DRUG DISCOVERY TODAY BIOSILICO

In silico strategies for modeling membrane transporter function

Cheng Chang, Abhijit Ray and Peter Swaan

Transporter proteins facilitate the transfer of solutes across the cell membrane and have an intricate role in drug absorption, distribution and excretion. Because of their substrate promiscuity, several transporters represent viable pharmacological targets for enhancing drug absorption, preventing drug toxicity or facilitating localized tissue delivery. However, the slow emergence of high-resolution structures for these proteins has hampered the intelligent design of transporter substrates. Nonetheless, currently available functional, as well as structural, data provide an attractive scaffold for generating fusion models that merge substrate-based SARs and protein-based homology structures. The resultant models offer features that extend single modality paradigms in predictive function.

Transporters are polytopic membrane proteins that are indispensable to the cellular uptake and homeostasis of many essential nutrients. Furthermore, it is now known that many drugs are transporter substrates, thus substantiating the intricate involvement of these proteins in all facets of drug absorption, tissue distribution, excretion and toxicity, as well as drug pharmacokinetics and pharmacodynamics – all widely accepted crucial parameters that a new drug must attain to survive the drug discovery and development pipeline. Clearly, a detailed understanding of transporter structure and function, as well as mechanism, would be of considerable value in drug discovery. Despite the significance of membrane transporters in drug discovery, most are poorly characterized at the atomic level because of difficulties associated with expressing [1] and crystallizing membrane proteins [2]. To date, there are only 86 membrane protein structures available (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html), of which only two are from the major facilitator superfamily (MFS) and three are from the ATP-binding cassette (ABC) superfamily. As a result, knowledge of transporter structure and mechanism has lagged far behind the insight into their biochemical and functional properties. *In silico* methodologies are a useful tool to circumvent the difficulties associated with traditional crystallization techniques and to fill the gap between knowledge of transporter protein structure and properties. This trend is confirmed by an increasing number of publications that project transporter structure and mechanism using single-approach and combinatorial *in silico* strategies that are consecutively validated by empirical methods.

Here, applications of transporter proteins in protein-based methods are discussed, including comparative modeling in which the transporter structure is directly modeled to gain insight into the transport process (Figure 1). Also reviewed are ligand-based methods, such as QSARs, in which transporter substrate specificity is analyzed in the absence of protein structure.

Comparative modeling

The primary amino acid sequences of many transporters are known and solution of their secondary

Cheng Chang
Biophysics Program, Ohio State University, Columbus, Ohio, USA
Abhijit Ray
Peter Swaan*
Department of Pharmaceutical Sciences, University of Maryland, Baltimore, Maryland, USA
*e-mail: pswaan@rx.umaryland.edu

1359-6446/05/$ – see front matter PII: S1359-6446(05)03429-X

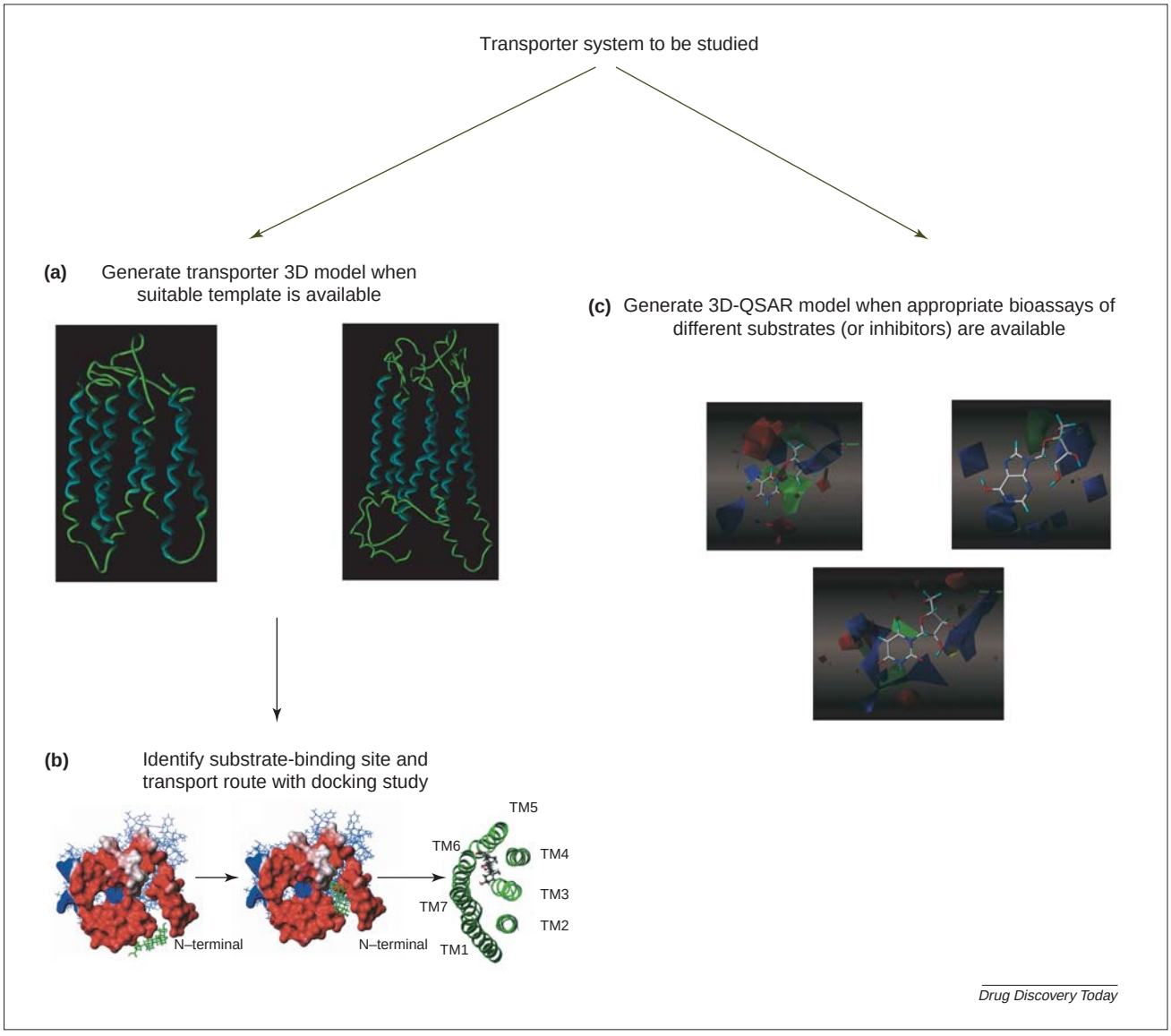

FIGURE 1

General protocol for modeling of transporter proteins. (a) Comparative modeling of human ASBT 3D model (right) based on its template bacteriorhodopsin crystal structure (left). The template and the target transporter must have identical transmembrane topology and ≥20% sequence identity. Although not essential, a template structure determined by X-ray crystallization is preferred. **(b)** Putative binding or ligand–protein interaction domains can be determined through a docking study. This figure illustrates three potential high-affinity binding domains on ASBT for cholic acid, a natural substrate. **(c)** General example of a 3D-QSAR modeling approach. Here, the CoMFA coefficient contour maps surround a model substrate for hCNT1 (upper left), hCNT2 (upper right) and hENT1 (bottom). For hCNT1 and hENT1, steric contour maps indicate that greater inhibition is correlated with reduced steric hindrance near green contours and increased steric hindrance near yellow. Electrostatic contours suggest that more negative electrostatic charge near blue and more positive charge near red will increase biological activity. For hCNT2, the contours of the hydrogen bond acceptor field map are shown in yellow and green and those of the hydrogen bond donor field map are shown in red and blue. Greater inhibition is correlated with weaker H-bond acceptor near green, stronger H-bond acceptor near yellow, stronger H-bond donor near blue and weaker H-bond donor near red. The number of substrates (or inhibitors) is ≥10, the spread of biological activity is >100 (i.e. two log unit differential) and all substrates are binding to the same site.

structures is facilitated by bioinformatics tools, such as hydropathy plotting, and, preferably, by a combination of experimental verification methods, such as *N*-glycosylation analysis or epitope insertion scanning [3,4]. Although only a few transporter proteins have been analyzed by X-ray crystallography and afforded high-resolution 3D information, new structures are emerging at a steady rate, which opens the opportunity for generating homology or comparative models. Typically, when two proteins have adequate sequence identity (generally >60%) and the experimentally determined 3D structure of one of these proteins is known, comparative protein models can be constructed, even when the two proteins are not functionally related [5].

The procedure of comparative modeling is illustrated best with the most widely used program for this approach – Modeller [6]. First, possible structural templates, as well as related sequences, are retrieved using the query

TABLE 1

Summary for membrane transporter comparative modeling studies

Transporter	Template	No. of TMDs	Method	Docked substrates	Refs
ASBT	Bacteriorhodopsin crystal structure	7	Knowledge based	Cholic acid	[3,4]
GLUT1	Helix packing model for LacY	12	Knowledge based	NR	[12]
	LacY crystal structure	12	NR	NR	[16–18]
	GlpT crystal structure	12	Nest Modeller	Glucose, forskolin, phloretin and cytochalasin B	[52,53]
G6PT	GlpT crystal structure	12	Modeller	NR	[11]
EmrE	Biochemical and biophysical study results	8	Custom method [54]	TPP^{+}	[54,55]
P-gp	MsbA crystal structure (*Escherichia coli*)	12	Modeller	NR	[19]
	MsbA crystal structure (*E. coli*) Cross-linking results	12	NR	NR	[20]
MRP1	MsbA crystal structure (*Vibrio cholera*)	17	Modeller[a]	NR	[23]
NhaA	NhaA electron density projection map	12	Knowledge based	Amiloride	[40]
DAT	NhaA model	12	Knowledge based	S-citalopram and cocaine	[56]
SERT	DAT model	12	*In silico* mutation	S-citalopram and cocaine	[57]
NET	SERT model	12	*In silico* mutation	S-citalopram and cocaine	[57]

[a]Only the 12 TMDs that are essential to function were modeled.

Abbreviations: NR, not reported; TPP, tetraphenylphosphonium.

sequence. All template structures are initially aligned to generate a sequence block, which is subsequently aligned to all sequences related to the unknown 3D structure. Extensive manual adjustment of the computer-generated multiple sequence alignment is required to ensure that conserved residues are aligned and to move gaps to the loop regions that connect the transmembrane domains (TMDs): this is particularly important because most amino acid variability is observed in the extramembranous loop regions. Crucially, high homology should not be the sole criterion for manually adjusting the alignment because this might not correspond to high amino acid identity. One or more templates are selected based on sequence similarity and phylogenetic analysis. Modeller generates several models based on the alignment, which are, in turn, evaluated by internal self-consistency checks and external programs that verify structural protein quality, such as ProCheck [7] or WHAT IF [8]. The cycle of template selection, sequence alignment adjustment, modeling and evaluation are repeated until no further improvements to the model are observed.

Application of comparative models to transporters

Comparative modeling techniques now enable the elucidation of 3D structures for a large number of transporters that would otherwise not be available to researchers (Table 1). This area has accelerated significantly with the recent publication of high-resolution crystal structures for two bacterial transporters, the lactose permease LacY (resolution 3.5 Å) [9] and the glycerol-3-phosphate transporter GlpT (resolution 3.3 Å) [10]. Both structures represent polytopic membrane proteins comprising 12 transmembrane segments. Importantly, although LacY and GlpT are functionally different transporters, they share a similar fold. This has encouraged researchers to speculate that all MFS transporters share folding patterns, thus initiating novel opportunities for comparative modeling and these models can be expected to appear in the literature in the near future. For example, Huang and co-workers [11] recently described a comparative model of the human glucose-6-phosphate transporter (G6PT), which is involved in the pathology of the glycogen storage disease type Ib. They identified two positively charged residues, Asp28 and Lys240, as being part of the substrate-binding site. Twenty-eight missenses and two deletion mutations that are associated with the disease were mapped onto the comparative model to provide information on the molecular mechanism of G6PT. Three different disease-causing mechanisms were proposed – mutations that could modify the substrate-binding site, change the N- and C-terminal domain interface or destabilize the protein. These observations agree with the available experimental data on G6PT, providing further support for the comparative modeling approach.

Glucose transporter 1

Because of its role in glucose transport and diabetes, glucose transporter 1 (GLUT1) is perhaps the most extensively studied membrane transporter, yet an atomic level structural

description has yet to emerge. Several *in silico* models have been generated to aid in explaining the wealth of experimental data. Zuniga *et al*. [12] constructed a 3D GLUT1 model using a 'piecemeal strategy': experimentally derived information was used to guide the step-by-step construction of the model, supplemented by comparative modeling of extramembranous loops and molecular dynamics (MD) optimization of the overall protein. MD simulations are ideally suited to optimize the overall conformation of the protein in a simulated membrane environment [13,14]. The final model assisted in clarifying experimental data not used in model building, including the observation of an interaction between Asn288 and glucose during a MD run, which explained the observed tenfold reduction in glucose transport of the mutant Asn288Cys [15]. With the availability of the LacY crystal structure, which is a potential template for GLUT1, a comparative model was generated [16] to illustrate mutagenesis results further. In addition, this model was continuously verified and optimized using new experimental results [17,18]. This exemplifies a new trend of feedback between *in silico* studies with bench experiments – *in silico* models help to illustrate and explain experimental discoveries, whereas experimental results aid in continuously optimizing theoretical models.

ATP-binding cassette transporters

The ABC superfamily of transporters has been extensively studied because of their role in drug efflux, chemoresistance and drug–drug interactions and, as a result, a wealth of experimental information is available. To date, all three available full-length ABC transporter crystal structures are of bacterial origin. However, to further understanding of human ABC transporter mechanisms, a human ABC transporter atomic structure is crucial. The most widely studied ABC transporter, P-glycoprotein (P-gp), was initially modeled using the crystal structure of another ABC protein, the *Escherichia coli* lipid A transporter (MsbA) [19]: this template shares ~30% homology with P-gp. It should be noted that, because of the relatively low resolution (4.5 Å), the original MsbA crystal structure file does not contain coordinates for TMD side chain or backbone atoms (with the exception of C-α) and does not account for 39 out of 582 amino acid residues. Indeed, the C-α coordinates of the nucleotide-binding domains (NBDs) are only partially available (60%). As a result, special emphasis was applied to the modeling of the missing backbone and side-chain atoms. Overall, the resultant model can only be used qualitatively, predominantly because of the missing template information and alignment uncertainties. However, the model provided additional information, revealing that the three intracellular domains of each half of the transporter have distinct and selective interactions with the core domain (which contains the ATP-binding site), the α-domain (i.e. helices α3–α5) and the conserved Q-loop that links the core- and the α-domain. It was also reported that aromatic residues in the internal chamber of the protein were highly conserved and there was a correlation of chamber cavity dimensions and substrate size and charge, which together suggest an important role for the internal chamber in drug transport.

Stenham and colleagues [20] challenged the validity of the *E. coli* MsbA crystal structure. Assessment of the crystal structure of MsbA in relation to their disulfide cross-linking studies and other non-crystallographic experimental evidence suggested that the orientation of *E. coli* MsbA TMD with respect to NBD observed in the crystal structure does not reflect the physiological association. Instead, the P-gp TMD and NBD were comparatively modeled, according to *E. coli* MsbA, independently from each other and the protein was subsequently assembled based on cross-linking data. Additionally, structural data from another bacterial ABC transporter, vitamin B12 transport system permease, BtuC, were incorporated, in particular, the correct orientation of the NBDs. The final model of P-gp contains a consensus NBD:NBD interface and a parallel TMD:TMD interface, which is structurally consistent with chemical cross-linking data, as well as previous electron microscopy data. The physiologically correct orientation of the NBD with respect to the TMDs was again confirmed in the crystal structure of MsbA in *Vibrio cholera* [21].

The two diverging comparative models of P-gp were later reconciled in accordance with a theory proposed by Lee and co-workers [22], who postulated that P-pg could exist in two dynamic conformations – an open conformation [19] and a closed conformation [20]. It was also suggested that by rotating the two monomers toward each other, the *E. coli* MsbA crystal structure (open) could be transformed into the closed conformation.

The *V. cholera* MsbA structure was used to model another major ABC transporter, multidrug-resistance-related protein (MRP) 1 [23]. The model implied that Phe594 forms a hydrophobic pocket with four residues previously identified as being important for activity. The hypothesis that Phe594 would also have a role in MRP1 transport activity was verified by site-directed mutagenesis. This exemplifies that *in silico* models could potentially guide experimental design to identify important functional amino acid residues, providing insight into function and mechanisms of biologically important processes. Using the same template structure, Ecker and colleagues [24] built a comparative model for the bacterial efflux transporter LmrA, which is a close orthologue to human P-gp. The drug-binding site was identified using photoaffinity labeling and a comparative model was constructed to illustrate the results and provide an overall transporter structure to guide further experimental design.

The relatively low-quality of the *E. coli* MsbA structure diminished its usefulness as a template for comparative modeling. With the assumption that molecular modeling could extend the usability of low-resolution crystal structures, Campbell *et al.* [25] optimized the 4.5 Å MsbA crystal

structure to atomic resolution by combining comparative modeling techniques and MD simulation. The missing backbone and side chain atoms were regenerated using MAXSPROUT [26] and Modeller. The coordinates for protein gaps that had readily available templates (e.g. NBD) were generated by comparative modeling. The positions of additional missing C-α were predicted by QUANTA (www.accelrys.com/quanta) and PSIPRED [27], resulting in a model with acceptable stereochemical quality parameters, such as dihedral and torsional angles. MsbA model stability was further verified by a 2 ns MD simulation after the protein was inserted into a membrane-mimetic octane slab with full solvation on both sides. During the MD simulation, significant structural rearrangement in the TMDs was observed, which is corroborated by previous evidence that repacking of the TMDs occurs during the P-gp transport cycle [28]. Based on the simulation data, it was suggested that the crystal structure of the MsbA dimer might not correspond to the MsbA dimer *in vivo* because of crystal packing effects, rather than the *in vivo* dimeric state.

Protein-mediated transport is a dynamic process involving several intermediary conformations. Thus, a static structure might not be adequate to study the molecular mechanisms of transport, but MD simulations could overcome these limitations. For example, to determine the catalytic mechanism of the NBD of ABC transporters, Jones and colleagues [29] performed a 390 ps MD simulation of the bacterial histidine permease, HisP, NBD domain. This simulation revealed that the peptide bond between Phe99 and Gln100 serves as the hinge point in ABC transporters, moving the conserved glutamine in and out of the catalytic site. Key interfaces involved in TMD and NBD communication were also identified, which, together with previous experimental data, could be translated to a detailed catalytic cycle of the ABC transporter superfamily.

Ligand-based methods

In silico approaches based on molecular level transporter models could be used to predict ligand- and inhibitor-binding modes, thus representing a useful tool in the discovery of novel transporter substrates. Although comparative models can be generated successfully for membrane proteins with seven or 12 TMDs, this approach is not yet feasible for proteins without a suitable TMD template. In these cases, techniques that do not require knowledge of transporter structure can be applied, such as ligand-based QSAR. Self-consistent models, such as 3D-QSAR models, are generated to correlate the variations in ligand (or inhibitor) binding affinities with their respective structural features. This approach has been applied successfully to generate pharmacophore and 3D-QSAR models for P-gp [30,31], organic cation transporters [32,33], bile acid transporters, nucleoside transporter [34], human peptide transporter 1 (hPEPT1) and organic anion transporters: a comprehensive overview of current applications is given in Table 2.

Several commercial tools are available to generate 3D-QSAR models. Based on the number of publications citing their use, the three most widely applied programs are perhaps Catalyst (Accelrys), comparative molecular field analysis (CoMFA) [35] and comparative molecular similarity index analysis (CoMSIA; Tripos Associates) [36]: these programs were reportedly used in 92, 515 and 113 publications for Catalyst, CoMFA and CoMSIA, respectively, of which 8, 31 and 10, respectively, explicitly dealt with transporter models. Ekins and Swaan [37] have recently reviewed these approaches in detail, including an overview of common pitfalls, advantages and disadvantages. The application of these programs to clinically relevant transporter systems is discussed here.

Monoamine transporters

Monoamine transporters consist of dopamine transporter (DAT), serotonin transporter (SERT) and norepinephrine transporter (NET), which are ion-coupled secondary transporters with 12 membrane-spanning domains. Because of their importance in substance abuse and depression, monoamine transporters have been studied extensively through QSAR analyses. Recently, Kulkarni and colleagues [38] generated highly correlative [cross-validated regression coefficient (q^2) of 0.695] and predictive [regression coefficient (r^2) of 0.75] CoMFA models for DAT using inhibition data from a set of 71 mazindol analogues [38]. The model identified the importance of the relative orientation of the two hydrophobic groups to the position of a heteroatom of the mazindol analogues, which could serve as a guiding principle in the design of novel analogues with higher DAT binding.

Using a different training set of 76 benztropine analogues, additional CoMFA models for DAT were generated [39]. The diphenyl ether moiety was found to be essential for binding of this particular class of compounds to DAT. In addition, it was noted that the N-8 position of these compounds tolerated various substitutions, and thus a series of N-8 substituted benztropine derivatives was synthesized for testing. As well as showing potent inhibition of dopamine uptake with reduced lipophilicity (a step forward toward increasing bioavailability) these compounds exhibited excellent selectivity of DAT over SERT and NET. Thus, this research provides an elegant example of CoMFA-guided drug discovery of novel, potent DAT inhibitors with favorable physicochemical properties. An interesting next step would be a combined model encompassing benztropine and mazindol analogues, thereby providing coverage for a much broader chemical space.

Despite the inferences that can be made from these ligand-based studies, details of DAT structure remain elusive. To generate an atomic description of the transporter, Ravna and colleagues [40] produced a DAT comparative model based on the electron density projection map of another ion-coupled secondary transporter with 12 TMDs,

TABLE 2

Summary for membrane transporter QSAR studies

Transporter	Model training set (model)	Model features	Correlation	Program	Refs
P-gp	27 inhibitors of digoxin transport (Caco-2)	4 hydrophobes 1 HB acceptor	0.77	Catalyst	[58]
	21 inhibitors of vinblastine binding (CEM/VLB100)	1 hydrophobe 3 aromatic ring features	0.88	Catalyst	[58]
	17 inhibitors of vinblastine accumulation (LLC-PK1)	4 hydrophobes 1 HB acceptor	0.86	Catalyst	[58]
	18 inhibitors of calcein accumulation (LLC-PK1)	2 hydrophobes 1 HB donor 1 aromatic ring feature	0.76	Catalyst	[58]
	16 inhibitors of verapamil binding (Caco-2)	2 hydrophobes 1 HB acceptor 1 aromatic ring feature	0.96	Catalyst	[31]
	9 glucocorticoids exported by L-MDR1	4 hydrophobes 3 HB acceptors	NR	DISCO	[30]
	10 glucocorticoids exported by L-MDR1	Importance of nonpolar bulky group at α6 position	0.99 (0.48[a]) 0.95 (0.41[a])	CoMFA CoMSIA	[30]
hOCT1	22 inhibitors of tetraethylammonium uptake (HeLa)	3 hydrophobes 1 positive ionizable feature	0.86	Catalyst	[32]
	23 inhibitors of tetraethylammonium uptake (HeLa)	5 molecular descriptors correlates best with inhibition	0.95	Cerius	[32]
hOCT2	30 inhibitors of TEA transport	Hydrophobic interaction most important	0.81	Cerius	[32]
	31 inhibitors of TEA transport	Molecular size and shape important OCT2 binding site could potentially be multispecific	0.97 (0.60[a])	CoMFA	[32]
hCNT1	27 nucleoside analogues inhibiting thymidine transport	2 hydrophobes 3 HB acceptors	NR	DISCO	[34]
	35 nucleoside analogues inhibiting thymidine transport	Steric and electrostatic interaction equally important	0.98 (0.65[a])	CoMFA	[34]
hCNT2	13 nucleoside analogues inhibiting inosine transport	2 hydrophobes 3 HB acceptors	NR	DISCO	[34]
	32 nucleoside analogues inhibiting inosine transport	Hydrogen bonding interaction determines inhibition	0.83 (0.52)	CoMFA	[34]
hENT1	27 nucleoside analogues inhibiting uridine transport	2 hydrophobes 2 HB acceptors	NR	DISCO	[34]
	39 nucleoside analogues inhibiting uridine transport	Steric and electrostatic interaction equally important	1.00 (0.74)	CoMFA	[34]
ASBT	17 bile acid reabsorption inhibitors	3 hydrophobes 1 HB acceptor 1 HB donor	0.94	Catalyst	[44]
	25 bile acid analogues	Tighter fit at ring system attached hydroxyl groups Sterically more favorable at C17 position Electrostatic interaction mainly located at C24–C27 region	0.96 (0.63[a])	CoMFA	[45]
DAT	50 mazindol analogues inhibiting dopamine transport	Sterically favorable around heterocyclic ring A and ring D Electrostatic interaction principally around ring D	0.90 (0.70[a])	CoMFA	[38]
	76 benztropine analogues inhibiting dopamine transport	Sterically favorable at the diphenylmethoxy terminus Electrostatic interaction is mainly located toward the diphenylmethoxy group	0.92 (0.63[a])	CoMFA	[39]

Continued on following page

TABLE 2 (continued)

Summary for membrane transporter QSAR studies

Transporter	Model training set (model)	Model features	Correlation	Program	Refs
SERT	12 imipramine analogues	Distance between fused ring and amine important for activity	0.53	CoMFA	[59]
	11 antidepressants	Distance between fused ring and amine important for activity	0.55	CoMFA	[59]
	10 fluoxetine analogues	Aromatic ring conformation determines potency	0.60	CoMFA	[59]
BBB choline transporter	33 inhibitors to choline uptake into brain	Positive charge Sterically favorable area around alkyl chain segment of choline and *N*-methyl group Binding pocket is larger than choline molecule	0.95 (0.47[a]) 0.85 (0.37[a])	CoMFA CoMSIA	[60]
PEPT1	79 dipeptide-type substrates with affinity data	Electrostatic, lipophilic and HB donor interaction important	0.90 (0.64[a]) 0.91 (0.78[a])	CoMFA CoMSIA	[61]

[a]Numbers in parentheses indicate cross-validated r^2 values, commonly referred to as q^2 values.
Abbreviations: BBB, blood–brain barrier; HB, hydrogen bond; MDR, multidrug resistance; NR, not reported; OCT, organic cation transporter; TEA, triethylamine.

Na^+/H^+-antiporter (NhaA), and site-directed mutagenesis data on DAT. In addition, cocaine and its analogue, (–)-2β-carbomethoxy-3β-(4-fluorophenyl)tropane (CFT), were docked to this model. Based on this model, it was suggested that Asp79, Tyr252 and Tyr274 are the primary cocaine binding residues and a dopamine transport mechanism involving transmembrane helices 1, 3, 4, 5, 7 and 11 was proposed. This study is unique in that a synergistic combination of comparative modeling and ligand-based design was applied; thus, additional insight was gained into the monoamine transport process, which will ultimately lead to improved streamlining of drug discovery for the treatment of substance abuse and depression.

Nucleoside transporters

Nucleoside transporters can be categorized broadly into equilibrative nucleoside transporters (ENT) and sodium-dependent concentrative nucleoside transporters (CNT). As their name implies, nucleoside transporters have an important role in the physiological and pharmacological activity and disposition of nucleosides and nucleoside drugs. The ENT homologues ENT1 (nitrobenzylthioinosine-sensitive) and ENT2 (nitrobenzylthioinosine-insensitive) are ubiquitously expressed. The CNT homologue CNT1 (pyrimidine-specific) is primarily expressed in epithelia tissues, including intestine, kidney and liver, whereas CNT2 (purine-specific) and CNT3 have more generalized distributions [41]. In the absence of high-resolution structural data for these transporters, Chang *et al*. [34] sought to generate and validate pharmacophore and 3D-QSAR models for the three different homologues of nucleoside transporters that are expressed in the human small intestine: hCNT1, hCNT2 and hENT1 [34]. The purpose of this study was to develop a cell-based pharmacokinetic model to explain intestinal nucleoside processing. The resulting models successfully discriminated the subtly different substrate preferences of the three homologous transporters (Figure 1). Furthermore, they identified the importance of steric and electrostatic interactions for hCNT1 and hENT1 inhibition, whereas hydrogen bonding was the dominant force in interactions between ligands and hCNT2. The hENT1 model was externally validated by successfully predicting affinity of p38 mitogen-activated protein kinase inhibitors for hENT1 [42]. These models should assist the design of high-affinity nucleoside transporter inhibitors and substrates for anticancer and antiviral therapies.

Combination *in silico* approaches

In silico methods represent one of numerous useful tools to study the transport mechanism and substrate affinity requirements of membrane transporters. It efficiently fills the gap between our knowledge of atomic level transporter structural mechanisms and *in vitro* transporter properties derived by biochemical experiments. This type of hybrid approach will eventually lead to the discovery of safer and more efficient drugs by targeting transporters or, in the case of efflux pumps, avoiding them for increased cellular permeability. Combining the results of an array of computational methods to provide insight into the complexities of transporter SAR is rapidly becoming an important approach to study transporter function.

Despite the utility of *in silico* approaches in the study of membrane transporters, investigators are cautioned to avoid overextending model interpretation or extrapolation. For example, the quality of a comparative model is limited by the validity of sequence alignment and the quality of template structure. Sequence homology between transporter families is usually too low to build a robust model with confidence. To overcome these limitations, researchers have opted to forego classical sequence homology rules that have been applied to soluble proteins and demanded >60% sequence identity for the construction of a reliable

model and >80% identity for a high-quality model. Instead, combination approaches are now actively developed that focus on shared membrane topology, satisfactory alignment of transmembrane regions and comparative modeling of extracellular protein segments. For such models to be acceptable, they have to be critically evaluated by biochemical data. By contrast, QSAR models are devoid of information on substrate–receptor interactions at the molecular level, although they could be used to infer such interactions. A recent development is the combination of 3D-QSAR models with comparative models to extract simultaneous information from substrate affinity data, as well as atomic level binding site models. Each technique is assumed to validate the other by fitting the QSAR model into the binding domain of its target protein. These combination approaches are highly synergistic and provide information beyond the individual models [43].

Combination modeling of the apical sodium-dependent bile acid transporter

Apical sodium-dependent bile acid transporter (ASBT) has a crucial role in lipid and cholesterol homeostasis and is a potential drug target for hypercholesterolemia (i.e. elevated plasma cholesterol levels), as well as a drug-transporting vector for increasing bioavailability. As a result, ASBT has been studied extensively using biochemical methods and *in silico* techniques. Baringhaus and co-workers [44] generated a Catalyst pharmacophore model based on 17 chemically diverse ASBT inhibitors, which described molecular features essential for ASBT affinity; these data are in good agreement with a previous CoMFA model [45]. The model described by Baringhaus *et al.* [44] also enabled the *in silico* discovery of substrate(s) for ASBT. Together, the two models provided the first step in the rational design of prodrugs for specific targeting to ASBT. To make the study more complete and visualize the transporter directly, an *in silico* model for ASBT was generated using knowledge-based comparative modeling [3]. The TMDs were modeled using bacteriorhodopsin as a scaffold and the extracellular loops were modeled by remote-threading homology modeling [46]. A subsequent docking study with the natural substrate, cholic acid, identified putative substrate-binding sites that provided rational leads for site-directed mutagenesis studies. A more recent publication reinforced the theoretical model with additional biochemical data [4].

Ab initio *methods*

An alternative modeling approach that is rapidly gaining attention and could be applied to transporter proteins is the template-independent *ab initio* technique. This approach attempts to model a protein structure solely based on its primary sequence. Different algorithms and modeling techniques have been developed to accomplish this goal, with moderate success [47–51]. However, the applicability of these approaches is still limited by relative model inaccuracy and the complexity of calculation.

Future perspectives

Clearly, the future of transporter modeling will lie in the realm of combination *in silico* approaches. The grouping of QSAR and homology modeling techniques has been pioneered in the receptor pharmacology area and could now be applied to transporter families. Furthermore, the emergence of expression systems and HTS methods has enabled the collection of functional data on many transporter homologues and orthologues. The era of modeling individual transporters will now give way to comparative models of transporter families, which would illuminate the subtleties of genetic evolution in the fine-tuning of the substrate affinity of homologous transport proteins.

References

1 Grisshammer, R. and Tate, C.G. (1995) Overexpression of integral membrane proteins for structural studies. *Q. Rev. Biophys.* 28, 315–422

2 Rosenbusch, J.P. (2001) Stability of membrane proteins: relevance for the selection of appropriate methods for high-resolution structure determinations. *J. Struct. Biol.* 136, 144–157

3 Zhang, E.Y. *et al.* (2002) Modeling of active transport systems. *Adv. Drug Deliv. Rev.* 54, 329–354

4 Zhang, E.Y. *et al.* (2004) Topology scanning and putative three-dimensional structure of the extracellular binding domains of the apical sodium-dependent bile acid transporter (SLC10A2). *Biochemistry* 43, 11380–11392

5 Leach, A.R. (2001) *Molecular Modelling. Principles and Applications* (2nd edn), Prentice Hall

6 Sali, A. and Blundell, T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815

7 Laskowski, R.A. *et al.* (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291

8 Vriend, G. (1990) WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* 8, 52–56

9 Abramson, J. *et al.* (2003) Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* 301, 610–615

10 Huang, Y. *et al.* (2003) Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. *Science* 301, 616–620

11 Almqvist, J. *et al.* (2004) Homology modeling of the human microsomal glucose 6-phosphate transporter explains the mutations that cause the glycogen storage disease type Ib. *Biochemistry* 43, 9289–9297

12 Zuniga, F.A. *et al.* (2001) A three-dimensional model of the human facilitative glucose transporter Glut1. *J. Biol. Chem.* 276, 44970–44975

13 Karplus, M. and McCammon, J.A. (2002) Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* 9, 646–652

14 Mackerell, A.D., Jr (2004) Empirical force fields for biological macromolecules: overview and issues. *J. Comput. Chem.* 25, 1584–1604

15 Hruz, P.W. and Mueckler, M.M. (2000) Cysteine-scanning mutagenesis of transmembrane segment 11 of the GLUT1 facilitative glucose transporter. *Biochemistry* 39, 9367–9372

16 Mueckler, M. and Makepeace, C. (2004) Analysis of transmembrane segment 8 of the GLUT1 glucose transporter by cysteine-scanning mutagenesis and substituted cysteine accessibility. *J. Biol. Chem.* 279, 10494–10499

17 Alisio, A. and Mueckler, M. (2004) Relative proximity and orientation of helices 4 and 8 of the GLUT1 glucose transporter. *J. Biol. Chem.* 279, 26540–26545

18 Mueckler, M. *et al.* (2004) Transmembrane segment 3 of the Glut1 glucose transporter is an outer helix. *J. Biol. Chem.* 279, 46876–46881

19 Seigneuret, M. and Garnier-Suillerot, A. (2003) A structural model for the open conformation of the mdr1 P-glycoprotein based on the MsbA crystal structure. *J. Biol. Chem.* 278, 30115–30124

20 Stenham, D.R. *et al.* (2003) An atomic detail model for the human ATP-binding cassette transporter P-glycoprotein derived from

disulfide cross-linking and homology modeling. *FASEB J.* 17, 2287–2289
21 Chang, G. (2003) Structure of MsbA from *Vibrio cholera*: a multidrug resistance ABC transporter homolog in a closed conformation. *J. Mol. Biol.* 330, 419–430
22 Lee, J.Y. *et al.* (2002) Projection structure of P-glycoprotein by electron microscopy. Evidence for a closed conformation of the nucleotide binding domains. *J. Biol. Chem.* 277, 40125–40131
23 Campbell, J.D. *et al.* (2004) Molecular modeling correctly predicts the functional importance of Phe594 in transmembrane helix 11 of the multidrug resistance protein, MRP1 (ABCC1). *J. Biol. Chem.* 279, 463–468
24 Ecker, G.F. *et al.* (2004) A three-dimensional model for the substrate binding domain of the multidrug ATP-binding cassette transporter LmrA. *Mol. Pharmacol.* 66, 1169–1179
25 Campbell, J.D. *et al.* (2003) Extending the structure of an ABC transporter to atomic resolution: modeling and simulation studies of MsbA. *Biochemistry* 42, 3666–3673
26 Holm, L. and Sander, C. (1991) Database algorithm for generating protein backbone and side-chain co-ordinates from a C alpha trace application to model building and detection of co-ordinate errors. *J. Mol. Biol.* 218, 183–194
27 McGuffin, L.J. *et al.* (2000) The PSIPRED protein structure prediction server. *Bioinformatics* 16, 404–405
28 Rosenberg, M.F. *et al.* (2001) Repacking of the transmembrane domains of P-glycoprotein during the transport ATPase cycle. *EMBO J.* 20, 5615–5625
29 Jones, P.M. and George, A.M. (2002) Mechanism of ABC transporters: a molecular dynamics simulation of a well characterized nucleotide-binding subunit. *Proc. Natl. Acad. Sci. U. S. A.* 99, 12639–12644
30 Yates, C.R. *et al.* (2003) Structural determinants of P-glycoprotein-mediated transport of glucocorticoids. *Pharm. Res.* 20, 1794–1803
31 Ekins, S. *et al.* (2002) Application of three-dimensional quantitative structure–activity relationships of P-glycoprotein inhibitors and substrates. *Mol. Pharmacol.* 61, 974–981
32 Bednarczyk, D. *et al.* (2003) Influence of molecular structure on substrate binding to the human organic cation transporter, hOCT1. *Mol. Pharmacol.* 63, 489–498
33 Suhre, W.M. *et al.* Molecular Determinants of Substrate/Inhibitor Binding to the Human and Rabbit Renal Organic Cation Transporters, hOCT2 and rbOCT2. *Mol. Pharmacol.* (in press)
34 Chang, C. *et al.* (2004) Molecular requirements of the human nucleoside transporters hCNT1, hCNT2, and hENT1. *Mol. Pharmacol.* 65, 558–570
35 Cramer, R.D. *et al.* (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* 110, 5959–5967
36 Klebe, G. *et al.* (1994) Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J. Med. Chem.* 37, 4130–4146
37 Ekins, S. and Swaan, P.W. (2004) Computational models for enzymes, transporters, channels and receptors relevant to absorption, distribution, metabolism, excretion and toxicology. *Rev. Comp. Chem.* 20, 333–415
38 Kulkarni, S.S. *et al.* (2002) Three-dimensional quantitative structure–activity relationships of mazindol analogues at the dopamine transporter. *J. Med. Chem.* 45, 4119–4127
39 Kulkarni, S.S. *et al.* (2004) Structure–activity relationships at monoamine transporters for a series of *N*-substituted 3α-(bis[4-fluorophenyl]methoxy)tropanes: comparative molecular field analysis, synthesis, and pharmacological evaluation. *J. Med. Chem.* 47, 3388–3398
40 Ravna, A.W. *et al.* (2001) Molecular model of the *Escherichia coli* Na1/H1 antiporter NhaA. *Receptors Channels* 7, 319–328
41 Gray, J.H. *et al.* (2004) The concentrative nucleoside transporter family, SLC28. *Pflugers Arch.* 447, 728–734
42 Huang, M. *et al.* (2002) Inhibition of nucleoside transport by p38 MAPK inhibitors. *J. Biol. Chem.* 277, 28364–28367
43 Bohl, C.E. *et al.* (2004) A ligand-based approach to identify quantitative structure–activity relationships for the androgen receptor. *J. Med. Chem.* 47, 3765–3776
44 Baringhaus, K.H. *et al.* (1999) Substrate specificity of the ileal and the hepatic Na^+/bile acid cotransporters of the rabbit. II. A reliable 3D QSAR pharmacophore model for the ileal Na^+/bile acid cotransporter. *J. Lipid Res.* 40, 2158–2168
45 Swaan, P.W. *et al.* (1997) Molecular modeling of the intestinal bile acid carrier: a comparative molecular field analysis study. *J. Comput. Aided Mol. Des.* 11, 581–588
46 Jones, D.T. *et al.* (1999) Successful recognition of protein folds using threading methods biased by sequence similarity and predicted secondary structure. *Proteins* (Suppl. 3), 104–111
47 Jahnig, F. and Edholm, O. (1992) Modeling of the structure of bacteriorhodopsin. A molecular dynamics study. *J. Mol. Biol.* 226, 837–850
48 Milik, M. and Skolnick, J. (1993) Insertion of peptide chains into lipid membranes: an off-lattice Monte Carlo dynamics model. *Proteins* 15, 10–25
49 Taylor, W.R. *et al.* (1994) A method for alpha-helical integral membrane protein fold prediction. *Proteins* 18, 281–294
50 Pappu, R.V. *et al.* (1999) A potential smoothing algorithm accurately predicts transmembrane helix packing. *Nat. Struct. Biol.* 6, 50–55
51 Pellegrini-Calace, M. *et al.* (2003) Folding in lipid membranes (FILM): a novel method for the prediction of small membrane protein 3D structures. *Proteins* 50, 537–545
52 Xiang, Z. and Honig, B. (2001) Extending the accuracy limits of prediction for side-chain conformations. *J. Mol. Biol.* 311, 421–430
53 Salas-Burgos, A. *et al.* (2004) Predicting the three-dimensional structure of the human facilitative glucose transporter Glut1 by a novel evolutionary homology strategy: insights on the molecular mechanism of substrate migration, and binding sites for glucose and inhibitory molecules. *Biophys. J.* 87, 2990–2999
54 Gottschalk, K.E. (2004) Structure prediction of small transmembrane helix bundles. *J. Mol. Graph. Model.* 23, 99–110
55 Gottschalk, K.E. *et al.* (2004) A structural model of EmrE, a multi-drug transporter from *Escherichia coli*. *Biophys. J.* 86, 3335–3348
56 Ravna, A.W. *et al.* (2003) Molecular model of the neural dopamine transporter. *J. Comput. Aided Mol. Des.* 17, 367–382
57 Ravna, A.W. *et al.* (2003) Molecular mechanism of citalopram and cocaine interactions with neurotransmitter transporters. *J. Pharmacol. Exp. Ther.* 307, 34–41
58 Ekins, S. *et al.* (2002) Three-dimensional quantitative structure–activity relationships of inhibitors of P-glycoprotein. *Mol. Pharmacol.* 61, 964–973
59 Roman, D.L. *et al.* (2003) Interactions of antidepressants with the serotonin transporter: a contemporary molecular analysis. *Eur. J. Pharmacol.* 479, 53–63
60 Geldenhuys, W.J. *et al.* (2004) Molecular modeling studies on the active binding site of the blood–brain barrier choline transporter. *Bioorg. Med. Chem. Lett.* 14, 3085–3092
61 Gebauer, S. *et al.* (2003) Three-dimensional quantitative structure–activity relationship analyses of peptide substrates of the mammalian H+/peptide cotransporter PEPT1. *J. Med. Chem.* 46, 5725–5734