

Original article

GATMD: γ -Aminobutyric Acid Transporter Mutagenesis Database

Cynthia M. Anderson, Patrick D. Kidd and Sepehr Eskandari*

Biological Sciences Department, California State Polytechnic University, Pomona, CA 91768-4032, USA

*Corresponding author: Tel.: +1 909 869 4182; Fax: +1 909 869 4078; Email: seskandari@csupomona.edu

Submitted 6 September 2010; Revised 8 November 2010; Accepted 11 November 2010

Since the cloning of the first γ -aminobutyric acid (GABA) transporter (GAT1; SLC6A1) from rat brain in 1990, more than 50 published studies have provided structure–function information on investigator-designed rat and mouse GAT1 mutants. To date, more than 200 of 599 GAT1 residues have been subjected to mutagenesis experiments by substitution with different amino acids, and the resulting transporter functional properties have significantly advanced our understanding of the mechanism of Na^+ - and Cl^- -coupled GABA transport by this important member of the neurotransmitter:sodium symporter family. Moreover, many studies have addressed the functional consequences of amino acid deletion or insertion at various positions along the primary sequence. The enormity of this growing body of structure–function information has prompted us to develop GABA Transporter Mutagenesis Database (GATMD), a web-accessible, relational database of manually annotated biochemical, functional and pharmacological data reported on GAT1—the most intensely studied GABA transporter isoform. As of the last update of GATMD, 52 GAT1 mutagenesis papers have yielded 3360 experimental records, which collectively contain a total of $\sim 100\,000$ annotated parameters.

Database URL: <http://physiology.sci.csupomona.edu/GATMD/>

Introduction

In recent years, the proliferation of structure–function studies of proteins via investigator-designed site-directed mutagenesis has led to significant strides in our understanding of molecular, biochemical, cell biological, physiological and pharmacological principles that underlie life processes. For any given protein, the totality of the available mutagenesis data is related to the number of laboratories devoted to its function, and the diversity and complexity of standard and novel tools used to study the properties of wild-type and mutant forms of the protein. As the size of this data set grows, it becomes increasingly difficult to remain abreast of new developments and, in the absence of a comprehensive database, it becomes an arduous task to make coherent connections between existing and new aspects of the available data. For many proteins with significant perceived physiological and/or clinical relevance, this complexity of

structure-function data has left new and established investigators struggling to organize, assimilate, and comprehend this information.

A case in point is the Na^+ - and Cl^- -coupled transporter for the inhibitory neurotransmitter γ -aminobutyric acid (GABA). The GABA transporters (GATs) are electrogenic Na^+ - and Cl^- -coupled transporters that are responsible for maintaining low resting levels of GABA in the central nervous system, as well as for modulating synaptic and extra-synaptic GABAergic neurotransmission (1–5). As the regulators of the concentration of the principal inhibitory neurotransmitter in the brain, the GABA transporters are potential sites for therapeutic intervention, particularly in epilepsy and stroke (1, 6). Indeed, the GABA transporters have been implicated in epilepsy (3), and one isoform (GAT1) is the target of the anti-epileptic drug tiagabine (7, 8).

The GABA transporters belong to the large neurotransmitter/Na⁺ symporter family (NSS; 2.A.22 according to the transporter classification system; SLC6 according to the Human Genome Organization classification) (2, 9–12). Solute transport in these transporters is driven by the electrochemical potential gradient of Na⁺ and Cl⁻ and, in fact, 2 Na⁺ ions and 1 Cl⁻ ion are cotranslocated with GABA during each transport cycle (13–19). Four GABA transporter isoforms are present in the mammalian brain (rat/human GAT-1, BGT-1, GAT-2, GAT-3 corresponding to mouse GAT1, GAT2, GAT3 and GAT4, respectively) exhibiting significant differences in function, pharmacology, and localization (1, 5, 7, 8, 20–32). GAT1 was the first neurotransmitter transporter cloned (20, 21), and set the stage for the cloning and characterization of other important family members such as the norepinephrine, serotonin, dopamine and glycine transporters (33–40). Recently, high-resolution crystal structures of a bacterial leucine transporter from this family (41–44) have catapulted structure–function studies of these proteins to a new level of structure-guided mutagenesis design (45–51). Detailed structure–function studies of these transporters not only enhance our mechanistic understanding of secondary active Na⁺- and Cl⁻-coupled solute transport (52–56), but will also pave the way for structure-guided design of selective drugs for the treatment of disorders such as epilepsy, depression and anxiety, orthostatic intolerance and Parkinson's disease.

To date, more than 50 published studies have provided detailed structure–function information on investigator-designed rat and mouse GAT1 mutants. Here, we report the development and web implementation of GABA Transporter Mutagenesis Database (GATMD), a comprehensive, web-accessible, relational database for managing published GAT1 mutagenesis data. As nearly all mutagenesis studies of the GABA transporters have focused on the GABA transporter isoform 1 (GAT1), we have limited the database to this isoform. To the best of our knowledge, GATMD is the first publicly accessible, comprehensive mutagenesis data resource on any membrane transport protein. It is our sincere hope that this database and its web interface will facilitate knowledge dissemination on this important member of the neurotransmitter:sodium symporter family of transporters. GATMD is available on the Internet at <http://physiology.sci.csupomona.edu/GATMD/>.

Methods

Data source

The data annotated in GATMD were obtained from peer-reviewed published papers, all of which are found in the PubMed database (<http://www.ncbi.nlm.nih.gov/pubmed/>) (Figure 1). To date, more than 50 (52 as of

Literature search for studies of investigator-designed GAT1 (P23978, P31648) mutants

52 peer-reviewed publications
(1990 to October 2010)

Full-content manual annotation and
data compilation into XML database

Database Statistics
52 published studies
3,360 experimental records
~100,000 annotated parameters

MySQL database

GATMD web interface
<http://physiology.sci.csupomona.edu/GATMD/>

Figure 1. General approach for the development of GATMD. A thorough search of the PubMed database revealed 52 published studies in which investigator-designed mutants of GAT1 have been studied. A complete list of the papers annotated in GATMD is available online. Using the data model shown in Table 1 and described in greater detail in the online tutorial, full annotation of these published studies yielded 3360 experimental records containing a total of ~100000 annotated parameters. The web interface of GATMD utilizes Perl/CGI scripts to access the GATMD MySQL database.

October 2010) published studies have reported data on investigator-designed mutants of the GABA transporters and nearly all of these studies have focused on rat GAT1 (UniProtKB, P23978) and mouse GAT1 (UniProtKB, P31648) which have identical amino acid sequences. No mutagenesis experiments of human GAT1 (UniProtKB, P30531) exist in the literature. We, therefore, have focused our efforts on rat and mouse GAT1 mutagenesis data.

The data included in GATMD are organized around individual 'experimental records' (i.e. database records), which correspond to physical experiments performed on wild-type (WT) and laboratory-designed GAT1 mutants in order to elucidate one or more aspects of transporter structural, biochemical, functional and/or pharmacological properties. An experimental record can be thought of as an independent data unit of GATMD, which includes the discrete set of physical parameters used (e.g. substrate, ionic and/or inhibitor concentrations, experimental temperature, etc.) and the results obtained (e.g. transport rate, total and surface expression levels, glycosylation state, etc.) for any

given experiment included in the published study. Each experimental record can accommodate many different types of data, which represent real-life experimental parameters, and whose values occupy predefined database fields (database fields are described in greater detail under 'Data structure and database development' section). Thus, each published paper yielded many GATMD experimental records.

Our criterion for including an experiment in the database was simple. Any published experiment that examined an investigator-designed mutated sequence of rat or mouse GAT1, as well as its corresponding wild-type and control experiments, were included in the database. Thus, the curated data include mutants that exhibit biochemical, functional and/or pharmacological deviations from the wild-type phenotype as well as those that appear indistinguishable from the wild-type transporter. We defined a mutated sequence as one in which (i) one or more wild-type residues were substituted by other amino acids, (ii) one or more wild-type residues were deleted and/or (iii) amino acid insertions were introduced at any position along the wild-type primary sequence. In addition, the related control and wild-type experiments were annotated and included in GATMD. Studies which focused primarily on the effects of mutagenesis were fully annotated. Where the mutagenesis experiments constituted only a portion of a much larger study, annotation was limited to the mutagenesis experiments and the corresponding wild-type and control experiments. Experiments, in which a fusion protein was added to the carboxy tail or amino tail (which generally also included a linker sequence) without introducing a mutation in the primary sequence of GAT1, were not considered mutagenesis experiments and were not included in GATMD. However, experiments were included in GATMD when fusion proteins were present in addition to mutations of the primary sequence of GAT1.

Data structure and database development

We first designed a comprehensive data model for deposition of biochemical, functional and pharmacological data involving the GABA transporters. While we have only used this database to annotate mutagenesis studies, the database can in fact be used to fully annotate all papers published on the GABA transporters. Moreover, the database can be scaled to accommodate data obtained for other members of the SLC6 family of transporters (e.g. serotonin, dopamine, norepinephrine, glycine transporters; 10), as well as other electrogenic Na⁺-coupled transporters belonging to other solute carrier families (e.g. Na⁺/glucose, Na⁺/iodide and Na⁺/phosphate cotransporters; 57, 58).

As described above, the data in GATMD are organized as experimental records, each of which includes all of the physical experimental parameters used as well as the results obtained for a published experiment. Because a given

experiment may utilize a number of different techniques and the data obtained may be presented by the authors in a number of different ways, many predefined database fields are needed for accurate and complete annotation of experimental parameters. Therefore, to accomplish comprehensive annotation of mutagenesis experiments on GAT1, we first identified broad categories of data related to published experiments on the GABA transporters (Table 1). These categories included information about the published study, protein under study (rat or mouse GAT1 in this case), basic experimental conditions, mutation information, general description of data, experimental method used, experiment conclusion, as well as others (see Table 1). Subsequently, within each of the broad categories seen in Table 1, many subcategories were introduced in order to accommodate important details related to that data category (see 'Search' page online). All in all, a total of 366 subcategories were defined, which formed the basis for the creation of corresponding database fields. Because these data subcategories relate to actual attributes of physical experiments, we will henceforth interchangeably refer to these as database fields or formal parameters. Collectively, these formal parameters comprehensively define the basic information about a given paper (i.e. citation), mutation information (substitution, deletion or insertion, as well as the mutation locus), experimental methods employed in the study (e.g. tracer flux, two-electrode voltage clamp electrophysiology, etc.), all aspects of experimental conditions (e.g. temperature, substrate concentrations, etc.), the results obtained (e.g. percentage of GABA uptake by mutant versus wild-type), as well as the conclusion of the experiment. As noted above, these parameters are logically grouped according to the experimental methodology employed. For example, all parameters related to tracer flux experiments are grouped together under the 'Tracer flux experiment' category (Table 1). Thus, the data from any given experimental record would populate only the fields related to that group, which comprises a subset of all available database fields. Clearly, the selection of these parameters has been guided by the papers annotated thus far, and it is reasonable to predict that this list will continue to evolve as additional papers are annotated and as newly published studies employ new methodologies not yet represented in the existing predefined set of database fields.

We used the Delphi[®] (Embarcadero; San Francisco, CA, USA) integrated development environment, running within the Windows operating system, to develop our own database management software, which we have named *GATMD-Manager*. *GATMD-Manager* stores the native GATMD database file in Extensible Markup Language (XML) format. We chose the XML data format because it provided the modularity needed when reading the data using standard XML parsers/interpreters, as well as

Table 1. General data categories in GATMD

 General data category

- (1) Study citation
 - (2) Protein studied
 - (3) WT or mutant protein, or control cells
 - (4) Mutation information
 - (5) Data presentation in publication
 - (6) Basic experimental information
 - (7) Experimental methods used
 - (a) Transporter operation mode and transport direction
 - (b) Cell-surface biotinylation, SDS-PAGE and immunoblot analysis
 - (c) Total and cell-surface transporter expression
 - (d) Reconstitution into proteoliposomes
 - (e) Na⁺- and Cl⁻-dependence examined
 - (f) Ion replacement experiment
 - (g) Inhibition experiment
 - (h) Sulphydryl modification experiment
 - (i) Treatment with other drugs
 - (j) Tracer flux experiment
 - (k) Electrophysiology experiment
 - (l) Tracer uptake under voltage clamp
 - (m) Transporter turnover rate determined
 - (n) Steady-state kinetic parameters determined
 - (o) Immunohistochemistry and/or immunofluorescence experiment
 - (p) Cellular localization and/or trafficking experiment
 - (q) Real-time optical measurements
 - (r) Binding assay and/or immunoprecipitation experiment
 - (8) Overall assessment of mutant transporter
 - (9) Main conclusion of experiment
 - (10) Main conclusion of study
 - (11) Comments
-

The general data categories shown in this table (each of which contains many subcategories; see 'Search' page online) were used to manually annotate all published mutagenesis studies of GAT1. Within the XML data document, each of the categories shown (modified to conform to XML tag naming conventions) served as a child element to a 'GATMDEXperiment' parent element. Each category contains one or more subcategories (i.e. child elements). Please see under the Advanced Search Functions menu on the 'Search' page online for a full listing of these subcategories. The entire hierarchical tree structure the 'GATMDEXperiment' parent element can be found online at: http://physiology.sci.csupomona.edu/GATMD/data_structure.xml, where for simplicity, element attributes have been omitted.

different software modules developed in house. In addition, this data format allowed us to rapidly update the data structure (i.e. modify, add or delete data fields). Within the XML data document, each experimental record is contained within a 'GATMDEXperiment' element, which in turn contains, as child elements, each of the broad data categories listed in Table 1. Each of the data category elements, in turn, contains one or more child elements related to that category (see the 'Search' page online). The entire hierarchical tree structure of an experimental record can be found online at: http://physiology.sci.csupomona.edu/GATMD/data_structure.xml, where for simplicity,

element attributes have been omitted. *GATMD-Manager* exports the native XML database to the web server in a proper relational format to be imported by MySQL. The major data categories noted above and shown in Table 1 are used to construct the relational database tables.

Annotation

Using our own database management software (*GATMD-Manager*) running within the Windows operating system, the categories shown in Table 1 (each of which contains many subcategories; see 'Search' page online) were used to manually annotate all published mutagenesis

studies on GAT1. For any given publication, the data included in the paper as well as any supplementary materials were curated. If data were discussed in the text but were not shown in figures, tables or summary diagrams, the data were annotated and appropriately marked as 'data not shown'. To ensure the fidelity of the data included in GATMD, each experimental record was independently annotated by each of the authors of this study and, in addition, all authors performed numerous quality control checks of all experimental records.

In order to curate quantitative experimental results (e.g. uptake rates, etc.), we relied on information provided in the form of tables, figures or numerical values explicitly reported in the text. When an experimental value was represented in a figure and its numerical value was also explicitly given in a table or in the text, we reported the value explicitly noted by the authors. On the other hand, if quantitative information was included only in the figures (e.g. bar or scatter plots, electrophysiological traces, etc.), we used the portable document format (PDF) file of published articles in order to obtain measurements of the experimental variables. We again used Delphi[®] to develop our own software (entitled *XYCapture*) to capture *x*-axis and/or *y*-axis coordinates from figures of published studies. Every effort was made to report these quantitative data as accurately as possible to reflect the original plotted values. When the data were reported only qualitatively, this was appropriately annotated in database fields designed to handle qualitative parameters (e.g. less or greater than, or similar to WT).

Every annotated GATMD experimental record also includes a summary statement based on the experiment and the results obtained and, in addition, an overall conclusion statement was included for every paper annotated in GATMD summarizing the totality of the study findings (Table 1). These statements were paraphrased from those provided in the published studies. In order to maintain objectivity in data curation, we refrained from interpreting the data beyond the results shown and/or conclusions drawn by the authors.

Web interface

GATMD is implemented on a Linux server (CentOS 5.2; www.centos.org) running Apache 2.2 (www.apache.org). Data queries are processed by MySQL 5.0 (www.mysql.com). Data requests are sent to MySQL from a Perl (www.perl.org) script which utilizes the Common Gateway Interface (CGI) for communication with the web server. The web interface utilizes HTML (HyperText Markup Language), CSS (Cascading Style Sheets) and JavaScript to provide static and dynamic content to the end user. From the start, a deliberate decision was made to maximize end-user access to the database by avoiding the use of additional programs, such as Java (Oracle Corporation) or Flash (Adobe Systems



Incorporated), for rendering the topology diagram or other features. Thus, for the end-user, the only strict requirement is to use a web browser in which JavaScript is enabled. While all annotation, curation, and database management functions may be performed via the web interface (using Perl-CGI scripts and MySQL), in practice and for security purposes, we only use *GATMD-Manager* for these operations.

Querying the database

The web interface of GATMD provides a rich set of functions for searching the database (<http://physiology.scic.pomona.edu/GATMD/search.html>). The GAT1 topology diagram shown on the 'Search' page provides a hyperlink for each of the 599 residues, which can be used to retrieve all experimental evidence associated with any given residue (Figure 2). The topology map shown conforms to the sequence alignment of Yamashita *et al.* (41); however, the user may select an alternative topology map according to the sequence alignment of Beuming *et al.* (59). In general, search results are initially presented as the summary statements of individual experimental records contained in GATMD, along with additional links that reveal all details of the experimental records returned by the query (for example, see Figure 3). Any residues implicated in the search results are also highlighted in the GAT1 topology diagram in order to provide a quick glance of the topological location targeted by the experiment. An online tutorial (<http://physiology.scic.pomona.edu/GATMD/tutorial.html>) provides detailed descriptions and examples for the basic and advanced search functions.

The database may also be searched by using functions available under the 'Basic Search Functions' (Figure 4) and 'Advanced Search Functions' (Figure 5) menus. A number of built-in search functions are available to the user under the 'Basic Search Functions' menu. For example, all mutagenized residues of GAT1 can be shown based on whether the mutation resulted from substitution, deletion, and/or insertion mutagenesis (Figure 6). All GAT1 functional sites inferred from the crystal structure of LeuT_{Aa} can be readily examined (41). These functional sites include the residues involved in coordinating Na⁺ and GABA in the transporter binding pocket, as well as those that serve as extracellular and cytoplasmic gates. Moreover, the user may examine the residues proposed to serve as the chloride binding site (47). Rat GAT1 mutants may also be searched according to phenotype (e.g. whether or not a given mutant mediates GABA uptake, GABA-evoked current, pre-steady-state charge movements, lithium leak, etc.). A simple text search function is also provided to enable maximum flexibility in searching all data (but not metadata, see below) of all database fields of all experimental records. Simple text searches are case-insensitive, may use the standard Boolean operators (AND, OR, NOT), may use parentheses

Search Results

Mutagenesis experiments involving **C399**  

Number of experiments: **212**

[All Mutations](#) [Substitutions Only](#) [Deletions Only](#) [Insertions Only](#)

Exp ID	Mutation	Experiment Summary	PubMed ID
779	sub:C399S	When expressed in HeLa cells, GABA uptake mediated by rat GAT-1 C399S is $28.0 \pm 0.2\%$ of that of WT rat GAT-1.	10438469
792	sub:C399S	Pretreatment with MTSEA (0.75 mM for 5 min) led to a $4.0 \pm 0.6\%$ inhibition of GABA uptake mediated by rat GAT-1 C399S expressed in HeLa cells.	10438469
805	sub:C399S	Pretreatment with MTSEA (0.75 mM for 5 min) led to a $3.8 \pm 0.5\%$ inhibition of GABA uptake mediated by rat GAT-1 C399S expressed in HeLa cells. GABA uptake mediated by rat GAT-1 C399S is $48 \pm 9\%$ of that of WT rat GAT-1.	10438469
806	sub:C399A	Pretreatment with MTSEA (0.75 mM for 5 min) led to a $4.0 \pm 0.7\%$ inhibition of GABA uptake mediated by rat GAT-1 C399A expressed in HeLa cells.	10438469
807	sub:C399G	Pretreatment with MTSEA (0.75 mM for 5 min) led to a $8.7 \pm 0.3\%$ inhibition of GABA uptake mediated by rat GAT-1 C399G expressed in HeLa cells.	10438469
809	sub:C399S	Pretreatment with NEM (1.5 mM for 5 min) led to a $8.0 \pm 0.4\%$ inhibition of GABA uptake mediated by rat GAT-1 C399S expressed in HeLa cells.	10438469

Figure 3. A representative screenshot is shown, where search results are initially presented as the summary statements of individual experimental records returned by the query. Any residues implicated in the search results are also highlighted in the GAT1 topology diagram (see Figure 6). Additional links are provided to refine the search and/or to reveal further details about any given experimental record.

for grouping search terms, and may use quotation marks to search for exact phrases. The default Boolean operator is AND. Other functions enable the user to rapidly locate a given amino acid or a sequence of amino acids. Detailed functionality for each of the built-in search functions is provided in the online tutorial.

The real advantages of full-content annotation and data placement into a comprehensive, relational database are realized by including a powerful interface for searching specific data fields of the database. The search engine can in principle reveal every aspect of the data in the database by searching for known database fields (i.e. formal parameters). In practice, not all parameters are scientifically useful as some relate to database management functions (e.g. record processing date, revision date, curator signature, etc.). The scientifically relevant parameters are available under the 'Advanced Search Functions' menu, where the parameters are presented according to their respective data categories (Figure 5). Also provided for each parameter is a brief definition, the type of data it contains and the appropriate values that may be passed to it for searching the particular database field it represents. For convenience to the end user, one or more parameters may be

selected from the menu and added (by double-clicking or using the 'Add' feature) to the search query, where the proper syntax is automatically rendered. The syntax of the search command is simple and follows the rules outlined below and described in greater detail in the online tutorial.

To search formal parameters specifically, the search query must enclose a formal parameter name within braces. For example, {ElectrophysiologyExperiment} will return all experiments in which electrophysiological methods were used. Formal parameters (but not the values passed to them; see below) are processed in a case-sensitive manner. A complete list of the formal parameters, their intended use, the data they contain, acceptable values and search phrase examples can be found in the online tutorial (http://physiology.sci.csupomona.edu/GATMD/formal_parameters.html).

A search can be further narrowed by specifying the desired value(s) of a formal parameter. This is done by passing a value of interest to the search command; enclosed in parentheses and immediately following a formal database parameter. For example, {ExpressionSystem(xenopus)} will return all experiments in which the expression system

Figure 4. Partial screenshot of GATMD basic search functions. A number of predefined search functions are available under the 'Basic Search Functions' menu. A text search box is also provided, where terms/phrases may be entered for searching all data in the database. Text searches must conform to the guidelines outlined in text and described in greater detail in the online tutorial. Advanced queries (see text and Figure 5) may also be entered in this search box.

used was the *Xenopus laevis* oocyte system. The values passed to formal parameters are treated in a case-insensitive manner.

Passing desired values to formal parameters is optional. Thus, `{SulfhydrylModificationExperiment()}` is equivalent to `{SulfhydrylModificationExperiment}`.

Both searches will return all experiments in which GAT1 was subjected to sulfhydryl modification. If no desired value is passed to the formal parameter, the search engine will return all experiments that have any data for that parameter. On the other hand, if a desired value is passed, the search will be more refined and limited to the

Advanced Search Functions

The general data categories are shown on the left below. Selecting from this list will display the relevant data sub-categories (i.e., formal parameters) on the right. Select any formal parameter to display a brief description, the type of data it contains, appropriate values that may be passed to it for searching the database, as well as a query example. Add a selected formal parameter to the search query (by double-clicking or using the button provided) and, if so desired, enter an appropriate value to be passed to the search engine. The search query may contain one or many formal parameters, and may also contain simple search words or phrases. Please see the [tutorial](#) as well as the complete list of the [formal parameters](#) for additional details.

General Data Categories in GATMD:

- Study citation
- Protein studied
- WT or mutant protein, or control cells
- Mutation information**
- Data presentation in publication
- Basic experimental information
- Transporter operation mode and transport direction
- Cell-surface biotinylation, SDS-PAGE and immunoblot analysis
- Total and cell-surface transporter expression
- Reconstitution into proteoliposomes
- Na⁺- and Cl⁻-dependence examined
- Ion replacement experiment
- Inhibition experiment
- Sulphydryl modification experiment
- Treatment with other drugs

Data Sub-categories in GATMD (i.e., formal parameters):

- MutationType**
- MutatedResidueNumber
- InsertionSequence
- Mutant_AminoAcidSequence
- MutationTopologicalLocation_Yamashita_2005
- MutationTopologicalLocation_Beuming_2006
- ResidueFunctionalSignificance_Yamashita_2005

MutationType

Description: Indicates whether a mutation was substitution, deletion, or insertion.

Data type: Simple text

Acceptable search input: Substitution, deletion, or insertion

Query result: All experimental records which involved the indicated mutation type.

Query example: {MutationType(insertion)}

Add Selected Formal Parameter To Search Query ---Insert Relational Operator---

{MutationType(substitution)}

Search Clear

Figure 5. Screenshot of GATMD advanced search functions. The general data categories of GATMD are shown on the left (see also Table 1). Selecting from this list will display the relevant data sub-categories (i.e. formal parameters) on the right. Selecting any formal parameter displays a brief description, the type of data it contains, appropriate values that may be passed to it for searching the database, as well as a query example. Advanced queries must conform to the guidelines outlined in text and described in greater detail in the online tutorial.

value of interest. For example, {SulphydrylReagentAbbreviatedName()} will return all experiments in which any sulphydryl reagent was used, whereas {SulphydrylReagentAbbreviatedName(mtset)} will return only those experiments in which MTSET was used as the sulphydryl reagent.

Boolean operators (AND, OR, NOT) may be used to refine the search of a formal parameter (with the exception of passing numbers for numerical searches; see below). For example, {SulphydrylReagentAbbreviatedName(mtset OR mtset OR mtsea)} will return all sulphydryl modification experiments in which the reagent used was

MTSET, MTSES, or MTSEA. As another example, {SulphydrylReagentAbbreviatedName(NOT mtset)} will return all experiments in which any sulphydryl modification reagent except MTSET was used.

Multiple advanced search commands may be combined by using the standard Boolean operators AND, OR, NOT. For example: {ElectrophysiologyExperiment} AND {IonReplacementExperiment} returns all electrophysiology experiments, which involved ion replacement (i.e. replacement of Na⁺ and/or Cl⁻ with an appropriate cation/anion). As another example, {ExpressionSystem(xenopus)}{SulphydrylModificationExperiment()}

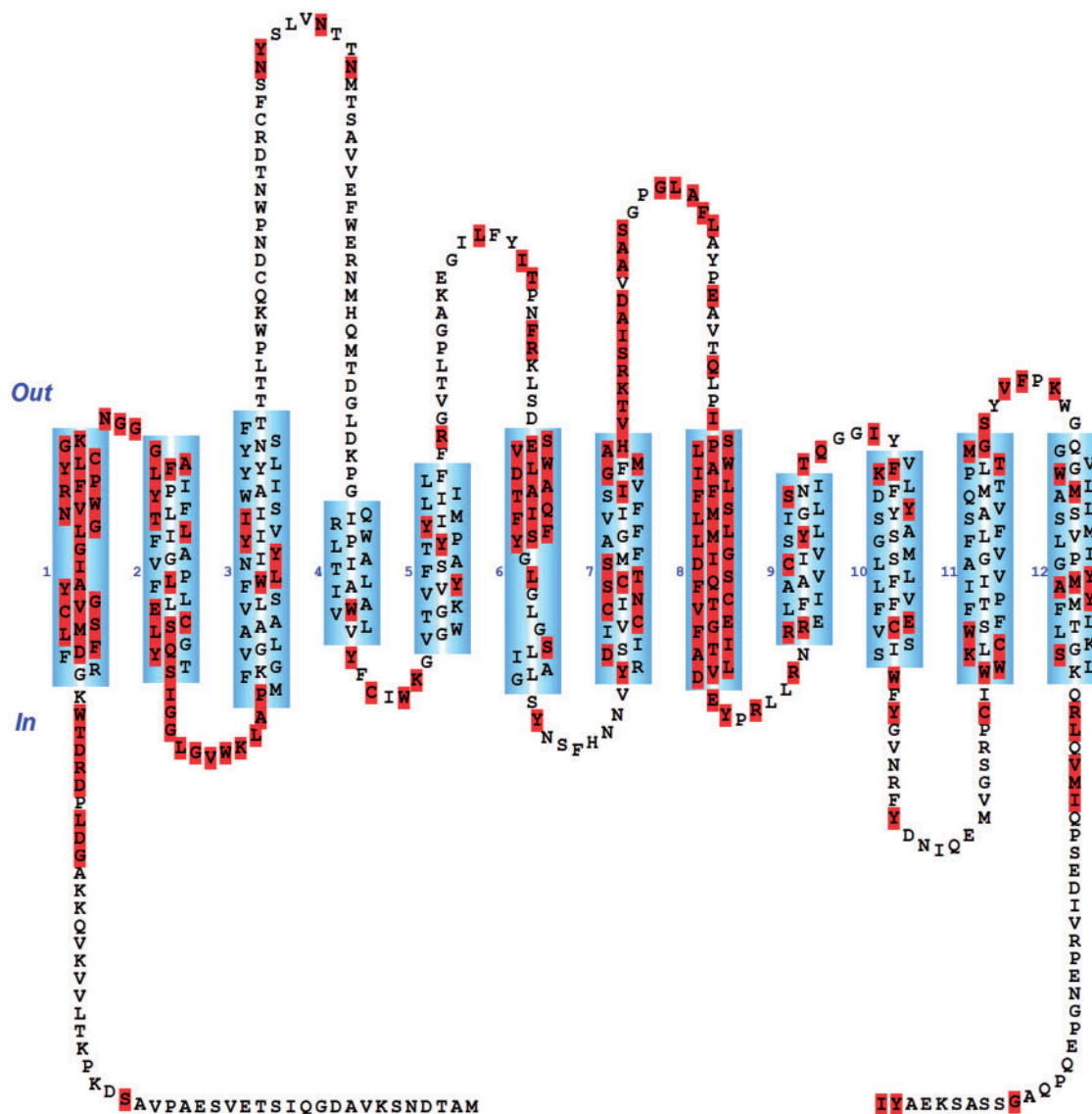


Figure 6. Using a predefined query under the Basic Search Functions menu of the ‘Search’ page, a partial screenshot of the resulting page is shown in which all rat GAT1 residues subjected to substitution mutagenesis are highlighted. As of the last update of GATMD, 218 of 599 rat GAT1 residues have been subjected to substitution mutagenesis. Each residue of the topology diagram (highlighted or not) has a hyperlink which may be used to retrieve the associated experimental evidence for that residue.

will return all sulfhydryl modification experiments performed in the *Xenopus* oocyte expression system. Advanced search commands may also be combined with simple search words/phrases; e.g. {Expression System(hela)} AND ‘similar to WT’. In this manner, complex queries can be formulated to search for very specific data related to mutants of GAT1.

Numerical searches may be evoked for formal parameters that are designed to contain integer or real numbers. For example, whereas {GABA_Km} will return all experiments that have reported a value for the steady-state

GABA K_m (i.e. concentration of GABA needed to evoke 50% of the maximum transport rate), {GABA_Km(>25)} {GABA_KmUnit(uM)} will return all experiments that have reported a GABA K_m value >25 μM . As another example, {InhibitorAbbreviatedName(no-711)} AND {InhibitorConcentration_o(>=1)} AND {InhibitorConcentration_oUnit(uM)} will return all experiments in which the concentration of the specific inhibitor of GAT1, NO-711, was equal to or greater than 1 μM . Finally, {Mutant_UptakeRate(<=5)}{UptakeRateUnit(%)} will return all experiments in which the rate of uptake

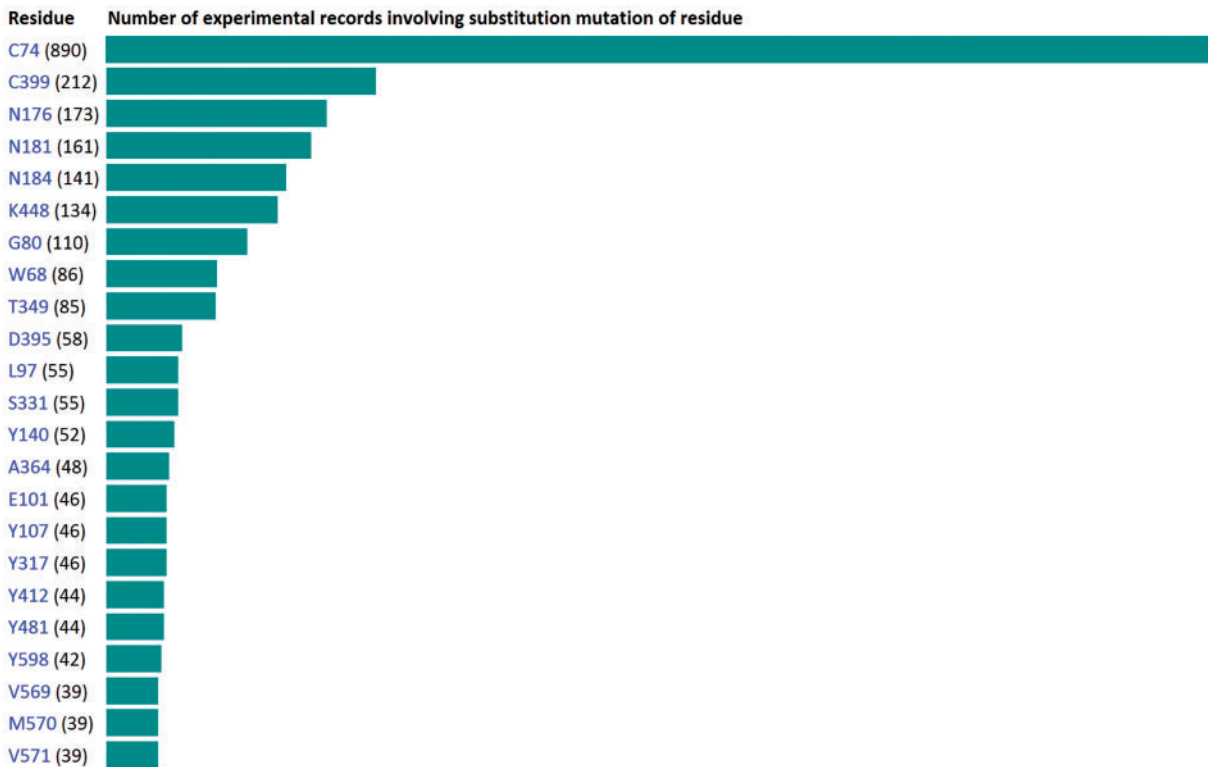


Figure 7. A partial screenshot of the ‘Statistics’ page shows the frequency with which each of the 599 residues of GAT1 has been subjected to substitution mutagenesis. In addition, general database statistics are shown on this page and reveal a quick glance of mutagenesis experiments, as well as the expression systems and experimental approaches used to study the GABA transporters. Links provided on this page utilize advanced search queries to retrieve the corresponding experimental records.

mediated by the mutant transporter was $\leq 5\%$ of that of the WT transporter.

General database statistics are shown on the ‘Statistics’ page and reveal a quick summary of mutagenesis experiments, as well as the expression systems and experimental approaches used to study the GABA transporters. Figure 7 is a partial screenshot of the ‘Statistics’ page highlighting the frequency with which each of the 599 residues of GAT1 has been subjected to substitution mutagenesis. Hyperlinks provided on this page enable the user to quickly access the corresponding experimental records. These links utilize advanced search commands to query the database and, for clarity, these commands are shown in hint boxes for each of the links.

Discussion

The complexity, enormity and rapid growth of GABA transporter structure–function information have led us to envisage, design, and implement a comprehensive, publicly accessible data resource for GABA transporter mutagenesis studies (GATMD). Currently, the database contains approximately 100 000 annotated parameters, collected from 3360

experimental records of 52 published studies. While very useful databases exist that present broad genomics and/or pharmacogenomics information about transporter families and their members (11, 60–67), GATMD contains comprehensive experiment-specific information, where search queries can reveal biochemical, functional, and pharmacological consequences of structure–function investigations. To the best of our knowledge, GATMD is the first database of its kind for any membrane transport protein. Another publicly available database that contains data for selected members of the neurotransmitter:sodium symporter family provides an extensive catalog of published mutagenesis papers for several family members, where the papers are linked to the residues investigated (68). Moreover, the database provides limited functional information on the dopamine, norepinephrine and serotonin transporters (68).

Our curation effort involves manual annotation of biochemical, functional and pharmacological data reported on GAT1 mutants. Most of the curated data contained in GATMD were found in the published papers in textual form, or were embedded in image graphics. While automated methods have been introduced to extract information about the existence of mutants of a given protein

(69, 70), current content annotation and text mining technologies do not allow us to use automation to reliably extract experiment-specific, structure–function information from published papers. Therefore, at least for the time being, we have elected to employ manual annotation of data. One disadvantage of manual annotation is that the database structure and its contents have to co-evolve manually with emerging datasets. Clearly, the most significant disadvantage of manual annotation is the time involved in delving deep into the published papers. This is highlighted by the fact that nearly half of the two years we have committed to this project has been devoted to annotation. Nevertheless, we believe that GATMD will be a useful resource for all investigators interested in electrogenic Na^+ -coupled transporters.

First and foremost, GATMD can be used as a general tool by experts as well as non-experts to perform a literature review on the GABA transporters. For example, no prior user knowledge is required to determine whether or not a specific residue has been subjected to mutagenesis, or to locate functional site residues that have been identified by homology modeling using the crystal structure of LeuT_{Aa} as a template. As other examples, using the ‘Basic Search Functions’ menu, it is simple to identify all mutants that are incapable of mediating Na^+ - and Cl^- -coupled GABA uptake, those that have lost sensitivity to known transport inhibitors, or those that play a role in transporter trafficking to the plasma membrane. To reveal additional experiment-specific information, these same searches may be performed by using the advanced search queries: `{MutantTransporterMediatesGABAUptake(false)}`, `{MutantTransporterExhibitsSensitivityToGAT1Inhibitors(false)}` and `{MutantTransporterIsTraffickedToThePlasmaMembrane(false)}`, respectively. Moreover, very specific questions can be asked by experts that can form the basis for planning future experiments. For example, if an investigator is interested in studying the partial reactions of the transport cycle that are responsible for voltage-induced pre-steady-state charge movements, then the ideal mutant to examine is one which exhibits pre-steady-state charge movements but not GABA transport because, in such a mutant, the transporter conformational changes are restricted to those involved in pre-steady-state charge movements, and not Na^+ - and Cl^- -coupled GABA translocation. In this case, using the mutant phenotype search available under the ‘Basic Search Functions’ menu, it can be seen that GAT1 Y140F would be the right choice for such studies.

GATMD may also be used as a tool to generate testable hypotheses, which may then be subjected to experimental verification. For example, if GATMD had been available prior to the publication of the high-resolution crystal structure of LeuT_{Aa} in 2005 (41), the advanced search query, `{Mutant_UptakeRate(<5)}{UptakeRateUnit(%)}`,

would have revealed seventeen residues to be critical for GABA transport. They are R44, W47, Y60, G63, W68, R69, Y86, E101, Y140, W222, Y226, W230, Y309, G343, L363, E467 and W500. Remarkably, the crystal structure identified seven of these residues as functional sites (R44, W47, Y60, R69, Y86, Y140 and Y309), and showed two others (G63 and W68) to be immediately adjacent to a functional site residue. Thus, this search, which retrieves data from ten papers in the database, would have represented at least a 50% success rate in predicting important residues. As the crystal structure does not reveal the full molecular dynamics of the transport cycle, it is clear that the remaining residues not identified by the crystal structure should be pursued in future real-time structure–function studies as they may be involved in critical conformational changes of the transport cycle.

We also hope that our database will provide some encouragement for future authors to include all of the necessary experimental information related to their findings. Our annotation efforts revealed many instances where key experimental parameters (e.g., substrate concentration, temperature, etc.) had not been specified. While in many cases, this had no bearing on the mechanistic conclusions reached by the authors, these omissions do provide an obstacle for those wishing to repeat the published experiments. From the perspective of the curator, these gaps in information leave important database fields blank, which in turn render otherwise important findings less amenable to targeted search queries.

Future implementations of GATMD will utilize three-dimensional homology models of GAT1 based on the crystal structure of the related bacterial transporter, LeuT_{Aa} (41), and will make spatial information available about residues as well as their interactions with neighboring residues (e.g. within a specified radius). Such information will further facilitate our structure–function understanding of this important member of the neurotransmitter:sodium symporter family of transporters. Future implementations will also extend the database to the other three GABA transporter isoforms (BGT-1, GAT-2, GAT-3). Surprisingly, there are only a few structure–function studies of these isoforms.

Acknowledgements

We thank Mr Jaison J. Omoto for providing insightful comments regarding this article and the GATMD web interface, Dr Craig A. Rich for helpful discussions on the overall architecture of GATMD and Ms Nemone Muster for providing input during the initial stages of database development and data annotation.

Funding

U.S. National Institute of General Medical Sciences at the National Institutes of Health (SC1GM086344 awarded to S.E.). Funding for open access charge: U.S. National Institute of General Medical Sciences at the National Institutes of Health (SC1GM086344 awarded to S.E.).

Conflict of interest statement. None declared.

References

- Borden, L.A. (1996) GABA transporter heterogeneity: pharmacology and cellular localization. *Neurochem. Int.*, **29**, 335–356.
- Nelson, N. (1998) The family of Na⁺/Cl⁻ neurotransmitter transporters. *J. Neurochem.*, **71**, 1785–1803.
- Dalby, N.O. (2003) Inhibition of γ -aminobutyric acid uptake: anatomy, physiology and effects against epileptic seizures. *Eur. J. Pharmacol.*, **479**, 127–137.
- Richerson, G.B. and Wu, Y. (2003) Dynamic equilibrium of neurotransmitter transporters: Not just for reuptake anymore. *J. Neurophysiol.*, **90**, 1363–1374.
- Conti, F., Minelli, A. and Melone, M. (2004) GABA transporters in the mammalian cerebral cortex: localization, development and pathological implications. *Brain Res. Brain Res. Rev.*, **45**, 196–212.
- Richerson, G.B. and Wu, Y. (2004) Role of the GABA transporter in epilepsy. *Adv. Exp. Med. Biol.*, **548**, 76–91.
- Madsen, K.K., Clausen, R.P., Larsson, O.M. et al. (2009) Synaptic and extrasynaptic GABA transporters as targets for anti-epileptic drugs. *J. Neurochem.*, **109**(Suppl. 1), 139–144.
- Madsen, K.K., White, H.S. and Schousboe, A. (2010) Neuronal and non-neuronal GABA transporters as targets for antiepileptic drugs. *Pharmacol. Ther.*, **125**, 394–401.
- Busch, W. and Saier, M.H. Jr (2002) The transporter classification (TC) system, 2002. *Crit. Rev. Biochem. Mol. Biol.*, **37**, 287–337.
- Chen, N.H., Reith, M.E. and Quick, M.W. (2004) Synaptic uptake and beyond: the sodium- and chloride-dependent neurotransmitter transporter family SLC6. *Pflugers Arch.*, **447**, 519–531.
- Saier, M.H. Jr, Tran, C.V. and Barabote, R.D. (2006) TCDB: the Transporter Classification Database for membrane transport protein analyses and information. *Nucleic Acids Res.*, **34**, D181–D186.
- Livesay, D.R., Kidd, P.D., Eskandari, S. et al. (2007) Assessing the ability of sequence-based methods to provide functional insight within membrane integral proteins: a case study analyzing the neurotransmitter/Na⁺ symporter family. *BMC Bioinformatics*, **8**, 397.
- Radian, R. and Kanner, B.I. (1983) Stoichiometry of sodium- and chloride-coupled γ -aminobutyric acid transport by synaptic plasma membrane vesicles isolated from rat brain. *Biochemistry*, **22**, 1236–1241.
- Kanner, B.I. and Schuldiner, S. (1987) Mechanism of transport and storage of neurotransmitters. *CRC Crit. Rev. Biochem.*, **22**, 1–38.
- Keynan, S. and Kanner, B.I. (1988) γ -Aminobutyric acid transport in reconstituted preparations from rat brain: coupled sodium and chloride fluxes. *Biochemistry*, **27**, 12–17.
- Loo, D.D., Eskandari, S., Boorer, K.J. et al. (2000) Role of Cl⁻ in electrogenic Na⁺-coupled cotransporters GAT1 and SGLT1. *J. Biol. Chem.*, **275**, 37414–37422.
- Gonzales, A.L., Lee, W., Spencer, S.R. et al. (2007) Turnover rate of the γ -aminobutyric acid transporter GAT1. *J. Membr. Biol.*, **220**, 33–51.
- Wu, Y., Wang, W., Díez-Sampedro, A. et al. (2007) Nonvesicular inhibitory neurotransmission via reversal of the GABA transporter GAT-1. *Neuron*, **56**, 851–865.
- Matthews, E. Jr, Rahnama-Vaghef, A. and Eskandari, S. (2009) Inhibitors of the γ -aminobutyric acid transporter 1 (GAT1) do not reveal a channel mode of conduction. *Neurochem. Int.*, **55**, 732–740.
- Guastella, J., Nelson, N., Nelson, H. et al. (1990) Cloning and expression of a rat brain GABA transporter. *Science*, **249**, 1303–1306.
- Nelson, H., Mandiyan, S. and Nelson, N. (1990) Cloning of the human brain GABA transporter. *FEBS Lett.*, **269**, 181–184.
- Borden, L.A., Smith, K.E., Hartig, P.R. et al. (1992) Molecular heterogeneity of the γ -aminobutyric acid (GABA) transport system. Cloning of two novel high affinity GABA transporters from rat brain. *J. Biol. Chem.*, **267**, 21098–21104.
- Borden, L.A., Dhar, T.G., Smith, K.E. et al. (1994) Cloning of the human homologue of the GABA transporter GAT-3 and identification of a novel inhibitor with selectivity for this site. *Recept. Channels*, **2**, 207–213.
- Borden, L.A., Smith, K.E., Gustafson, E.L. et al. (1995) Cloning and expression of a betaine/GABA transporter from human brain. *J. Neurochem.*, **64**, 977–984.
- Clark, J.A., Deutch, A.Y., Gallipoli, P.Z. et al. (1992) Functional expression and CNS distribution of a β -alanine-sensitive neuronal GABA transporter. *Neuron*, **9**, 337–348.
- Lopez-Corcuera, B., Liu, Q.R., Mandiyan, S. et al. (1992) Expression of a mouse brain cDNA encoding novel γ -aminobutyric acid transporter. *J. Biol. Chem.*, **267**, 17491–17493.
- Yamauchi, A., Uchida, S., Kwon, H.M. et al. (1992) Cloning of a Na⁺- and Cl⁻ dependent betaine transporter that is regulated by hypertonicity. *J. Biol. Chem.*, **267**, 649–652.
- Liu, Q.R., Mandiyan, S., Nelson, H. et al. (1992) A family of genes encoding neurotransmitter transporters. *Proc. Natl Acad. Sci. USA*, **89**, 6639–6643.
- Liu, Q.R., López-Corcuera, B., Mandiyan, S. et al. (1993) Molecular characterization of four pharmacologically distinct γ -aminobutyric acid transporters in mouse brain. *J. Biol. Chem.*, **268**, 2106–2112.
- Sacher, A., Nelson, N., Ogi, J.T. et al. (2002) Presteady-state and steady-state kinetics and turnover rate of the mouse γ -aminobutyric acid transporter (mGAT3). *J. Membr. Biol.*, **190**, 57–73.
- Cohen-Kfir, E., Lee, W., Eskandari, S. et al. (2005) Zinc inhibition of γ -aminobutyric acid transporter 4 (GAT4) reveals a link between excitatory and inhibitory neurotransmission. *Proc. Natl Acad. Sci. USA*, **102**, 6154–6159.
- Karakossian, M.H., Spencer, S.R., Gomez, A.Q. et al. (2005) Novel properties of a mouse γ -aminobutyric acid transporter (GAT4). *J. Membr. Biol.*, **203**, 65–82.
- Blakely, R.D., Berson, H.E., Freneau, R.T. Jr et al. (1991) Cloning and expression of a functional serotonin transporter from rat brain. *Nature*, **354**, 66–70.
- Hoffman, B.J., Mezey, E. and Brownstein, M.J. (1991) Cloning of a serotonin transporter affected by antidepressants. *Science*, **254**, 579–580.
- Kilty, J.E., Lorang, D. and Amara, S.G. (1991) Cloning and expression of a cocaine-sensitive rat dopamine transporter. *Science*, **254**, 578–579.

36. Pacholczyk, T., Blakely, R.D. and Amara, S.G. (1991) Expression cloning of a cocaine- and antidepressant-sensitive human noradrenaline transporter. *Nature*, **350**, 350–354.
37. Shimada, S., Kitayama, S., Lin, C.L. et al. (1991) Cloning and expression of a cocaine-sensitive dopamine transporter complementary DNA. *Science*, **254**, 576–578.
38. Usdin, T.B., Mezey, E., Chen, C. et al. (1991) Cloning of the cocaine-sensitive bovine dopamine transporter. *Proc. Natl Acad. Sci. USA*, **88**, 11168–11171.
39. Guastella, J., Brecha, N., Weigmann, C. et al. (1992) Cloning, expression, and localization of a rat brain high-affinity glycine transporter. *Proc. Natl Acad. Sci. USA*, **89**, 7189–7193.
40. Liu, Q.R., Nelson, H., Mandiyan, S. et al. (1992) Cloning and expression of a glycine transporter from mouse brain. *FEBS Lett.*, **305**, 110–114.
41. Yamashita, A., Singh, S.K., Kawate, T. et al. (2005) Crystal structure of a bacterial homologue of Na⁺/Cl⁻ dependent neurotransmitter transporters. *Nature*, **437**, 215–223.
42. Singh, S.K., Yamashita, A. and Gouaux, E. (2007) Antidepressant binding site in a bacterial homologue of neurotransmitter transporters. *Nature*, **448**, 952–956.
43. Singh, S.K., Piscitelli, C.L., Yamashita, A. et al. (2008) A competitive inhibitor traps LeuT in an open-to-out conformation. *Science*, **322**, 1655–1661.
44. Zhou, Z., Zhen, J., Karpowich, N.K. et al. (2007) LeuT-desipramine structure reveals how antidepressants block neurotransmitter reuptake. *Science*, **317**, 1390–1393.
45. Zhou, Y., Zomot, E. and Kanner, B.I. (2006) Identification of a lithium interaction site in the γ -aminobutyric acid (GABA) transporter GAT-1. *J. Biol. Chem.*, **281**, 22092–22099.
46. Forrest, L.R., Tavoulari, S., Zhang, Y.W. et al. (2007) Identification of a chloride ion binding site in Na⁺/Cl⁻ dependent transporters. *Proc. Natl Acad. Sci. USA*, **104**, 12761–12766.
47. Zomot, E., Bendahan, A., Quick, M. et al. (2007) Mechanism of chloride interaction with neurotransmitter:sodium symporters. *Nature*, **449**, 726–730.
48. Forrest, L.R., Zhang, Y.W., Jacobs, M.T. et al. (2008) Mechanism for alternating access in neurotransmitter transporters. *Proc. Natl Acad. Sci. USA*, **105**, 10338–10343.
49. Shi, L., Quick, M., Zhao, Y. et al. (2008) The mechanism of a neurotransmitter:sodium symporter-inward release of Na⁺ and substrate is triggered by substrate in a second binding site. *Mol. Cell*, **30**, 667–677.
50. Tavoulari, S., Forrest, L.R. and Rudnick, G. (2009) Fluoxetine (Prozac) binding to serotonin transporter is modulated by chloride and conformational changes. *J. Neurosci.*, **29**, 9635–9643.
51. Zhao, Y., Quick, M., Shi, L. et al. (2010) Substrate-dependent proton antiport in neurotransmitter:sodium symporters. *Nat. Chem. Biol.*, **6**, 109–116.
52. Kanner, B.I. and Zomot, E. (2008) Sodium-coupled neurotransmitter transporters. *Chem. Rev.*, **108**, 1654–1668.
53. Abramson, J. and Wright, E.M. (2009) Structure and function of Na⁺-symporters with inverted repeats. *Curr. Opin. Struct. Biol.*, **19**, 425–432.
54. Forrest, L.R. and Rudnick, G. (2009) The rocking bundle: a mechanism for ion-coupled solute flux by symmetrical transporters. *Physiology*, **24**, 377–386.
55. Gouaux, E. (2009) Review. The molecular logic of sodium-coupled neurotransmitter transporters. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, **364**, 149–154.
56. Krishnamurthy, H., Piscitelli, C.L. and Gouaux, E. (2009) Unlocking the molecular secrets of sodium-coupled transporters. *Nature*, **459**, 347–355.
57. Murer, H., Forster, I. and Biber, J. (2004) The sodium phosphate cotransporter family SLC34. *Pflugers Arch.*, **447**, 763–767.
58. Wright, E.M. and Turk, E. (2004) The sodium/glucose cotransport family SLC5. *Pflugers Arch.*, **447**, 510–518.
59. Beuming, T., Shi, L., Javitch, J.A. et al. (2006) A comprehensive structure-based alignment of prokaryotic and eukaryotic neurotransmitter/Na⁺ symporters (NSS) aids in the use of the LeuT structure to probe NSS structure and function. *Mol. Pharmacol.*, **70**, 1630–1642.
60. Yan, Q. and Sadée, W. (2000) Human membrane transporter database: A Web-accessible relational database for drug transport studies and pharmacogenomics. *AAPS PharmSci.*, **2**, E20.
61. Ren, Q., Kang, K.H. and Paulsen, I.T. (2004) TransportDB: a relational database of cellular membrane transport systems. *Nucleic Acids Res.*, **32**, D284–D288.
62. Ren, Q., Chen, K. and Paulsen, I.T. (2007) TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. *Nucleic Acids Res.*, **35**, D274–D279.
63. Fichant, G., Basse, M.J. and Quentin, Y. (2006) ABCdb: an online resource for ABC transporter repertoires from sequenced archaeal and bacterial genomes. *FEMS Microbiol. Lett.*, **256**, 333–339.
64. Pharmacogenetics of Membrane Transporters Database. <http://pharmacogenetics.ucsf.edu/> (11 November 2010, date last accessed).
65. Transporter Classification Database (TCDB). <http://www.tcdb.org/> (11 November 2010, date last accessed).
66. TransportDB. <http://www.membranetransport.org/> (11 November 2010, date last accessed).
67. Archaeal and Bacterial ABC Transporter Database (ABCDB). <http://www-abcdb.biotoul.fr/> (11 November 2010, date last accessed).
68. TRansporter Annotation Center (TRAC). <http://icb.med.cornell.edu/trac/> (11 November 2010, date last accessed).
69. Horn, F., Lau, A.L. and Cohen, F.E. (2004) Automated extraction of mutation data from the literature: application of MuteXt to G protein-coupled receptors and nuclear hormone receptors. *Bioinformatics*, **20**, 557–568.
70. Lee, L.C., Horn, F. and Cohen, F.E. (2007) Automatic extraction of protein point mutations using a graph bigram association. *PLoS Comput. Biol.*, **3**, e16.