxamide, 3,5-diamino-6-chloro-N-[(2S)-2-ethanimidamido-3-(2-methylphenyl)propyl]pyrazine-2-carboxamide, | |

| | | | |
|---|---|---|---|
| | | 3-(carbamoylamino)-5-(2-hydroxyphenyl)thiophene-2-carboxamide, 3-[(hydroxydiphenylacetyl)oxy]-1,1-dimethylpiperidinium, 3-{2-[4-carbamoyl-5-(carbamoylamino)thiophen-2-yl]phenoxy}-1,1-dimethylpyrrolidin-1-ium, 3-{2-amino-3-cyano-6-[2-(cyclopropylmethoxy)-6-hydroxyphenyl]pyridin-4-yl}-1,1-dimethylpiperidin-1-ium, 3-{7-[2-(cyclopropylmethoxy)-6-hydroxyphenyl]-2-oxo-1H,2H,4H-pyrido[2,3-d][1,3]oxazin-5-yl}-1,1-dimethylpiperidin-1-ium, 4-({[2-({6-amino-9-[(2R,3R,4S,5S)-5-(ethylcarbamoyl)-3,4-dihydroxyoxolan-2-yl]-9H-purin-2-yl}formamido)ethyl]carbamoyl}amino)-1-methylpyridin-1-ium, 4-amino-2-[2-(pyrrolidin-3-yloxy)phenyl]thieno[3,2-c]pyridine-7-carboxamide, 5-(2-{3-[(butan-2-yl)amino]-2-hydroxypropoxy}phenyl)-2-(carbamoylamino)thiophene-3-carboxamide, 5-(carbamoylamino)-2-[2-(pyrrolidin-3-yloxy)phenyl]-1,3-thiazole-4-carboxamide, 5-(carbamoylamino)-2-{[(propan-2-yl)carbamoyl]amino}-1,3-thiazole-4-carboxamide, 5-[2-(2-aminoethoxy)phenyl]-2-(carbamoylamino)thiophene-3-carboxamide, 5-[2-(3-amino-2-hydroxypropoxy)phenyl]-2-(carbamoylamino)thiophene-3-carboxamide, 5-[2-(3-aminopropoxy)phenyl]-2-(carbamoylamino)thiophene-3-carboxamide, 5-[2-(azepan-4-yloxy)phenyl]-2-(carbamoylamino)thiophene-3-carboxamide, 5-[2-(azetidin-3-yloxy)phenyl]-2-(carbamoylamino)thiophene-3-carboxamide, 5-[2-(benzyloxy)-4-(carbamoylmethoxy)phenyl]-2-(carbamoylamino)thiophene-3-carboxamide, 5-[4-(2-aminoethoxy)phenyl]-2-(carbamoylamino)thiophene-3-carboxamide, 5-{2-[(2R)-2-aminopropoxy]phenyl}-2-(carbamoylamino)thiophene-3-carboxamide, 5-{2-[(4-aminocyclohexyl)oxy]phenyl}-2-(carbamoylamino)thiophene-3-carboxamide, 5-{3-bromo-5-fluoro-2-[(3S)-pyrrolidin-3-yloxy]phenyl}-2-(carbamoylamino)thiophene-3-carboxamide, Acebutolol, Acetylcholine, Agmatine, Amiloride, Atenolol, Atropine, Azidoprocainamide methoiodide, Bamet-R2, | |

| | | | |
|---|---|---|---|
| | | Benzamil, Buformin, Camostat, Choline, Cimetidine, Citalopram, Clidinium, Clonidine, Cyanine-863, Cytarabine, DSM, Debrisoquine, Desipramine, Dinoprostone, Disopyramide, Dopamine, Epinephrine, Ethanaminium, 2-[(4-azidobenzoyl)amino]-N,N-diethyl-N-methyl-, Ethidium bromide, Fenoterol, Fexofenadine, Formoterol, Furentomin, Glycopyrrolate, Guanidine, Indacaterol, Ipratropium, Lamivudine, N-methyl-quinidine, N-methylnicotinamide, N1-methylnicotinamide, Nizatidine, Norepinephrine, Oxyphenonium, Perphenazine, Phenformin, Phenoxybenzamine, Prazosin, Procainamide, Procaterol, Quinidine, Quinine, S-Atenolol, Salbutamol, Serotonin, Spermidine, Terbutaline, Tetramethylammonium, Tetrapropylammonium, Tiotropium, Triamterene, Tributylmethylammonium, Trimethoprim, Tubocurarine, Tyramine, Varenicline, Verapamil, Zidovudine, berberine, cladribine, ethyl N-[4-carbamoyl-5-(carbamoylamino)-1,3-thiazol-2-yl]carbamate, ganciclovir, imatinib, lamotrigine, methyl (2S)-4-{2-[4-carbamoyl-5-(carbamoylamino)thiophen-2-yl]phenoxy}pyrrolidine-2-carboxylate, methyl 4-[4-carbamoyl-5-(carbamoylamino)thiophen-2-yl]-3-hydroxybenzoate, metoclopramide, prostaglandin F2alpha, sepantronium bromide, sitagliptin, sumatriptan, 5-hydroxytryptamine, choline, desipramine, metformin, tetraethylammonium, (R)-Xamoterol, 1'-Methylquinidinium, 1-Methylpyridinium cation, Bamet-UD2, DASPMI, N-(4,4-azo-n-pentyl)-21-deoxyajmalinium, N-(4,4-azo-n-pentyl)-quinuclidine, N-methyl-quinine, O-demethly tramadol, R-atenolol, phencycladine, Furamidine, Ganciclovir, N-methylpyridinium, Prostaglandin E2, Prostaglandin F2alpha, YM155, MPP+, PGE2, PGF2alpha, aciclovir, 4-(4-dimethylamino)styryl-N-methylpyridinium , Cisplatin, organic cations | |
| SLCO1B1 | 130 | Atorvastatin, Bilirubin, Fluo-3, Fluvastatin, Glycoursodeoxycholate, Methotr- | UCSF-FDA, Metrabase, |

| | | | |
|---|---|---|---|
| | | exate, Bilirubin monoglucuronide, Olmesartan, Pravastatin, Rosuvastatin, 17beta-estradiol 17beta-D-glucuronide, 8-fluorescein-cAMP, Atrasentan, BDE 47, BQ-123, Bamet-R2, Benzenesulfonamide, 4-(2-hydroxy-1,1-dimethylethyl)-N-[6-(2-hydroxyethoxy)-5-(2-methoxyphenoxy)[2,2'-bipyrimidin]-4-yl]-, Bromosulfophthalein, Caspofungin, Cefazoline, Cerivastatin, Cholecystokinin 8, Cholic acid, Darunavir, Demethylphalloin, Dinoprostone, Enalapril, Estrone-3-sulfate, Etoposide, Ezetimibe glucuronide, Fluorescein, Gimatecan, Glycochenodeoxycholic acid, Glycocholate, Hydroxyurea, Karenitecin, Leukotrien C4, Levothyroxine, Liothyronine, Liotrix, Lopinavir, Mebrofenin, Nafcillin, Nateglinide, Nodularin, Paclitaxel, Pazopanib, Phalloidin, Primovist, Rifampin, SN-38, Simvastatin, Sirolimus, TR-14035, Taurochenodeoxycholic acid, Taurocholic acid, Temocapril, Thromboxane B2, Torasemide, YM758, [D-Ala2, D-Leu5]-enkephalin, bilirubin bisglucuronoside, bosentan, cefditoren, cefoperazone, cholyl lysyl fluorescein, dehydroepiandrosterone sulfate, flavopiridol, leukotriene E4, lovastatin, microcystin LF, microcystin LW, microcystin-LR, penicillin G, pitavastatin, repaglinide, saquinavir, taurolithocholic acid 3-sulfate, triiodothyronine sulfate, valsartan, bilirubin, bromsulphthalein, fexofenadine, rifampicin, ACU154, BDE153, BDE99, Bamet-UD2, Bromosulfophthalein disodium, Cholylglycylamidofluorescein, Ethinylestradiol-3-O-sulfate, Fluorescein methotrexate, Mycophenolic acid-O-glucuronide, S8921G, Troglitazone sulfate, [D-Pen2,D-Pen5]-enkephalin, chenodeoxycholyl-(Nepsilon-NBD)-lysine, flavopiridol glucuronide, microcystein RR, Estradiol-17beta-glucuronide, Bilirubin diglucuronide, Bosentan, Bromsulphthalein, Cholate, Dehydroepiandrosterone sulfate, Estrone 3-sulfate, Pitavastatin, Rifampicin, Taurour-sodeoxycholate, Valsartan, opioids, thyroid hormones, bile salts, Cholyl taurine, | IUPHAR, Bioparadigms |

| | | | |
|---|---|---|---|
| | | Cholyl-glycylamido-fluorescein, Arsenite, arsenate, beta-lactam antibiotics, ACE inhibitors, HIV protease inhibitors, anticancer drugs, antifungals, bile acid derivatives and conjugates, bile acids, endothelin receptor antagonists, leukotrienes, sartans, statins, steroid conjugates, organic anions | |
| SLCO1B3 | 96 | Bilirubin, Cholecystokinin 8, Digoxin, Fexofenadine, Fluo-3, Fluvastatin, Glycoursodeoxycholate, Methotrexate, Bilirubin monoglucuronide, Olmesartan, Paclitaxel, Rosuvastatin, Telmisartan glucuronide, (-)-epicatechin-3-gallate, (-)-epigallocatechin gallate, 17beta-estradiol 17beta-D-glucuronide, 8-fluorescein-cAMP, Atorvastatin, Atrasentan, BQ-123, Benzenesulfonamide, 4-(2-hydroxy-1,1-dimethylethyl)-N-[6-(2-hydroxyethoxy)-5-(2-methoxyphenoxy)[2,2'-bipyrimidin]-4-yl]-, Bromosulfophthalein, Deltorphin II, Demethylphalloin, Diclofenac, Docetaxel, Erythromycin, Estrone-3-sulfate, Etoposide, Fluorescein, Glycocholate, Hydroxyurea, Indocyanine green, Leukotrien C4, Levothyroxine, Liothyronine, Mebrofenin, Ouabain, Primovist, Rifampin, SN-38, Taurocholic acid, Warfarin, alpha-amanitin, bosentan, dehydroepiandrosterone sulfate, imatinib, microcystin LF, microcystin LW, microcystin-LR, penicillin G, pitavastatin, telmisartan, testosterone, valsartan, CCK-8, amanitin, bilirubin, bromsulphthalein, digoxin, erythromycin, fexofenadine, ouabain, phalloidin, rifampicin, saquinavir, Bromosulfophthalein disodium, Fluorescein methotrexate, [D-Pen2,D-Pen5]-enkephalin, chenodeoxycholyl-(Nepsilon-NBD)-lysine, microcystein RR, Estradiol-17beta-glucuronide, Amanitin, Bosentan, Bromsulphthalein, Estrone 3-sulfate, Pitavastatin, Rifampicin, Taurocholate, Tauroursodeoxycholate, Telmisartan, Triiodothyronine, Valsartan, LTC4, opioids, thyroid hormones, bile salts, Cholyl-glycylamido-fluorescein, beta-lactam antibiotics, anticancer drugs, bile acid derivatives and | UCSF-FDA, Metrabase, IUPHAR, Bioparadigms |

| | | | |
|---|---|---|---|
| | | conjugates, bile acids, sartans, statins, steroid conjugates, organic anions | |
| SLCO1A2 | 83 | Fexofenadine, Methotrexate, Ouabain, Rosuvastatin, (-)-epicatechin-3-gallate, (-)-epigallocatechin gallate, 17beta-estradiol 17beta-D-glucuronide, Acebutolol, Atenolol, Atorvastatin, Atrasentan, BQ-123, Bamet-R2, Bromosulfophthalein, Celiprolol, Chlorambucil, Cholic acid, Darunavir, Deltorphin II, Dinoprostone, Enoxacin, Estrone-3-sulfate, Gatifloxacin, Glycocholate, Grepafloxacin, Hydroxyurea, Labetalol, Levothyroxine, Liothyronine, Lomefloxacin, N-methylquinidine, Nadolol, Norfloxacin, Ochratoxin A, P-aminohippuric acid, Rhodamine 123, Rocuronium, Sotalol, TR-14035, Talinolol, Taurocholic acid, Tauroursodeoxycholic acid, ciprofloxacin, dehydroepiandrosterone sulfate, levofloxacin, microcystin-LR, pitavastatin, reverse triiodothyronine, saquinavir, tebipenem pivoxil, bilirubin, bromsulphthalein, deltorphin II, fexofenadine, ouabain, rosuvastatin, talinolol, Bamet-UD2, Bromosulfophthalein disodium, Gd-B 20790, N-methyl-quinine, [D-Pen2,D-Pen5]-enkephalin, unoprostone carboxylate, Bromsulphthalein, Dehydroepiandrosterone sulfate, Estrone 3-sulfate, Levofloxacin, Pitavastatin, Saquinavir, Taurocholate, Tauroursodeoxycholate, Thyroxine, PGE2, thyroid hormones, bile salts, HIV protease inhibitors, antibiotics, anticancer drugs, beta blockers, bile acids, fluoroquinolones, steroid conjugates, organic anions and cations | UCSF-FDA, Metrabase, IUPHAR, Bioparadigms |
| SLCO2B1 | 65 | Atorvastatin, Fexofenadine, Fluvastatin, Pravastatin, Rosuvastatin, Telmisartan glucuronide, (betaR,deltaR)-5-cyclopropyl-2-(4-fluorophenyl)-4-[[[(3-fluorophenyl)methyl]amino]carbonyl]-beta,delta-dihydroxy-1H-imidazole-1-heptanoic acid, 3-Pyridinecarboxylic acid, 6-[[(2S)-3-cyclopentyl-1-oxo-2-[4-(trifluoromethyl)-1H-imidazol-1-yl]propyl]amino]-, BDE 47, Bromosulfophthalein, Cerivastatin, Digoxin, Dinoprostone, Eltrombopag, Estrone-3-sulfate, Etoposide, E- | UCSF-FDA, Metrabase, IUPHAR, Bioparadigms |

| | | | |
|---|---|---|---|
| | | zetimibe glucuronide, Iloprost, Latanoprost acid, Levothyroxine, M17055, Montelukast, N-methyl-quinidine, Pregnenolone sulfate, Rocuronium, Scutellarein 7-O-beta-D-glucuronide, Sulfasalazine, Talinolol, Taurocholic acid, Thromboxane B2, aliskiren, bosentan, dehydroepiandrosterone sulfate, glyburide, pemetrexed, penicillin G, pitavastatin, tebipenem pivoxil, telmisartan, amiodarone, bromsulphthalein, dehydroepiandrosterone sulphate, estrone-3-sulphate, fexofenadine, glibenclamide, talinolol, BDE153, BDE99, Bromosulfophthalein disodium, CP 671305, Ethinylestradiol-3-O-sulfate, Methanone, [4-[4-(3,4-dihydro-6,7-dimethoxy-2(1H)-isoquinolinyl)-6,7-dimethoxy-2-quinazolinyl]hexahydro-1H-1,4-diazepin-1-yl](3,4-dihydroxy-1-pyrrolidinyl), Salozinal, Scutellarein 6-O-beta-D-glucuronide, unoprostone carboxylate, Bosentan, Bromsulphthalein, Estrone 3-sulfate, Glyburide, Pitavastatin, Taurocholate, T4, DHEAS, statins, E-3-S | |
| SLC10A2 | 58 | Glycodeoxycholate, Glycoursodeoxycholate, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(2-fluoropyridin-3-yl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(2-hydroxyphenyl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(2-methylpyridin-3-yl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(2-methylpyridin-4-yl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4- | UCSF-FDA, Metrabase, Biopara-digms |

| | | | |
|---|---|---|---|
| | | [(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(3-hydroxyphenyl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(3-hydroxypyridin-2-yl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(4,6-dimethylpyridin-2-yl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(4-hydroxyphenyl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(5-fluoropyridin-3-yl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(5-methylpyridin-3-yl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(6-fluoropyridin-3-yl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(6-methoxypyridin-3-yl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetra- | |

| | | | |
|---|---|---|---|
| | | cyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(6-methylpyridin-3-yl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(pyridin-2-yl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(pyridin-3-yl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-[(pyridin-4-yl)carbamoyl]butanoic acid, (2S)-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]-4-{[2-(pyridin-2-yl)ethyl]carbamoyl}butanoic acid, (2S)-4-[(5-bromopyridin-3-yl)carbamoyl]-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]butanoic acid, (2S)-4-[(6-aminopyridin-2-yl)carbamoyl]-2-[(4R)-4-[(1S,2S,5R,7S,9R,10R,11S,14R,15R)-5,9-dihydroxy-2,15-dimethyltetracyclo[8.7.0.0^{2,7}.0^{11,15}]heptadecan-14-yl]pentanamido]butanoic acid, Benzoic acid, 2-[[(4S)-4-carboxy-4-[[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]amino]-1-oxobutyl]amino]-, 1-ethyl ester, Benzoic acid, 4-[[(4S)-4-carboxy-4-[[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]amino]-1-oxobutyl]amino]-, 1-ethyl ester, Cholic acid, Glycine, N-[(3alpha,5beta)-3-hydroxy-7,24-dioxocholan-24-yl]-, Glycine, N-[(3alpha,5beta,7alpha)-7-hydroxy-24-oxo-3-[(3-pyridinylcarbonyl)oxy]cholan- | |

| | | | |
|---|---|---|---|
| | | 24-yl]-, Glycochenodeoxycholic acid, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-(2,3-dimethylphenyl)-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-(2,6-dimethylphenyl)-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-(2-fluorophenyl)-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-(2-methoxyphenyl)-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-(2-methylphenyl)-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-(3,5-dimethoxyphenyl)-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-(3-fluorophenyl)-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-(3-methoxyphenyl)-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-(3-methylphenyl)-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-(4-fluorophenyl)-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-(4-methoxyphenyl)-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-(4-methylphenyl)-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-[3-[[(1,1-dimethylethoxy)carbonyl]amino]phenyl]-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-[4-(1,1-dimethylethyl)phenyl]-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-[4-(trifluoromethyl)phenyl]-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-[4-[[(1,1-dimethylethoxy)carbonyl]amino]phenyl]-, L-Glutamine, N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-N5-phenyl-, L-Glutamine, N5-(2,4-difluorophenyl)-N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-, L- | |

| | | | |
|---|---|---|---|
| | | Glutamine, N5-(2-aminophenyl)-N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-, L-Glutamine, N5-(3-aminophenyl)-N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-, L-Glutamine, N5-(4-aminophenyl)-N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-, L-Glutamine, N5-(4-chlorophenyl)-N2-[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]-, Taurocholic acid, Acyclovir valylchenodeoxycholic acid, Benzoic acid, 2-[[(4S)-4-carboxy-4-[[(3alpha,5beta,7alpha)-3,7-dihydroxy-24-oxocholan-24-yl]amino]-1-oxobutyl]amino]-, 1-methyl ester, Glycine,N-[(5-beta)-1'-[(3-hydroxyphenyl)methyl]-24-oxo-1'H-cholano[3,2-c]pyrazol-24-yl]-, Glycine,N-[(5-beta)-1'-[2-[[(3,4-dihydro-6,7-dimethoxy-2(1H)-isoquinolinyl)carbonyl]oxy]ethyl]-24-oxo-1'H-cholano[3,2-c]pyrazol-24-yl]-, Cholate, Glycochenodeoxycholate, Taurocholate, bile acids | |
| SLC22A2 | 30 | Amiloride, Cimetidine, Dopamine, Epinephrine, Lamivudine, Metformin, Norepinephrine, Ranitidine, Serotonin, Tetraethylammonium, Varenicline, dopamine, histamine, metformin, pancuronium, tetraethylammonium, tubocurarine, Amantadine, Famotidine, Histamine, Memantine, N-methylpyridinium, Prostaglandin E2, Prostaglandin F2alpha, YM155, MPP+, PGE2, 4-(4-dimethylamino)styryl-N-methylpyridinium , cisplatin, organic cations | UCSF-FDA, IUPHAR, Bioparadigms |
| SLC22A8 | 30 | Cefaclor, Cimetidine, Fexofenadine, Furosemide, Methotrexate, Ochratoxin A, Olmesartan, Pravastatin, Rosuvastatin, Zidovudine, aminohippuric acid, cimetidine, estrone-3-sulphate, ochratoxin A, tenofovir, uric acid, Bumetanide, Ceftizoxime, Cortisol, Dimesna, Edaravone sulfate, Estrone 3-sulfate, Para-aminohippurate, Pitavastatin, Prostaglandin E2, Prostaglandin F2alpha, Sitagliptin, Tetracycline, Uric acid, organic anions | UCSF-FDA, IUPHAR, Bioparadigms |
| SLC36A1 | 25 | 5-aminolevulinic acid, D-alanine, D-cycloserine, D-cysteine, D-proline, D-serine, GABA, L-azetidine-2-carboxylate, L-cycloserine, MeAIB, THPO, arecaidine, betaine, gaboxadol, glycine, | IUPHAR, Bioparadigms |

| | | | |
|---|---|---|---|
| | | muscimol, sarcosine, taurine, trans-4-hydroxy-proline, vigabatrin, beta-alanine, beta-guanidinopropionic acid, L-alanine, Glycine, Proline | |
| SLC22A6 | 23 | Acyclovir, Methotrexate, Ochratoxin A, Olmesartan, Zidovudine, aminohippuric acid, tenofovir, uric acid, 6-carboxyfluorescein, Adefovir, Cidofovir, Dimesna, Edaravone sulfate, Ganciclovir, Glutarate, Para-aminohippurate, Prostaglandin E2, Prostaglandin F2alpha, Tenofovir, Uric acid, Zalcitabine, non-steroidal anti-inflammatory drugs, organic anions | UCSF-FDA, IUPHAR, Bioparadigms |
| SLC29A1 | 23 | 2-chloroadenosine, adenine, adenosine, atenolol, cladribine, cytarabine, cytidine, didanosine, floxuridine, gemcitabine, guanosine, hypoxanthine, inosine, pentostatin, ribavirin, thymidine, thymine, tubercidin, uridine, vidarabine, abacavir, formycin B, zalcitabine | IUPHAR |
| SLC47A1 | 23 | Acyclovir, Cimetidine, Guanidine, Metformin, Procainamide, Tetraethylammonium, cephalexin, cimetidine, creatine, metformin, paraquat, quinidine, thiamine, 1-methyl-4-phenylpyridinium, guanidine, procainamide, tetraethylammonium, Estrone 3-sulfate, Ganciclovir, N-methylpyridinium, Paraquat, Topotecan, cephradine | UCSF-FDA, IUPHAR, Bioparadigms |
| SLC47A2 | 21 | Acyclovir, Cimetidine, Guanidine, Metformin, Procainamide, Tetraethylammonium, N1-methylnicotinamide, cimetidine, creatine, guanidine, metformin, procainamide, thiamine, Estrone 3-sulfate, Ganciclovir, N-methylpyridinium, Topotecan, MPP+, aciclovir, MPP, TEA | UCSF-FDA, IUPHAR, Bioparadigms |
| SLC5A8 | 20 | 3-bromopyruvate, 5-aminosalicylate, D-lactic acid, L-lactic acid, acetoacetic acid, benzoate, butyric acid, dichloroacetate, nicotinic acid, propanoic acid, pyroglutamic acid, pyruvic acid, salicylic acid, alpha-ketoisocaproate, beta-D-hydroxybutyric acid, beta-L-hydroxybutyric acid, gamma-hydroxybutyric acid, 2-oxothiazolidine-4-carboxylate, acetic acid, short chain fatty acids | IUPHAR, Bioparadigms |
| SLC15A2 | 19 | 5-aminolevulinic acid, Amoxicillin, Bestatin, Cefaclor, Enalapril, alafosfalin, cefadroxil, cyclacillin, muramyl dipeptide, Glycylsarcosine, beta-Ala-Lys- | UCSF-FDA, IUPHAR, Bioparadigms |

| | | | |
|---|---|---|---|
| | | AMCA, gamma-iE-DAP, D-Ala-Lys-AMCA, dipeptides, protons, tripeptides, beta-lactam antibiotics, di-peptides, tri-peptides | |
| SLC29A2 | 18 | 2-chloroadenosine, adenine, adenosine, cladribine, cytarabine, cytosine, gemcitabine, guanine, guanosine, hypoxanthine, inosine, thymidine, thymine, tubercidin, uridine, vidarabine, zidovudine, formycin B | IUPHAR |
| SLC22A11 | 16 | Methotrexate, Ochratoxin A, Pravastatin, Zidovudine, dehydroepiandrosterone sulphate, estrone-3-sulphate, ochratoxin A, uric acid, Bumetanide, Dehydroepiandrosterone sulfate, Estrone 3-sulfate, Prostaglandin E2, Prostaglandin F2alpha, Tetracycline, Uric acid, organic anions | UCSF-FDA, IUPHAR, Bioparadigms |
| SLC16A1 | 15 | Atorvastatin, Butyrate, Gabapentin enacarbil, Glycolic acid, L-lactate, Pyruvic acid, Salicylic acid, doxorubicin, L-lactic acid, pyruvic acid, beta-D-hydroxybutyric acid, gamma-hydroxybutyric acid, lactate, pyruvate, ketone bodies | Metrabase, IUPHAR, Bioparadigms |
| SLC22A7 | 15 | 5-fluorouracil, Erythromycin, Paclitaxel, Zidovudine, aminohippuric acid, 2'-deoxyguanosine, Bumetanide, Cyclic GMP, Prostaglandin E2, Prostaglandin F2alpha, Tetracycline, Theophylline, PGE2, non-steroidal anti-inflammatory drugs, organic anions | UCSF-FDA, IUPHAR, Bioparadigms |
| SLC25A2 | 15 | D-arginine, D-citrulline, D-histidine, D-lysine, D-ornithine, L-arginine, L-citrulline, L-histidine, L-lysine, L-ornithine, arginine, citrulline, histidine, lysine, ornithine | IUPHAR, Bioparadigms |
| SLC28A3 | 15 | 5-fluorouridine, adenosine, cladribine, cytidine, didanosine, floxuridine, gemcitabine, guanosine, inosine, thymidine, uridine, zebularine, zidovudine, formycin B, zalcitabine | IUPHAR |
| SLC29A3 | 15 | adenine, adenosine, cladribine, cordycepin, didanosine, floxuridine, fludarabine, guanosine, inosine, thymidine, tubercidin, uridine, zebularine, zidovudine, zalcitabine | IUPHAR |
| SLC18A1 | 14 | (-)-adrenaline, (-)-noradrenaline, 5-hydroxytryptamine, MDMA, dexamfeta- | IUPHAR, Bioparadigms |

| | | | |
|---|---|---|---|
| | | mine, dopamine, fenfluramine, histamine, beta-phenylethylamine, MPP+, 5-HT, DA, NE, epinephrine | |
| SLC18A2 | 14 | (-)-adrenaline, (-)-noradrenaline, 5-hydroxytryptamine, MDMA, dexamfetamine, dopamine, fenfluramine, histamine, beta-phenylethylamine, MPP+, 5-HT, DA, NE, epinephrine | IUPHAR, Bioparadigms |
| SLC22A3 | 14 | Epinephrine, Metformin, Norepinephrine, (-)-noradrenaline, 5-hydroxytryptamine, dopamine, metformin, quinidine, tetraethylammonium, Etilefrine, Histamine, N-methylpyridinium, MPP+, organic cations | UCSF-FDA, IUPHAR, Bioparadigms |
| SLC22A4 | 13 | Ipratropium, Quinidine, Tetraethylammonium, Verapamil, L-carnitine, mepyramine, tetraethylammonium, verapamil, Ergothioneine, MPP+, ergothioneine, organic cations, zwitterions | UCSF-FDA, IUPHAR, Bioparadigms |
| SLC22A5 | 13 | Ipratropium, L-carnitine, Quinidine, Verapamil, acetyl-L-carnitine, mepyramine, tetraethylammonium, verapamil, Acetyl-L-carnitine, D-carnitine, MPP+, organic cations, zwitterions | UCSF-FDA, IUPHAR, Bioparadigms |
| SLC36A2 | 13 | D-cycloserine, L-azetidine-2-carboxylate, L-cycloserine, L-proline, MeAIB, glycine, sarcosine, trans-4-hydroxy-proline, L-alanine, Alanine, Glycine, Proline, hydroxyproline | IUPHAR, Bioparadigms |
| SLC10A1 | 12 | Glycoursodeoxycholate, Rosuvastatin, dehydroepiandrosterone sulphate, estrone-3-sulphate, Cholate, Estrone 3-sulfate, Pitavastatin, Taurocholate, Tauroursodeoxycholate, T4, iodothyronine sulphates, bile acids | UCSF-FDA, IUPHAR, Bioparadigms |
| SLC15A4 | 12 | His-Leu-lopinavir, L-histidine, MDP-rhodamine, carnosine, glycyl-sarcosine, muramyl dipeptide, valacyclovir, histidine, C12-iE-DAP, Tri-DAP, di- and tripeptides, protons | IUPHAR, Bioparadigms |
| SLC19A1 | 11 | folic acid, folinic acid, methotrexate, tetrahydrofolic acid, thiamine monophosphate, N5-formyltetrahydrofolate, N5-methylfolate, Organic phosphates; in particular, adenine nucleotides, Other tetrahydrofolate-cofactors, antifolates, reduced folates | IUPHAR, Bioparadigms |
| SLC2A1 | 11 | 2-Deoxyglucose, 3-O-methyl-D-Glucose, D-glucose, Dehydroascorbic acid, D-glucosamine, dehydroascorbic acid, | Metrabase, IUPHAR, Bioparadigms |

| | | | |
|---|---|---|---|
| | | 2-deoxy-d-glucose (acyclic), galactose, glucosamine, glucose, mannose | |
| SLCO4A1 | 11 | 17beta-estradiol 17beta-D-glucuronide, Dinoprostone, Estrone-3-sulfate, Levothyroxine, Liothyronine, Taurocholic acid, penicillin G, thyroid hormones, bile acids, prostaglandins, steroid conjugates | Metrabase, IUPHAR |
| SLC15A3 | 10 | L-histidine, MDP-rhodamine, histidyl-leucine, muramyl dipeptide, histidine, Tri-DAP, dipeptides, protons, tripeptides, di- and tri-peptides | IUPHAR, Bioparadigms |
| SLC25A17 | 10 | ADP, ATP, adenosine 5'-monophosphate, AMP, CoA, FAD, FMN, NAD+, PAP, dPCoA | IUPHAR, Bioparadigms |
| SLC38A2 | 10 | MeAIB, Alanine, Asparagine, Cysteine, Glutamine, Glycine, Histidine, Methionine, Proline, Serine | IUPHAR, Bioparadigms |
| SLC43A1 | 10 | L-isoleucine, L-leucine, L-leucinol, L-methionine, L-phenylalanine, L-phenylalaninol, L-valine, L-valinol, L-BCAAs, amino alcohols | IUPHAR, Bioparadigms |
| SLC11A2 | 9 | Cd2+, Co2+, Cu2+, Fe2+, Mn2+, Cu1+, Ni2+, Pb2+, Zn2+ | IUPHAR, Bioparadigms |
| SLC25A10 | 9 | malic acid, succinic acid, malate, succinate, S2O32-, SO42-, phosphate, sulphate, thiosulphate | IUPHAR, Bioparadigms |
| SLC28A2 | 9 | adenosine, cladribine, didanosine, fludarabine, guanosine, inosine, thymidine, vidarabine, formycin B | IUPHAR |
| SLC29A4 | 9 | 5-hydroxytryptamine, adenosine, atenolol, dopamine, histamine, metformin, tetraethylammonium, tyramine, MPP+ | IUPHAR |
| SLC43A2 | 9 | L-isoleucine, L-leucine, L-leucinol, L-methionine, L-phenylalanine, L-valine, L-valinol, L-BCAAs, amino alcohols | IUPHAR, Bioparadigms |
| SLCO2A1 | 9 | Dinoprostone, Prostaglandin D2, Thromboxane B2, alprostadil, prostaglandin F2alpha, lactate, eicosanoids, prostaglandins, synthetic prostaglandin derivatives | Metrabase, Bioparadigms, IUPHAR |
| SLCO3A1 | 9 | Dinoprostone, Estrone-3-sulfate, alprostadil, penicillin G, prostaglandin F2alpha, BQ123, vasopressin, thyroid hormones, prostaglandins | Metrabase, IUPHAR |
| SLC17A1 | 8 | penicillin G, probenecid, uric acid, Cl-, chloride, organic acids, phosphate, organic anions | IUPHAR, Bioparadigms |

| | | | |
|---|---|---|---|
| SLC1A1 | 8 | L-aspartic acid, L-cysteine, L-glutamic acid, L-trans-2,4-pyrolidine dicarboxylate, D-aspartic acid, DL-threo-beta-hydroxyaspartate, L-Glu, D/L-Asp | IUPHAR, Bioparadigms |
| SLC25A15 | 8 | L-arginine, L-citrulline, L-lysine, L-ornithine, arginine, citrulline, lysine, ornithine | IUPHAR, Bioparadigms |
| SLC28A1 | 8 | adenosine, cytidine, gemcitabine, ribavirin, thymidine, uridine, zidovudine, zalcitabine | IUPHAR |
| SLC41A1 | 8 | Ba2+, Cd2+, Co2+, Cu2+, Fe2+, Mg2+, Sr2+, Zn2+ | IUPHAR, Bioparadigms |
| SLC5A4 | 8 | 1-deoxynojirimycin, 1-deoxynojirimycin-1-sulfonic acid, D-glucose, N-ethyl-1-deoxynojirimycin, miglitol, miglustat, H+, Na+ | IUPHAR, Bioparadigms |
| SLC63A2 | 8 | dihydrosphingosine-1-phosphate, sphingosine-1-phosphate, Phosphorylated fingolimod (FTY720-P), Phosphorylated sphingolipids, Sphingosine-1-phosphate (S1P), dihydrosphingosine-1-phosphate (DH-S1P), phyto-S1P, C17-S1P, C17-S1P, phosphorylated Fingolimod, phyto-S1P | Bioparadigms, IUPHAR |
| SLC10A6 | 7 | dehydroepiandrosterone sulphate, estrone-3-sulphate, pregnenolone sulphate, taurolithocholic acid-3-sulphate, dehydroepiandrosterone sulfate, estrone-3-sulfate, pregnenolone sulfate | IUPHAR, Bioparadigms |
| SLC13A3 | 7 | citric acid, succinic acid, NALA, citrate, glutarate, ketoglutarate, succinate | IUPHAR, Bioparadigms |
| SLC16A10 | 7 | L-phenylalanine, L-tryptophan, L-tyrosine, levodopa, T3, T4, aromatic amino acids | IUPHAR, Bioparadigms |
| SLC17A5 | 7 | L-aspartic acid, L-glutamic acid, L-lactic acid, gluconate, glucuronic acid, sialic acid, other acidic sugars | IUPHAR, Bioparadigms |
| SLC1A2 | 7 | L-aspartic acid, L-glutamic acid, L-trans-2,4-pyrolidine dicarboxylate, D-aspartic acid, DL-threo-beta-hydroxyaspartate, L-Glu, D/L-Asp | IUPHAR, Bioparadigms |
| SLC1A3 | 7 | L-aspartic acid, L-glutamic acid, L-trans-2,4-pyrolidine dicarboxylate, D-aspartic acid, DL-threo-beta-hydroxyaspartate, L-Glu, D/L-Asp | IUPHAR, Bioparadigms |
| SLC1A6 | 7 | L-aspartic acid, L-glutamic acid, L-trans-2,4-pyrolidine dicarboxylate, D-aspartic acid, DL-threo-beta-hydroxyaspartate, L-Glu, D/L-Asp | IUPHAR, Bioparadigms |

| | | | |
|---|---|---|---|
| SLC1A7 | 7 | L-aspartic acid, L-glutamic acid, L-trans-2,4-pyrolidine dicarboxylate, D-aspartic acid, DL-threo-beta-hydroxyaspartate, L-Glu, D/L-Asp | IUPHAR, Bioparadigms |
| SLC25A1 | 7 | citric acid, malic acid, phosphoenolpyruvic acid, PEP, citrate, isocitrate, malate | IUPHAR, Bioparadigms |
| SLC25A19 | 7 | thiamine monophosphate, thiamine pyrophosphate, Deoxynucleotide Diphosphates (dNDPs), Deoxynucleotide Triphosphates (dNTPs), Dideoxynucleotide Triphosphates (ddNTPs), Nucleotide Diphosphates (NDPs), deoxynucleotides | Bioparadigms, IUPHAR |
| SLC26A6 | 7 | formate, Cl-, HCO3-, I-, OH-, SO42-, oxalate | IUPHAR, Bioparadigms |
| SLC2A2 | 7 | D-glucosamine, D-glucose, fructose, galactose, glucosamine, glucose, mannose | IUPHAR, Bioparadigms |
| SLC38A1 | 7 | MeAIB, Alanine, Asparagine, Cysteine, Glutamine, Histidine, Serine | IUPHAR, Bioparadigms |
| SLC38A4 | 7 | MeAIB, Alanine, Asparagine, Cysteine, Glycine, Serine, Threonine | IUPHAR, Bioparadigms |
| SLC38A9 | 7 | Arginine, Glutamic acid, Glutamine, Histidine, Leucine, Lysine, Proline | Bioparadigms |
| SLC39A14 | 7 | Cd2+, Fe2+, Mn2+, Cd, Fe, Mn, Zn | IUPHAR, Bioparadigms |
| SLC46A1 | 7 | folic acid, methotrexate, pemetrexed, N-formyltetrahydrofolate, N5-methyltetrafolate, Reduced folates, antifolates | IUPHAR, Bioparadigms |
| SLC5A1 | 7 | D-galactose, D-glucose, urea, alpha-MDG, galactose, glucose, water | IUPHAR, Bioparadigms |
| SLC6A14 | 7 | 1-methyltryptophan, BCH, valganciclovir, beta-alanine, zwitterionic or cationic NOS inhibitors, cationic amino acids, neutral | IUPHAR, Bioparadigms |
| SLC6A2 | 7 | (-)-adrenaline, (-)-noradrenaline, amphetamine, dopamine, methamphetamine, MPP+, norepinephrine | IUPHAR, Bioparadigms |
| SLC7A5 | 7 | L-leucine, BCH, 3-fluoro-L-alpha-methyl-tyrosine, L-DOPA, T3, T4, large neutral L-amino acids | Metrabase, Bioparadigms |
| SLCO1C1 | 7 | bromsulphthalein, thyroid hormones, T3, T4, rT3, statins, steroid conjugates | IUPHAR, Bioparadigms |
| SLC13A1 | 6 | S2O32-, SO42-, SeO42-, selenate, sulfate, thiosulfate | IUPHAR, Bioparadigms |
| SLC14A1 | 6 | acetamide, acrylamide, ammonium carbonate, methylurea, urea, formamide | IUPHAR |
| SLC18A3 | 6 | N-(4'-pentanonyl)-4-(4''-dimethylaminostyryl)pyridinium, acetylcholine, choline, | IUPHAR, Bioparadigms |

| | | | |
|---|---|---|---|
| | | ethidium, N-methyl-pyridinium-2-aldoxime, TPP+ | |
| SLC1A5 | 6 | L-Ala, L-Asn, L-Cys, L-Gln, L-Ser, L-Thr | Biopara-digms |
| SLC22A12 | 6 | orotic acid, uric acid, Oxypurinol, Uric acid, urate, organic anions | IUPHAR, UCSF-FDA, Biopara-digms |
| SLC25A42 | 6 | ADP, ATP, CoA, 5'-diphosphate, adenosine 3', dPCoA | IUPHAR, Bioparadigms |
| SLC41A2 | 6 | Ba2+, Co2+, Fe2+, Mg2+, Mn2+, Ni2+ | IUPHAR, Bioparadigms |
| SLC57A3 | 6 | Ba2+, Cu2+, Fe2+, Mg2+, Sr2+, Mg2+, Sr2+, Ba2+, Fe2+, Cu2+ | Biopara-digms, IUPHAR |
| SLC5A5 | 6 | I-, SCN-, Br-, ClO4-, NO3-, pertechnetate | IUPHAR, Bioparadigms |
| SLC5A9 | 6 | D-glucose, D-mannose, alpha-MDG, fructose, glucose, mannose | IUPHAR, Bioparadigms |
| SLC6A3 | 6 | (-)-adrenaline, (-)-noradrenaline, amphetamine, dopamine, methamphetamine, MPP+ | IUPHAR, Bioparadigms |
| SLCO4C1 | 6 | cyclic AMP, thyroid hormones, anticancer drugs, cardiac glycosides, dipeptidyl peptidase-4 inhibitors, steroid conjugates | IUPHAR |
| SLC13A2 | 5 | citric acid, succinic acid, citrate, ketoglutarate, succinate | IUPHAR, Bioparadigms |
| SLC13A5 | 5 | citric acid, pyruvic acid, citrate, pyruvate, succinate | IUPHAR, Bioparadigms |
| SLC16A2 | 5 | triiodothyronine, T4, T2, T3, rT3 | IUPHAR, Bioparadigms |
| SLC16A7 | 5 | L-lactic acid, pyruvic acid, lactate, pyruvate, ketone bodies | IUPHAR, Bioparadigms |
| SLC22A32 | 5 | diclofenac, etodolac, indomethacin, mefenamic acid, p-aminohippuric acid | Biopara-digms |
| SLC25A12 | 5 | 2-amino-3-sulfinopropanoic acid, L-aspartic acid, L-glutamic acid, aspartate, glutamate | IUPHAR, Bioparadigms |
| SLC25A13 | 5 | 2-amino-3-sulfinopropanoic acid, L-aspartic acid, L-glutamic acid, aspartate, glutamate | IUPHAR, Bioparadigms |
| SLC25A23 | 5 | ADP, ATP, AMP, ATP-Mg2+, Pi | Biopara-digms |
| SLC25A24 | 5 | ADP, ATP, AMP, ATP-Mg2+, Pi | Biopara-digms |
| SLC26A11 | 5 | Cl-, HCO3-, HSO4-, SO42-, oxalate | Biopara-digms, IUPHAR |

| | | | |
|---|---|---|---|
| SLC26A4 | 5 | formate, Cl-, HCO3-, I-, OH- | IUPHAR, Bioparadigms |
| SLC26A5 | 5 | formate, Cl-, HCO3-, SO42-, oxalate | Biopara-digms, IUPHAR |
| SLC26A7 | 5 | Cl-, HCO3-, OH-, Ch: Cl-, SO42- | Biopara-digms |
| SLC26A8 | 5 | Cl-, HCO3-, OH-, SO42-, oxalate | IUPHAR, Bioparadigms |
| SLC2A13 | 5 | *muco*-inositol, *myo*-inositol, *scyllo*-inositol, D-*chiro*-inositol, myo-inositol | IUPHAR, Bioparadigms |
| SLC2A3 | 5 | D-glucose, galactose, glucose, mannose, xylose | IUPHAR, Bioparadigms |
| SLC2A9 | 5 | D-fructose, uric acid, fructose, glucose, urate | IUPHAR, Bioparadigms |
| SLC32A1 | 5 | GABA, glycine, beta-alanine, gamma-hydroxybutyric acid, GABA / glycine | IUPHAR, Bioparadigms |
| SLC38A3 | 5 | MeAIB, Alanine, Asparagine, Glutamine, Histidine | IUPHAR, Bioparadigms |
| SLC38A5 | 5 | MeAIB, Asparagine, Glutamine, Histidine, Serine | IUPHAR, Bioparadigms |
| SLC38A7 | 5 | Alanine, Asparagine, Glutamine, Histidine, Serine | Biopara-digms |
| SLC45A1 | 5 | L-glucose, Galactose, galactose, glucose, sucrose | IUPHAR, Bioparadigms |
| SLC57A1 | 5 | Mg2+, Co2+, Fe2+, Sr2+, Sr2+, Fe2+ and Co2+ to a lesser extent | IUPHAR, Bioparadigms |
| SLC58A2 | 5 | Cu2+, Fe2+, Mg2+, Mn2+, Mg2+, Fe2+, Cu2+, Mn2+ | Biopara-digms, IUPHAR |
| SLC5A10 | 5 | D-galactose, D-glucose, fructose, glucose, mannose | IUPHAR, Bioparadigms |
| SLC5A6 | 5 | biotin, lipoic acid, pantothenic acid, I-, lipoate panthothenate | IUPHAR, Bioparadigms |
| SLC7A1 | 5 | L-arginine, L-histidine, L-lysine, L-ornithine, cationic L-amino acids | IUPHAR, Bioparadigms |
| SLC7A2 | 5 | L-arginine, L-histidine, L-lysine, L-ornithine, cationic L-amino acids | IUPHAR, Bioparadigms |
| SLC9A7 | 5 | H+, K+, Li+, NH4+, Na+ | Biopara-digms |
| SLC12A8 | 4 | L-aspartic acid, L-glutamic acid, spermidine, spermine | IUPHAR |
| SLC13A4 | 4 | chromium, SO42-, oxyanions selenium, sulfate | Biopara-digms, IUPHAR |
| SLC16A3 | 4 | L-lactic acid, pyruvic acid, lactate, ketone bodies | IUPHAR, Bioparadigms |
| SLC17A9 | 4 | ATP, guanosine 5'-diphosphate, guanosine-5'-triphosphate, purine nucleotides | IUPHAR, Bioparadigms |

| | | | |
|---|---|---|---|
| SLC1A4 | 4 | L-Ala, L-Cys, L-Ser, L-Thr | Biopara-digms |
| SLC20A1 | 4 | AsO43-, phosphate, inorganic phos-phate, monovalent | IUPHAR, Bi-oparadigms |
| SLC22B1 | 4 | Galactose, brivaracetam, levetiracetam, selectracetam | IUPHAR, Bi-oparadigms |
| SLC25A11 | 4 | malic acid, alpha-ketoglutaric acid, 2-oxoglutarate, malate | IUPHAR, Bi-oparadigms |
| SLC25A21 | 4 | alpha-ketoglutaric acid, alpha-oxoadipic acid, oxoadipate, oxoglutarate | IUPHAR, Bi-oparadigms |
| SLC26A9 | 4 | Cl-, HCO3, HCO3-, Ch: Cl- | Biopara-digms |
| SLC2A10 | 4 | D-glucose, dehydroascorbic acid, galac-tose, glucose | IUPHAR, Bi-oparadigms |
| SLC2A11 | 4 | D-fructose, D-glucose, fructose, glucose | IUPHAR, Bi-oparadigms |
| SLC2A4 | 4 | D-glucosamine, D-glucose, glucosa-mine, glucose | IUPHAR, Bi-oparadigms |
| SLC2A7 | 4 | D-fructose, D-glucose, fructose, glucose | IUPHAR, Bi-oparadigms |
| SLC2A8 | 4 | D-glucose, fructose, galactose, glucose | IUPHAR, Bi-oparadigms |
| SLC36A4 | 4 | L-proline, L-tryptophan, Proline, trypto-phan | IUPHAR, Bi-oparadigms |
| SLC38A8 | 4 | Alanine, Asparagine, Glutamine, Histi-dine | Biopara-digms |
| SLC39A8 | 4 | Cd2+, Cd, Mn, Zn | IUPHAR, Bi-oparadigms |
| SLC42A2 | 4 | NH3, NH4+, methyl amine, methyl am-monium | Biopara-digms |
| SLC4A11 | 4 | Cl-, sodium, NaHCO3-, borate | IUPHAR, Bi-oparadigms |
| SLC51A-SLC51B | 4 | Digoxin, Dinoprostone, Estrone-3-sul-fate, Taurocholic acid | Metrabase |
| SLC55A1 | 4 | Ca2+, H+, K+, Ca2+, K+, H+ | Biopara-digms, IUPHAR |
| SLC57A4 | 4 | Ba2+, Mg2+, Sr2+, Mg2+, Sr2+, Ba2+ | Biopara-digms, IUPHAR |
| SLC5A12 | 4 | L-lactic acid, nicotinic acid, pyruvic acid, short chain fatty acids | IUPHAR, Bi-oparadigms |
| SLC64A1 | 4 | Mn2+, Ca2+, H+, Ca2+, H+ | IUPHAR, Bi-oparadigms |
| SLC6A11 | 4 | GABA, guvacine, nipecotic acid, beta-alanine | IUPHAR, Bi-oparadigms |
| SLC6A13 | 4 | GABA, guvacine, nipecotic acid, beta-alanine | IUPHAR, Bi-oparadigms |
| SLC6A20 | 4 | L-proline, sarcosine, pipecolate, proline | IUPHAR, Bi-oparadigms |

| | | | |
|---|---|---|---|
| SLC6A4 | 4 | 5-hydroxytryptamine, MDMA, p-chloro-amphetamine, serotonin | IUPHAR, Bi-oparadigms |
| SLC7A3 | 4 | L-arginine, L-lysine, L-ornithine, cationic L-amino acids | IUPHAR, Bi-oparadigms |
| SLC7A8 | 4 | BCH, T3, T4, neutral L-amino acids | Biopara-digms |
| SLC9A1 | 4 | H+, Li+, NH4+, Na+ | Biopara-digms |
| SLC9A2 | 4 | H+, Li+, NH4+, Na+ | Biopara-digms |
| SLC9A3 | 4 | H+, Li+, NH4+, Na+ | Biopara-digms |
| SLC9A4 | 4 | H+, Li+, NH4+, Na+ | Biopara-digms |
| SLC9A5 | 4 | H+, Li+, NH4+, Na+ | Biopara-digms |
| SLC12A2 | 3 | Cl-, K+, Na+ | Biopara-digms |
| SLC16A5 | 3 | probenecid, bumetanide, nateglinide | Biopara-digms |
| SLC20A2 | 3 | phosphate, inorganic phosphate, mono-valent | IUPHAR, Bi-oparadigms |
| SLC24A1 | 3 | Ca2+, K+, Na+ | Biopara-digms |
| SLC24A2 | 3 | Ca2+, K+, Na+ | Biopara-digms |
| SLC24A3 | 3 | Ca2+, K+, Na+ | Biopara-digms |
| SLC24A4 | 3 | Ca2+, K+, Na+ | Biopara-digms |
| SLC24A5 | 3 | Ca2+, K+, Na+ | Biopara-digms |
| SLC25A26 | 3 | S-adenosyl methionine, S-adenosyl-ho-mocysteine, S-adenosyl-methionine | IUPHAR, Bi-oparadigms |
| SLC26A1 | 3 | glyoxylate, SO42-, oxalate | Biopara-digms, IUPHAR |
| SLC26A2 | 3 | Cl-, SO42-, oxalate | Biopara-digms, IUPHAR |
| SLC26A3 | 3 | Cl-, HCO3-, oxalate | IUPHAR, Bi-oparadigms |
| SLC2A5 | 3 | D-fructose, D-glucose, fructose | IUPHAR, Bi-oparadigms |
| SLC35A2 | 3 | UDP N-acetyl-glucosamine, UDP-galac-tose, UDP-N-acetylgalactosamine | IUPHAR, Bi-oparadigms |
| SLC35B4 | 3 | UDP N-acetyl-glucosamine, UDP-xy-lose, UDP-N-acetylglucosamine | IUPHAR, Bi-oparadigms |
| SLC42A1 | 3 | CO2, NH3, NH4+ | IUPHAR, Bi-oparadigms |

| | | | |
|---|---|---|---|
| SLC45A2 | 3 | fructose, glucose, sucrose | Biopara-digms |
| SLC45A3 | 3 | fructose, glucose, sucrose | Biopara-digms |
| SLC45A4 | 3 | fructose, glucose, sucrose | Biopara-digms |
| SLC4A1 | 3 | Cl-, HCO3-, chloride bicarbonate | IUPHAR, Bi-oparadigms |
| SLC4A2 | 3 | Cl-, HCO3-, chloride bicarbonate | IUPHAR, Bi-oparadigms |
| SLC4A3 | 3 | Cl-, HCO3-, chloride bicarbonate | IUPHAR, Bi-oparadigms |
| SLC4A4 | 3 | carbonate, sodium bicarbonate, NaHCO3- | Biopara-digms, IUPHAR |
| SLC4A5 | 3 | carbonate, sodium bicarbonate, NaHCO3- | Biopara-digms, IUPHAR |
| SLC4A8 | 3 | Cl-, sodium bicarbonate chloride, NaHCO3- | IUPHAR, Bi-oparadigms |
| SLC5A2 | 3 | D-glucose, alpha-MDG, glucose | IUPHAR, Bi-oparadigms |
| SLC60A2 | 3 | D-glucose, alpha-Me-glucose, D-glu-cose, α-Me-glucose | Biopara-digms, IUPHAR |
| SLC6A1 | 3 | GABA, guvacine, nipecotic acid | IUPHAR, Bi-oparadigms |
| SLC6A6 | 3 | GABA, taurine, beta-alanine | IUPHAR, Bi-oparadigms |
| SLC8B1 | 3 | Ca2+, Li+, Na+ | Biopara-digms |
| SLC9A6 | 3 | H+, K+, Na+ | Biopara-digms |
| SLC9A8 | 3 | H+, K+, Na+ | Biopara-digms |
| SLC9A9 | 3 | H+, K+, Na+ | Biopara-digms |
| SLC11A1 | 2 | Fe2+, Mn2+ | IUPHAR, Bi-oparadigms |
| SLC12A3 | 2 | Cl-, Na+ | Biopara-digms |
| SLC12A4 | 2 | Cl-, K+ | Biopara-digms |
| SLC12A5 | 2 | Cl-, K+ | Biopara-digms |
| SLC12A6 | 2 | Cl-, K+ | Biopara-digms |
| SLC12A7 | 2 | Cl-, K+ | Biopara-digms |

| | | | |
|---|---|---|---|
| SLC16A8 | 2 | L-lactic acid, lactate | IUPHAR, Bioparadigms |
| SLC18B1 | 2 | spermidine, spermine | IUPHAR |
| SLC22A13 | 2 | urate, organic anions | Biopara-digms |
| SLC22A16 | 2 | L-carnitine, noncharged compounds | IUPHAR, Bioparadigms |
| SLC25A18 | 2 | L-glutamic acid, glutamate | IUPHAR, Bioparadigms |
| SLC25A20 | 2 | acylcarnitine, carnitine | Biopara-digms |
| SLC25A22 | 2 | L-glutamic acid, glutamate | IUPHAR, Bioparadigms |
| SLC25A29 | 2 | acylcarnitine, ornithine | Biopara-digms |
| SLC25A31 | 2 | ADP, ATP | Biopara-digms |
| SLC25A4 | 2 | ADP, ATP | Biopara-digms |
| SLC25A5 | 2 | ADP, ATP | Biopara-digms |
| SLC25A6 | 2 | ADP, ATP | Biopara-digms |
| SLC27A1 | 2 | LCFA, VLCFA | Biopara-digms |
| SLC27A2 | 2 | LCFA, VLCFA | Biopara-digms |
| SLC27A3 | 2 | LCFA, VLCFA | Biopara-digms |
| SLC27A4 | 2 | LCFA, VLCFA | Biopara-digms |
| SLC27A5 | 2 | LCFA, bile acids | Biopara-digms |
| SLC27A6 | 2 | LCFA, VLCFA | Biopara-digms |
| SLC2A12 | 2 | D-glucose, glucose | IUPHAR, Bioparadigms |
| SLC31A1 | 2 | copper, cisplatin | IUPHAR, Bioparadigms |
| SLC31A2 | 2 | copper, cisplatin | IUPHAR, Bioparadigms |
| SLC33A1 | 2 | acetyl CoA, acetyl-CoA | IUPHAR, Bioparadigms |
| SLC35A3 | 2 | UDP N-acetyl-glucosamine, UDP-N-acetylglucosamine | IUPHAR, Bioparadigms |
| SLC35B2 | 2 | A3P5PS, PAPS | IUPHAR, Bioparadigms |
| SLC35B3 | 2 | A3P5PS, PAPS | IUPHAR, Bioparadigms |

| | | | |
|---|---|---|---|
| SLC35D1 | 2 | UDP-N-acetylgalactosamine, UDP-glucuronic acid | Biopara-digms |
| SLC35D2 | 2 | UDP-N-acetylgalactosamine, UDP-N-acetylglucosamine | IUPHAR, Bioparadigms |
| SLC37A1 | 2 | glucose 6-phosphate, glycerol 3-phosphate | IUPHAR |
| SLC39A7 | 2 | Mn, Zn | Biopara-digms |
| SLC40A1 | 2 | Fe2+, ferrous iron | IUPHAR, Bioparadigms |
| SLC42A3 | 2 | NH3, NH4+ | IUPHAR, Bioparadigms |
| SLC4A10 | 2 | sodium bicarbonate chloride, NaHCO3- | Biopara-digms, IUPHAR |
| SLC4A7 | 2 | chloride bicarbonate, NaHCO3- | Biopara-digms, IUPHAR |
| SLC59A1 | 2 | lysophosphatidylcholine, LPC (lysophosphatidylcholine) form of DHA (docosahexaenoic acid) | Biopara-digms, IUPHAR |
| SLC5A11 | 2 | chiro-inositol, myoinositol | Biopara-digms |
| SLC5A3 | 2 | glucose, myoinositol | Biopara-digms |
| SLC5A7 | 2 | choline, triethylcholine | IUPHAR, Bioparadigms |
| SLC61A1 | 2 | molybdate, Molybdate | IUPHAR, Bioparadigms |
| SLC6A12 | 2 | GABA, betaine | IUPHAR, Bioparadigms |
| SLC6A15 | 2 | large, neutral amino acids | Biopara-digms |
| SLC6A7 | 2 | L-proline, proline | IUPHAR, Bioparadigms |
| SLC6A9 | 2 | glycine, sarcosine | IUPHAR, Bioparadigms |
| SLC7A11 | 2 | L-glutamate, cystine anionic form | Biopara-digms |
| SLC7A13 | 2 | L-aspartate, L-glutamate | Biopara-digms |
| SLC7A6 | 2 | cationic amino acids, large neutral L-amino acids | Biopara-digms |
| SLC7A7 | 2 | cationic amino acids, large neutral L-amino acids | Biopara-digms |
| SLC7A9 | 2 | cationic amino acids, large neutral amino acids | Biopara-digms |
| SLC8A1 | 2 | Ca2+, Na+ | Biopara-digms |

| | | | |
|---|---|---|---|
| SLC8A2 | 2 | Ca2+, Na+ | Biopara-digms |
| SLC8A3 | 2 | Ca2+, Na+ | Biopara-digms |
| SLC9B2 | 2 | Li+, Na+ | Biopara-digms |
| SLC9C1 | 2 | H+, Na+ | Biopara-digms |
| SLC12A9 | 1 | polyamines | Biopara-digms |
| SLC14A2 | 1 | urea | IUPHAR |
| SLC17A3 | 1 | organic anions | Biopara-digms |
| SLC17A6 | 1 | glutamate | Biopara-digms |
| SLC17A7 | 1 | glutamate | Biopara-digms |
| SLC17A8 | 1 | glutamate | Biopara-digms |
| SLC19A2 | 1 | thiamine | IUPHAR, Bi-oparadigms |
| SLC19A3 | 1 | thiamine | IUPHAR, Bi-oparadigms |
| SLC22A10 | 1 | ochratoxin A | IUPHAR |
| SLC22A18 | 1 | probably organic anions | Biopara-digms |
| SLC22A20 | 1 | probably organic anions | Biopara-digms |
| SLC22A9 | 1 | organic anions | Biopara-digms |
| SLC22B4 | 1 | Nicotinate | Biopara-digms |
| SLC23A1 | 1 | L-ascorbic acid | IUPHAR, Bi-oparadigms |
| SLC23A2 | 1 | L-ascorbic acid | IUPHAR, Bi-oparadigms |
| SLC23A4 | 1 | 5-fluorouracil | IUPHAR |
| SLC25A16 | 1 | CoA and congeners | IUPHAR |
| SLC25A28 | 1 | Fe2+ | Biopara-digms |
| SLC25A3 | 1 | phosphate | Biopara-digms |
| SLC25A32 | 1 | folate | Biopara-digms |
| SLC25A33 | 1 | UTP | Biopara-digms |
| SLC25A36 | 1 | pyrimidine nucleotides | Biopara-digms |
| SLC25A37 | 1 | Fe2+ | Biopara-digms |

| | | | |
|---|---|---|---|
| SLC25A38 | 1 | glycine | Biopara-digms |
| SLC25A41 | 1 | ATP-Mg / Pi | Biopara-digms |
| SLC25A7 | 1 | H+ | Biopara-digms |
| SLC25A8 | 1 | H+ | Biopara-digms |
| SLC25A9 | 1 | H+ | Biopara-digms |
| SLC2A6 | 1 | glucose | Biopara-digms |
| SLC34A1 | 1 | inorganic phosphate | Biopara-digms |
| SLC34A2 | 1 | inorganic phosphate | Biopara-digms |
| SLC34A3 | 1 | inorganic phosphate | Biopara-digms |
| SLC35A1 | 1 | CMP-sialic acid | IUPHAR, Bi-oparadigms |
| SLC35C1 | 1 | GDP-fucose | IUPHAR, Bi-oparadigms |
| SLC35C2 | 1 | GDP-fucose | Biopara-digms |
| SLC37A2 | 1 | glucose 6-phosphate | IUPHAR |
| SLC37A4 | 1 | glucose 6-phosphate | IUPHAR |
| SLC39A1 | 1 | Zn | Biopara-digms |
| SLC39A10 | 1 | Zn | Biopara-digms |
| SLC39A12 | 1 | Zn | Biopara-digms |
| SLC39A13 | 1 | Zn | Biopara-digms |
| SLC39A2 | 1 | Zn | Biopara-digms |
| SLC39A3 | 1 | Zn | Biopara-digms |
| SLC39A4 | 1 | Zn | Biopara-digms |
| SLC39A5 | 1 | Zn | Biopara-digms |
| SLC39A6 | 1 | Zn | Biopara-digms |
| SLC44A1 | 1 | choline | IUPHAR, Bi-oparadigms |
| SLC44A2 | 1 | choline | IUPHAR, Bi-oparadigms |
| SLC46A3 | 1 | Uncertain | Biopara-digms |

| | | | |
|---|---|---|---|
| SLC48A1 | 1 | heme | Biopara-digms |
| SLC49A1 | 1 | heme | IUPHAR, Bi-oparadigms |
| SLC49A2 | 1 | heme | IUPHAR, Bi-oparadigms |
| SLC50A1 | 1 | glucose | Biopara-digms |
| SLC51A | 1 | bile acids | Biopara-digms |
| SLC51B | 1 | steroids | Biopara-digms |
| SLC52A1 | 1 | riboflavin | IUPHAR |
| SLC52A2 | 1 | riboflavin | IUPHAR |
| SLC52A3 | 1 | riboflavin | IUPHAR |
| SLC53A1 | 1 | Phosphate | IUPHAR, Bi-oparadigms |
| SLC54A1 | 1 | Pyruvate | IUPHAR, Bi-oparadigms |
| SLC54A2 | 1 | Pyruvate | IUPHAR, Bi-oparadigms |
| SLC54A3 | 1 | Pyruvate | IUPHAR, Bi-oparadigms |
| SLC56A1 | 1 | pyridoxin | Biopara-digms |
| SLC57A2 | 1 | Mg2+ | IUPHAR, Bi-oparadigms |
| SLC58A1 | 1 | Mg2+ | IUPHAR, Bi-oparadigms |
| SLC62A1 | 1 | Pyrophosphate | IUPHAR, Bi-oparadigms |
| SLC65A1 | 1 | Cholesterol | IUPHAR, Bi-oparadigms |
| SLC65A2 | 1 | Cholesterol | IUPHAR, Bi-oparadigms |
| SLC6A17 | 1 | neutral amino acids | Biopara-digms |
| SLC6A18 | 1 | neutral amino acids | Biopara-digms |
| SLC6A19 | 1 | neutral amino acids | Biopara-digms |
| SLC6A5 | 1 | glycine | IUPHAR, Bi-oparadigms |
| SLC6A8 | 1 | creatine | IUPHAR, Bi-oparadigms |

***Table 18 – Classification of substrates for SLCs according to Murcko Scaffolds***

| SLC | Unique count*(RDKit Mol (from InChI code) (Murcko)) |
|---|---|
| SLC22A1 | 82 |
| SLCO1B1 | 62 |
| SLC15A1 | 59 |
| SLCO1B3 | 49 |
| SLCO2B1 | 39 |
| SLCO1A2 | 35 |
| SLC22A8 | 22 |
| SLC22A2 | 16 |
| SLC22A6 | 13 |
| SLC29A1 | 13 |
| SLC22A7 | 12 |
| SLC47A1 | 12 |
| SLC22A11 | 11 |
| SLC29A2 | 11 |
| SLC10A2 | 10 |
| SLC47A2 | 9 |
| SLC15A2 | 7 |
| SLC36A1 | 7 |
| SLCO3A1 | 7 |
| SLC10A1 | 6 |
| SLC18A1 | 6 |
| SLC18A2 | 6 |
| SLC22A3 | 6 |
| SLC22A4 | 6 |
| SLC29A3 | 6 |
| SLCO4A1 | 6 |
| SLC15A4 | 5 |
| SLC22A5 | 5 |
| SLC28A3 | 5 |
| SLC29A4 | 5 |
| SLC10A6 | 4 |
| SLC16A1 | 4 |
| SLC18A3 | 4 |
| SLC19A1 | 4 |
| SLC22A32 | 4 |
| SLC25A17 | 4 |
| SLC28A1 | 4 |
| SLC28A2 | 4 |
| SLC46A1 | 4 |
| SLC51A-SLC51B | 4 |
| SLC5A8 | 4 |
| SLC15A3 | 3 |
| SLC16A10 | 3 |
| SLC16A5 | 3 |
| SLC17A1 | 3 |
| SLC22A12 | 3 |

| | |
|---|---|
| SLC36A2 | 3 |
| SLC6A14 | 3 |
| SLC6A4 | 3 |
| SLC7A5 | 3 |
| SLCO2A1 | 3 |
| SLC17A9 | 2 |
| SLC2A1 | 2 |
| SLC2A10 | 2 |
| SLC2A9 | 2 |
| SLC35B2 | 2 |
| SLC35B3 | 2 |
| SLC36A4 | 2 |
| SLC38A2 | 2 |
| SLC38A9 | 2 |
| SLC45A1 | 2 |
| SLC45A2 | 2 |
| SLC45A3 | 2 |
| SLC45A4 | 2 |
| SLC5A3 | 2 |
| SLC5A4 | 2 |
| SLC5A6 | 2 |
| SLC6A1 | 2 |
| SLC6A11 | 2 |
| SLC6A13 | 2 |
| SLC6A2 | 2 |
| SLC6A20 | 2 |
| SLC6A3 | 2 |
| SLC7A8 | 2 |
| SLCO1C1 | 2 |
| SLCO4C1 | 2 |
| SLC16A2 | 1 |
| SLC17A5 | 1 |
| SLC19A2 | 1 |
| SLC19A3 | 1 |
| SLC1A1 | 1 |
| SLC1A2 | 1 |
| SLC1A3 | 1 |
| SLC1A6 | 1 |
| SLC1A7 | 1 |
| SLC22A10 | 1 |
| SLC22A13 | 1 |
| SLC22B1 | 1 |
| SLC22B4 | 1 |
| SLC23A1 | 1 |
| SLC23A2 | 1 |
| SLC23A4 | 1 |
| SLC25A19 | 1 |
| SLC25A2 | 1 |
| SLC25A20 | 1 |

| | |
|---|---|
| SLC25A23 | 1 |
| SLC25A24 | 1 |
| SLC25A26 | 1 |
| SLC25A29 | 1 |
| SLC25A31 | 1 |
| SLC25A32 | 1 |
| SLC25A33 | 1 |
| SLC25A4 | 1 |
| SLC25A42 | 1 |
| SLC25A5 | 1 |
| SLC25A6 | 1 |
| SLC2A11 | 1 |
| SLC2A12 | 1 |
| SLC2A13 | 1 |
| SLC2A2 | 1 |
| SLC2A3 | 1 |
| SLC2A4 | 1 |
| SLC2A5 | 1 |
| SLC2A6 | 1 |
| SLC2A7 | 1 |
| SLC2A8 | 1 |
| SLC33A1 | 1 |
| SLC35A1 | 1 |
| SLC35A2 | 1 |
| SLC35A3 | 1 |
| SLC35B4 | 1 |
| SLC35C1 | 1 |
| SLC35C2 | 1 |
| SLC35D1 | 1 |
| SLC35D2 | 1 |
| SLC38A1 | 1 |
| SLC38A3 | 1 |
| SLC38A5 | 1 |
| SLC38A7 | 1 |
| SLC38A8 | 1 |
| SLC43A1 | 1 |
| SLC43A2 | 1 |
| SLC48A1 | 1 |
| SLC49A1 | 1 |
| SLC49A2 | 1 |
| SLC50A1 | 1 |
| SLC52A1 | 1 |
| SLC52A2 | 1 |
| SLC52A3 | 1 |
| SLC56A1 | 1 |
| SLC5A1 | 1 |
| SLC5A10 | 1 |
| SLC5A11 | 1 |
| SLC5A12 | 1 |

| | |
|---|---|
| SLC5A2 | 1 |
| SLC5A9 | 1 |
| SLC60A2 | 1 |
| SLC65A1 | 1 |
| SLC65A2 | 1 |
| SLC6A7 | 1 |
| SLC7A1 | 1 |
| SLC7A2 | 1 |

# Table list

*Table 1- Results of the Metabolism and Transporter database substrate collection.* ..... 21
*Table 2 – Results of the UCSF-FDA database substrate collection.* ..... 24
*Table 3 – Results of the Bioparadigms database substrate collection* ..... 26
*Table 4 – Results of the GtoPbd database substrate collection* ..... 30
*Table 5 – Configuration for the five Stringe Replacer nodes for the text correction.* ..... 33
*Table 6 – Data overlap of the databases* ..... 41
*Table 7 – Top 10 SLCs with the most substrates* ..... 43
*Table 8 – Top 10 SLCs with the most different Murcko Scaffolds* ..... 43
*Table 9 – Replacement list* ..... 46
*Table 10 – Replacement list* ..... 46
*Table 11 – Expression for the Rule-based Row Splitter for the Bioparadigms workflow* ..... 47
*Table 12 – Bioparadigms_SLC and substrate list (02.02.2021)* ..... 47
*Table 13 – Replacement list* ..... 62
*Table 14- Replacement list* ..... 65
*Table 15 – SLC families* ..... 66
*Table 16 – Configuration of the Rule Engine node* ..... 75
*Table 17 – Results SLC and their substrates* ..... 76
*Table 18 – Classification of substrates for SLCs according to Murcko Scaffolds* ..... 107

# Figure list

*Figure 1 - Transporter protein list of the Metabolism and Transporter Database. The list includes data about solute carries and ATP-binding cassette sub-families (25.5.21).* ..... 10
*Figure 2 – Download as text file from the Metabolism and Transporter Database for the solute carrier SLC2A1* ..... 11
*Figure 3 – UCSF-FDA database: drug transporters in human liver and kidney – PMC Copyright Notice (public domain material) (25.5.21)* ..... 12
*Figure 4 – SLC list of the Bioparadigms Database. The list includes data about SLC1 – High-affinity glutamate and neutral amino acid transporter family (06.07.21).* ..... 13
*Figure 5 – Example page of the IUPHAR/BPS database.* ..... 14
*Figure 6 – Search Interface of the Chemical Identifier Resolver.* ..... 15
*Figure 7 – Search Interface of the Chemical Translation Service.* ..... 16
*Figure 8 – Chemical Translation Service_ Restweb Service for InChIkeys.* ..... 16
*Figure 9 - KNIME start screen page.* ..... 18
*Figure 10 – KNIME_workflow_Analysis of SLC families according to substrates similarity.* ..... 19
*Figure 11 – KNIME workflow_Data collection from UCSF-FDA, Metrabase, IUPHAR/BPS and Bioparadigms.* ..... 20
*Figure 12 – KNIME workflow collecting data from the Metabolism and Transporter Database.* ..... 20
*Figure 13 – Configuring the GroupBy node for the workflow of the Metabolism and Transporter database data collection.* ..... 21
*Figure 14 – KNIME workflow collecting data from the UCSF-FDA database.* ..... 22
*Figure 15 – Configuring the GroupBy node for the workflow of the USCF-FDA data collection* ..... 23
*Figure 16 – KNIME workflow collecting data from the Bioparadigms database.* ..... 25
*Figure 17 – Removing the non-breaking space by a String Replacer node.* ..... 28
*Figure 18- Configuring the String Replacer node to remove the non-breaking space.* ..... 28
*Figure 19 - KNIME workflow collecting data from the IUPHAR/BPS database created by Daniela Digles, a member of the Pharmacoinformatic Research Group.* ..... 29
*Figure 20 – The workflow for the ligand structure search of the IUPHAR-BPS query.* ..... 29
*Figure 21 – Concatenate node* ..... 32
*Figure 22 – Expert of the concatenated table, which shows Amiloride from two different databases.* ..... 32
*Figure 23 – workflow overview of finding a standardized molecular structure for a given substrate name.* ..... 33
*Figure 24 – structure tools query* ..... 34
*Figure 25 – Removing the "+"-sign with a String Replacer node by using the code "%20".* ..... 34
*Figure 26 – CIR query for the structure search.* ..... 34
*Figure 27 – The workflow of the PubChem query for the structure search.* ..... 35
*Figure 28 – Knime Workflow_structure text search* ..... 35

*Figure 29 – Problem solving for the removed salts and ions by inserting a Molecular properties node.* ..... 36
*Figure 30 – Configuring the Molecular Properties node to calculate the Number of the Heavy Atoms.* 36
*Figure 31 – The workflow for the metanode InChI_original_database* ..... 37
*Figure 32 - Merging the substrates to the SLC names.* ..... 37
*Figure 33 – KNIME workflow_ visualization the data.* ..... 38
*Figure 34 – Count InChIs match* ..... 38
*Figure 35 – Knime Workflow_Visualisation of structures (Scaffolds)* ..... 39
*Figure 36 – Extract from the Renderer to Image node results.* ..... 39
*Figure 37 - Data overlap of the databases.* ..... 40
*Figure 38 – Extract of the data-visualization as a heatmap. The x-axis represents the Murcko Scaffolds, and the y-axis represents the different solute carriers. The blue squares indicate from which SLC the substrate is transported, where a black area appears, the substrate is not transported.* ..... 44

## References


(1) Meixner, E.; Goldmann, U.; Sedlyarov, V.; Scorzoni, S.; Rebsamen, M.; Girardi, E.; Superti-Furga, G. A Substrate-Based Ontology for Human Solute Carriers. *Molecular Systems Biology* **2020**, *16* (7), e9652. https://doi.org/10.15252/msb.20209652.

(2) Hediger, M. A.; Clémençon, B.; Burrier, R. E.; Bruford, E. A. The ABCs of Membrane Transporters in Health and Disease (SLC Series): Introduction. *Mol Aspects Med* **2013**, *34* (2–3), 95–107. https://doi.org/10.1016/j.mam.2012.12.009.

(3) Schumann, T.; König, J.; Henke, C.; Willmes, D. M.; Bornstein, S. R.; Jordan, J.; Fromm, M. F.; Birkenfeld, A. L. Solute Carrier Transporters as Potential Targets for the Treatment of Metabolic Disease. *Pharmacol Rev* **2020**, *72* (1), 343–379. https://doi.org/10.1124/pr.118.015735.

(4) Pramod, A. B.; Foster, J.; Carvelli, L.; Henry, L. K. SLC6 Transporters: Structure, Function, Regulation, Disease Association and Therapeutics. *Mol Aspects Med* **2013**, *34* (2–3), 197–219. https://doi.org/10.1016/j.mam.2012.07.002.

(5) Grewer, C.; Gameiro, A.; Rauen, T. SLC1 Glutamate Transporters. *Pflugers Arch* **2014**, *466* (1), 3–24. https://doi.org/10.1007/s00424-013-1397-7.

(6) Superti-Furga, G.; Lackner, D.; Wiedmer, T.; Ingles-Prieto, A.; Barbosa, B.; Girardi, E.; Goldmann, U.; Gürtl, B.; Klavins, K.; Klimek, C.; Lindinger, S.; Liñeiro-Retes, E.; Müller, A. C.; Onstein, S.; Redinger, G.; Reil, D.; Sedlyarov, V.; Wolf, G.; Crawford, M.; Everley, R.; Hepworth, D.; Liu, S.; Noell, S.; Piotrowski, M.; Stanton, R.; Zhang, H.; Corallino, S.; Faedo, A.; Insidioso, M.; Maresca, G.; Redaelli, L.; Sassone, F.; Scarabottolo, L.; Stucchi, M.; Tarroni, P.; Tremolada, S.; Batoulis, H.; Becker, A.; Bender, E.; Chang, Y.-N.; Ehrmann, A.; Müller-Fahrnow, A.; Pütter, V.; Zindel, D.; Hamilton, B.; Lenter, M.; Santacruz, D.; Viollet, C.; Whitehurst, C.; Johnsson, K.; Leippe, P.; Baumgarten, B.; Chang, L.; Ibig, Y.; Pfeifer, M.; Reinhardt, J.; Schönbett, J.; Selzer, P.; Seuwen, K.; Bettembourg, C.; Biton, B.; Czech, J.; de Foucauld, H.; Didier, M.; Licher, T.; Mikol, V.; Pommereau, A.; Puech, F.; Yaligara, V.; Edwards, A.; Bongers, B. J.; Heitman, L. H.; IJzerman, A. P.; Sijben, H. J.; van Westen, G. J. P.; Grixti, J.; Kell, D. B.; Mughal, F.; Swainston, N.; Wright-Muelas, M.; Bohstedt, T.; Burgess-Brown, N.; Carpenter, L.; Dürr, K.; Hansen, J.; Scacioc, A.; Banci, G.; Colas, C.; Digles, D.; Ecker, G.; Füzi, B.; Gamsjäger, V.; Grandits, M.; Martini, R.; Troger, F.; Altermatt, P.; Doucerain, C.; Dürrenberger, F.; Manolova, V.; Steck, A.-L.; Sundström, H.; Wilhelm, M.; Steppan, C. M. The RESOLUTE Consortium: Unlocking SLC Transporters for Drug Discovery. *Nat Rev Drug Discov* **2020**, *19* (7), 429–430. https://doi.org/10.1038/d41573-020-00056-6.

(7) Armstrong, J. F.; Faccenda, E.; Harding, S. D.; Pawson, A. J.; Southan, C.; Sharman, J. L.; Campo, B.; Cavanagh, D. R.; Alexander, S. P. H.; Davenport, A. P.; Spedding, M.; Davies, J. A.; NC-IUPHAR. The IUPHAR/BPS Guide to PHARMACOLOGY in 2020: Extending Immunopharmacology Content and Introducing the IUPHAR/MMV Guide to MALARIA PHARMACOLOGY. *Nucleic Acids Res* **2020**, *48* (D1), D1006–D1021. https://doi.org/10.1093/nar/gkz951.

(8) SLCtables http://slc.bioparadigms.org/ (accessed 2021 -06 -05).

(9) Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J. R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; Assempour, N.; Iynkkaran, I.; Liu, Y.; Maciejewski, A.; Gale, N.; Wilson, A.; Chin, L.; Cummings, R.; Le, D.; Pon, A.; Knox, C.; Wilson, M. DrugBank 5.0: A Major Update to the DrugBank Database

for 2018. *Nucleic Acids Research* **2018**, *46* (D1), D1074–D1082. https://doi.org/10.1093/nar/gkx1037.
(10) Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A. P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L. J.; Cibrián-Uhalte, E.; Davies, M.; Dedman, N.; Karlsson, A.; Magariños, M. P.; Overington, J. P.; Papadatos, G.; Smit, I.; Leach, A. R. The ChEMBL Database in 2017. *Nucleic Acids Res* **2017**, *45* (D1), D945–D954. https://doi.org/10.1093/nar/gkw1074.
(11) Morrissey, K. M.; Wen, C. C.; Johns, S. J.; Zhang, L.; Huang, S.-M.; Giacomini, K. M. The UCSF-FDA TransPortal: A Public Drug Transporter Database. *Clin Pharmacol Ther* **2012**, *92* (5), 545–546. https://doi.org/10.1038/clpt.2012.44.
(12) Mak, L.; Marcus, D.; Howlett, A.; Yarova, G.; Duchateau, G.; Klaffke, W.; Bender, A.; Glen, R. C. Metrabase: A Cheminformatics and Bioinformatics Database for Small Molecule Transporter Data Analysis and (Q)SAR Modeling. *Journal of Cheminformatics* **2015**, *7* (1), 31. https://doi.org/10.1186/s13321-015-0083-5.
(13) KNIME Analytics Platform https://www.knime.com/knime-analytics-platform (accessed 2021 -06 -24).
(14) Sarah Foth, An Analysis of SLC-Transporter Substrates “Diploma Thesis”, 2021.
(15) BioParadigms https://www.bioparadigms.org/ (accessed 2021 -06 -05).
(16) Alexander, S. P. H.; Kelly, E.; Mathie, A.; Peters, J. A.; Veale, E. L.; Armstrong, J. F.; Faccenda, E.; Harding, S. D.; Pawson, A. J.; Sharman, J. L.; Southan, C.; Davies, J. A. THE CONCISE GUIDE TO PHARMACOLOGY 2019/20: Transporters. *British Journal of Pharmacology* **2019**, *176* (S1), S397–S493. https://doi.org/10.1111/bph.14753.
(17) Alexander, S. P. H.; Kelly, E.; Mathie, A.; Peters, J. A.; Veale, E. L.; Armstrong, J. F.; Faccenda, E.; Harding, S. D.; Pawson, A. J.; Sharman, J. L.; Southan, C.; Buneman, O. P.; Cidlowski, J. A.; Christopoulos, A.; Davenport, A. P.; Fabbro, D.; Spedding, M.; Striessnig, J.; Davies, J. A. THE CONCISE GUIDE TO PHARMACOLOGY 2019/20: Introduction and Other Protein Targets. *British Journal of Pharmacology* **2019**, *176* (S1), S1–S20. https://doi.org/10.1111/bph.14747.
(18) NCI/CADD Group Chemoinformatics Tools and User Services https://cactus.nci.nih.gov/ncicadd/about.html (accessed 2021 -06 -23).
(19) NCI/CADD Chemical Identifier Resolver https://cactus.nci.nih.gov/chemical/structure?identifier=glucose&representation=stdinchikey (accessed 2021 -06 -23).
(20) CTS Proxy https://cts.fiehnlab.ucdavis.edu/services (accessed 2021 -06 -23).
(21) Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E. E. PubChem in 2021: New Data Content and Improved Web Interfaces. *Nucleic Acids Research* **2021**, *49* (D1), D1388–D1395. https://doi.org/10.1093/nar/gkaa971.
(22) About PubChem https://pubchemdocs.ncbi.nlm.nih.gov/about (accessed 2021 -06 -24).
(23) About https://www.knime.com/about (accessed 2021 -06 -24).
(24) *PharminfoVienna/Chemical-Structure-Standardisation*; PharminfoVienna, 2020.
(25) SLC15A1 solute carrier family 15 member 1 [Homo sapiens (human)] - Gene - NCBI https://www.ncbi.nlm.nih.gov/gene/6564 (accessed 2021 -07 -31).
(26) SLC22A1 solute carrier family 22 member 1 [Homo sapiens (human)] - Gene - NCBI https://www.ncbi.nlm.nih.gov/gene/6580#gene-expression (accessed 2021 -07 -31).

(27) Türková, A.; Jain, S.; Zdrazil, B. Integrative Data Mining, Scaffold Analysis, and Sequential Binary Classification Models for Exploring Ligand Profiles of Hepatic Organic Anion Transporting Polypeptides. *J. Chem. Inf. Model.* **2019**, *59* (5), 1811–1825. https://doi.org/10.1021/acs.jcim.8b00466.

I have tried to locate all holders of the image rights and obtained their consent to the use of the images in this work. Should a copyright infringement become known, I would like to request a report from me.