Mini Review

Structure and function of nucleotide sugar transporters: Current progress

Barbara Hadley a, Andrea Maggioni a, Angel Ashikov b, c, Christopher J. Day a, Thomas Haselhorst a, Joe Tiralongo a,⁎

a Institute for Glycomics, Griffith University, Gold Coast Campus, Queensland 4222, Australia
b Institut für Zelluläre Chemie, Zentrum Biochemie, Medizinische Hochschule Hannover, Carl-Neuberg Strasse 1, 30625 Hannover, Germany
c Laboratory of Genetic, Endocrine and Metabolic Diseases, Department of Neurology, Radboud University Medical Center, Geert Grooteplein Zuid 10 (route 830), Nijmegen, The Netherlands

ARTICLE INFO
Available online 11 June 2014

Keywords:
Nucleotide sugar transporters
CMP-sialic acid transporter
Golgi apparatus
Endoplasmic reticulum
STD NMR spectroscopy

ABSTRACT

The proteomes of eukaryotes, bacteria and archaea are highly diverse due, in part, to the complex post-translational modification of protein glycosylation. The diversity of glycosylation in eukaryotes is reliant on nucleotide sugar transporters to translocate specific nucleotide sugars that are synthesised in the cytosol and nucleus, into the endoplasmic reticulum and Golgi apparatus where glycosylation reactions occur. Thirty years of research utilising multidisciplinary approaches has contributed to our current understanding of NST function and structure. In this review, the structure and function, with reference to various disease states, of several NSTs including the UDP-galactose, UDP-N-acetylgalactosamine, UDP-N-acetylgalactosamine, GDP-fucose, UDP-N-acetylgalactosamine/UDP-glucose/GDP-mannose and CMP-sialic acid transporters will be described. Little is known regarding the exact structure of NSTs due to difficulties associated with crystallising membrane proteins. To date, no three-dimensional structure of any NST has been elucidated. What is known is based on computer predictions, mutagenesis experiments, epitope-tagging studies, in-vitro assays and phylogenetic analysis. In this regard the best-characterised NST to date is the CMP-sialic acid transporter (CST). Therefore in this review we will provide the current state-of-play with respect to the structure-function relationship of the (CST). In particular we have summarised work performed by a number groups detailing the affect of various mutations on CST transport activity, efficiency, and substrate specificity.

© 2014 Hadley et al. Published by Elsevier B.V. on behalf of the Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Contents
1. Introduction ............................................................... 23
2. General features of the nucleotide sugar transporter family ................................................................. 25
3. SLC35A2: UDP-galactose (UGT) and SLC35A3: UDP-N-acetylgalactosamine (NGT) transporters. ........................................ 26
4. SLC35C1: GDP-fucose transporter (GFT). ...................................................... 26
5. SLC35D2: UDP-N-acetylgalactosamine/UDP-glucose transporter (HGFC1) also known as GDP-mannose transporter in non-humans (GMIT). ........................ 27
6. SLC35A1 CMP-sialic acid transporter (CST) .................................................. 28
7. Summary and outlook. .................................................. 29
References ................................................................ 31

1. Introduction

Protein glycosylation, which in eukaryotes occurs predominantly in endoplasmic reticulum (ER) and the Golgi apparatus, is the most prevalent and complex post-translational modification. It was originally believed that only eukaryotic membrane-bound or secreted proteins were glycosylated, however it is now known that this process occurs in a range of eukaryotic nuclear and cytoplasmic proteins, as well as in bacteria and archaea. The process of glycosylation covalently attaches a glycan, a reaction catalysed by glycosyltransferases, to the growing end of a carbohydrate chain on a nascent protein or lipid. These complex modifications modulate the properties of the proteins they decorate. These modifications play a crucial role in every aspect of biology including increasing protein solubility [1]: increasing structural stability and

http://dx.doi.org/10.1016/j.csbj.2014.05.003
2001-0370/© 2014 Hadley et al. Published by Elsevier B.V. on behalf of the Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
<table>
<thead>
<tr>
<th>SLC nomenclature</th>
<th>NCBI preferred name</th>
<th>Substrate</th>
<th>Sub-cellular localisation</th>
<th>Link to disease</th>
<th>Ensembl ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLC35A1</td>
<td>CMP-Sia transporter (CST)</td>
<td>CMP-Sia</td>
<td>Exclusively Golgi</td>
<td>Congenital disorder of glycosylation (CDG2F) (OMIM #603585) [15,16]</td>
<td>ENSG00000164414</td>
</tr>
<tr>
<td>SLC35A2</td>
<td>UDP-Gal transporter (UGT)</td>
<td>UDP-Gal; UDP-GlcNAc</td>
<td>Golgi and/or ER</td>
<td>Colon cancer [17]; Wiskott-Aldrich syndrome [18]; CDG2M (OMIM #300896) [19,20]</td>
<td>ENSG00000102100</td>
</tr>
<tr>
<td>SLC35A3</td>
<td>UDP-GlcNAc transporter (NGT)</td>
<td>UDP-GlcNAc</td>
<td>Predominantly Golgi</td>
<td>Possible link to malaria through UDP-GlcNAc transporter homolog [21]; Musculoskeletal abnormalities in cattle [22]; Arthrogryposis, mental retardation, and seizures (OMIM #615533) [23]</td>
<td>ENSG00000117620</td>
</tr>
<tr>
<td>SLC35A4</td>
<td>Probable UDP-sugar transporter; MGC2541</td>
<td>Putative UDP-Gal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC35A5</td>
<td>Probable UDP-sugar transporter</td>
<td></td>
<td></td>
<td></td>
<td>ENSG00000138459</td>
</tr>
<tr>
<td>SLC35B1</td>
<td>UGTREL1</td>
<td>Putative sugar transporter</td>
<td></td>
<td></td>
<td>ENSG000000121073</td>
</tr>
<tr>
<td>SLC35B2</td>
<td>PAPS transporter 1</td>
<td>PAPS</td>
<td>Exclusively Golgi</td>
<td>Colorectal cancer [24]; Dysplasia [25]; Overexpression in hepatocarcinoma cell line [26]; Colon cancer [27]; Dysplasia [25]</td>
<td>ENSG00000124786</td>
</tr>
<tr>
<td>SLC35B3</td>
<td>PAPS transporter 2</td>
<td>PAPS</td>
<td>Exclusively Golgi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC35B4</td>
<td>UDP-Xyl transporter (YEA)</td>
<td>UDP-Xyl; UDP-GlcNAc</td>
<td>Golgi and/or ER</td>
<td>Regulation of obesity &amp; glucose homeostasis in mice [28]; PHACE syndrome [29]</td>
<td>ENSG00000203060</td>
</tr>
<tr>
<td>SLC35C1</td>
<td>GDP-Fuc transporter (GFT)</td>
<td>GDP-Fuc</td>
<td>Predominantly Golgi</td>
<td>Leukocyte adhesion deficiency (CDG2C) (OMIM #266265) [30,31]; hepatocellular carcinoma [32]; Ovarian cancer [34]</td>
<td>ENSG00000181830</td>
</tr>
<tr>
<td>SLC35C2</td>
<td>OVC0V1</td>
<td>Putative GDP-Fuc transporter. Promotes Notch1 fucosylation [33]</td>
<td></td>
<td></td>
<td>ENSG00000080189</td>
</tr>
<tr>
<td>SLC35D1</td>
<td>UDP-GlcA/UDP-GalNAc dual transporter</td>
<td>UDP-GlcA; UDP-GalNAc</td>
<td>Exclusively ER</td>
<td>Schneckenbecken dysplasia (OMIM #269250) [35]</td>
<td>ENSG00000116704</td>
</tr>
<tr>
<td>SLC35D2</td>
<td>UDP-GlcNAc/UDP-Glc/GDP-Man transporter (HFRC1)</td>
<td>UDP-GlcNAc; UDP-Glc; GDP-Man (not humans)</td>
<td>Exclusively Golgi</td>
<td></td>
<td>ENSG00000130958</td>
</tr>
<tr>
<td>SLC35D3</td>
<td>FRCL1</td>
<td>Substrate unknown</td>
<td></td>
<td>Chediak-Higashi syndrome [36]; Hermansky-Pudlak syndrome [36]</td>
<td>ENSG0000182747</td>
</tr>
</tbody>
</table>
protection from proteolysis [2]; assistance in protein folding [1]; participation in immune responses [3]; cell–cell and cell–extra cellular matrix (ECM) recognition [4]; and selective protein targeting in both intra- or extracellular destinations [5]. Glycosylation permits diversity of the proteome by encompassing a wide range of variables such as glycosidic linkages (anomeric configuration; carbon/carbon linkage between sugars, N- or O-linked), composition, structure and length. In eukaryotes, before any of these glycosylation reactions can occur, the activated sugar must be transported into the Golgi or ER lumen where it can be used as a substrate by glycosyltranferases, a task performed by a family of transport proteins called nucleotide sugar transporters (NSTs).

Cellular membranes, including those that enclose organelles, are biological barriers that selectively either allow, inhibit, restrict or dictate the rate of flow of a range of solutes such as charged organic or inorganic molecules. Transporter proteins are an effective solution to the movement of selected solutes across these hydrophobic barriers that would otherwise be excluded. The second largest family of membrane proteins is the solute carriers (SLC). The SLCs, which is a classification of human transporters, now include 52 families [6]. With such a range of SLC families, there are a wide variety of solutes that can be transported, from amino acids to sugars, to complex organic molecules. As such, SLCs also contain different transport strategies and mechanisms to achieve their function, such as operating as antiporters, symporters or simple carriers [7]. The solute carrier family SLC35 (HUGO Gene Nomenclature Committee) comprises members of the evolutionary conserved family of human NSTs. The solute carrier family SLC35 of human NSTs is divided into 7 subfamilies (SLC35A-G), identified on the basis of sequence similarity (SLC35E-G are orphan transporters, that is, their physiological functions are yet to be determined). Each NST subfamily is then divided further to differentiate the type of substrate(s) that is/are transported (Table 1).

2. General features of the nucleotide sugar transporter family

NSTs are highly conserved type III trans-membrane (TM) proteins that provide a link between the synthesis of nucleotide sugars (in the ER, nucleus or cytosol), and the glycosylation process that occurs in the Golgi or ER lumen. It is well-established that NSTs function as antiporters, exchanging cytosolic nucleotide sugar for the corresponding lumenal nucleotide monophosphate (Fig. 1) [8–12]. That is, a constant level of nucleotide sugar is maintained in the Golgi or ER lumen through the equimolar exchange of nucleotide sugar with nucleotide monophosphate. The NST antiporter mechanism has been investigated in NSTs reconstituted into proteoliposomes [12], yeast Golgi vesicles [13], and directly in Golgi fractions isolated from rat liver [10]. Studies using CST reconstituted into phosphatidylcholine proteoliposomes preloaded with CMP significantly stimulated the uptake of CMP-sialic acid in a phenomenon known as trans-stimulation [12]. However, the ability of the CST (and other NSTs) to translocate its corresponding nucleotide sugar in the presence and absence of the antiporter molecule (nucleotide monophosphate) has lead to the characterisation of this transport system as “leaky” [9,13,14].

Thirty years of NST research aimed at identification and biochemical characterisation has identified a number of features that are common to all currently known NSTs (reviewed in [14]), including:

- Translocation of the entire nucleotide sugar;
- Translocation is saturable; temperature, concentration and time dependent with apparent $K_m$ in the order of 1–10 μM; and is able to concentrate the nucleotide sugar within the lumen of the ER or Golgi;
- Translocation is insensitive to the presence of ATP and ionophores and are energised by the coupled translocation of the corresponding nucleoside monophosphate in the opposite direction (antiporter);
- Translocation is competitively inhibited by the corresponding nucleoside monomono- and diphosphate, but not by the free sugar;
- Some nucleotide sugars are translocated exclusively into the Golgi apparatus, some exclusively in the ER, while others are translocated in both, including some being splice variant dependent.

The initial identification of a range of NSTs was achieved through complementation analysis [31,39–42]. Subsequent characterisation of the majority of these NSTs was with respect to their ability to translocate a single nucleotide sugar [43,44], and it was commonly accepted that NSTs had absolute substrate specificities. More recently however, multi-substrate transporters of nucleotide sugars have been described in vitro (Table 1). In Caenorhabditis elegans for example, there are 18 putative NSTs, three of which have been well characterised. All three of these have been shown to have multi-substrate specificity including that encoded by the gene ZK896.9, which is capable of transporting UDP-glucose (UDP-Glc), UDP-galactose (UDP-Gal), UDP-N-acetylglucosamine (UDP-GlcNAc) and UDP-N-acetylgalactosamine (UDP-GalNAc) [45]. Multi-substrate specificity may be partially explained by a common evolutionary ancestor [38], or alternatively recent studies have proposed that NST redundancy may be an evolutionary backup mechanism in case of NST impairment, deletion or mutation [15].

Fig. 1. The general transport mechanism of NSTs. The XDP-sugar (nucleotide sugar donor) enters the lumen of the organelle in exchange for the corresponding nucleoside monophosphate (XMP). After entering the lumen the sugar is transferred to either a protein or lipid in a reaction catalysed by glycosyltransferases. The diphosphate nucleotide (XDP) is then acted upon by a membrane-bound nucleotide diphosphatase [37] producing the XMP that is subsequently exported [38]. In some cases where the nucleotide sugar donor is a monophosphate, the diphosphorylation reaction performed by the diphosphatase is not required.
Although the amino acid sequence of a number of NSTs from a range of species has been determined, this information however has not proven to be a good indicator of substrate specificity. For example, the mammalian CMP-sialic acid transporter (CST) and UDP-Gal transporter (UGT) are 43% identical, yet are only able to transport the corresponding nucleotide sugars (CMP-Neu5Ac and UDP-Gal, respectively) [46], whereas the UDP-GlcNAc transporter (NGT) from Kluyveromycetes lactis, which shares only 22% identity with the human NGT, has the same nucleotide sugar substrate [47]. Similarly, in vitro studies show that the transport of UDP-GlcNAc in humans is maintained by 3 different NSTs (SLC35A3, SLC35B4 and SLC35D2) from 3 subfamilies that share very low amino acid identity (see Table 1 and references therein).

Little is known regarding the exact structure of NSTs due to the difficulty associated with crystallising membrane proteins. To date, no three-dimensional structure of any NST has been elucidated. What is known is based on computer predictions, mutagenesis experiments, epitope-tagging studies and evolutionary analysis [14, 48, 49]. In general, NST membrane topology has been predicted to comprise between six to ten trans-membrane (TM) domains linked by hydrophilic loops on both sides of the Golgi membrane [43, 44, 50]. All NST topologies predicted to date suggest that the C- and N-termini are present on the cytosolic side [51–53], corresponding to an even number of TM domains. A distinct exception to this classical NST topology is the Aspergillus fumigatus UDP-galactofuranose transporter which has 11 predicted TM domains [54]. It has been shown that several Golgi apparatus NSTs such as those that transport UDP-GalNAC, GDP-Fuc, ATP and PAPS appear as homodimers [8], whereas the GDP-Man transporter (GFT) from Leishmania donovani is presumed to be a hexamer in solution [55]. As well as potential homo-oligomers being formed in the membrane, there are also reports describing interactions, or possible complexes being formed between NSTs and glycosyltransferases [56]. Sprong et al. also concluded that the ceramide galactosyltransferase guarantees an adequate supply of UDP-Gal in the ER lumen by retaining the UGT in a molecular complex [57].

Thus far the oligomeric state of a functional NST has not been conclusively determined. For readers interested in inverted membrane protein topologies and conformational dynamics of antiporters we recommend the following reviews [58, 59].

3. SLC35A2: UDP-galactose (UGT) and SLC35A3: UDP-N-acetylgalcosamine (NGT) transporters

The cDNA that encodes for the human UDP-Gal transporter (UGT) was first cloned and characterised by Muira et al., in 1996 [42], and was believed to have been the first mammalian NST cDNA sequence described. Detailed characterisation of the UGT was possible using the mutant cell lines MDCK-RCA’ [60, 61], CHO-Lec8 [62, 63] and Had-1 [51]. Complementation of these UGT defective cell lines restored transport activity, and expression of recombinant UGT in mammalian and yeast cells confirmed its localisation and specificity [41, 42, 64, 65]. Interestingly, two isoforms of gene encoding the UGT, UGT1 and UGT2 have been identified in humans [41, 42]. Analyses of these human splice variants show that the only difference is confined to the proteins extreme C-termini. UGT1 is localised only in the Golgi apparatus, whereas the UGT2 C-terminus contains a dilsyne motif that is responsible for dual localisation in the Golgi and ER [66]. A recent study concluded that although UGT2 is more abundant in nearly all mammalian tissues and cell lines tested, expression of both splice variants is important for glycosylation of proteins in mammalian cells [67].

Compared to several other well-characterised transporters, the UDP-GlcNAc transporter (NGT) shows limited amino acid sequence identity to other NGTs that transport the same substrate, in particular yeast and mammals [42]. It has been proposed that the transport mechanisms of the UGT and NGT may be coupled. The overexpression of NGT in MDCK-RCA’ (Madin-Darby canine kidney-rin resistant) and CHO-Lec8 mutant cells defective in UGT has been found to restore galactosylation of N-glycans [68]. These cells lack UDP-Gal transport in the Golgi apparatus and therefore are unable to add Gal to glycans [69]. Although NGT overexpression restored UDP-Gal transport, it also resulted in the decrease of transport of its natural substrate UDP-GlcNac into the Golgi. This data suggested that the biological function of both the NGT and UGT in galactosylation might be coupled. Recent investigations into substrate specificity of the UGT have shown that the UGT/CST can function as a chimeric transporter [70], this is addressed in more detail later in this review.

Using co-immunoprecipitation analysis and FLIM-FRET measurements on living cells, it was demonstrated that NGT and UGT form complexes when overexpressed in MDCK-RCA’ cells [71]. This suggested that NGT/UGT complexes either mediate transport of both substrates (UDP-Gal and UDP-GlcNac) or alternatively these complexes just bring the NGT and UGT homodimers together. Either way, the ability of NGT and UGT to interact with each other may be a regulation mechanism of N-glycan biosynthesis in the Golgi by ensuring adequate supply of both natural substrates to their respective glycosyltransferases. It was concluded that the NGT and UGT function in glycosylation is combined via their mutual interaction [71, 72]. However, it must be stressed that these studies are based on overexpression of the NGT and UGT, and therefore may not truly reflect the physiological situation. Interestingly, overexpression of certain receptors [73–75] has been shown to create a “Brefeldin effect” [76]. Brefeldin A (BFA) is a fungal metabolite that affects the molecular mechanisms regulating membrane traffic and organelle structure [75]. Treatment with BFA leads to a rapid accumulation of proteins in the ER and a collapse of the Golgi stacks [77]. The result is that the Golgi apparatus largely disappears leaving Golgi proteins to intermix with those in the ER. With overexpression of these particular proteins, the effects were phenotypically indistinguishable from those treated with the addition of BFA.

More recently a number of CDGs have been identified due to SLC35A2 (UGT) and SLC35A3 (NGT) mutation [19, 20, 23], specifically SLC35A2 has been implicated in early-onset epileptic encephalopathy and SLC35A3 in autism.

4. SLC35C1: GDP-fucose transporter (GFT)

The GDP-fucose transporter (GFT) regulates the fucosylation of glycans predominantly in the Golgi. It was first identified using complementation cloning during investigations into the Congenital Disease of Glycosylation-IIc (CDGIIc), now known as Leukocyte Adhesion Deficiency II (LADII). This disease is characterised by a lack of fucosylated glycoconjugates [30, 31] resulting in immunodeficiency and severe mental and growth retardation [78]. It was purported that a deficiency in GFT was responsible for this disease state [30, 31]. Interestingly, the GFT shows a substantial level of amino acid conservation with both the CST and UGT [30, 31] however, even now, the elements essential for activity and localisation of the GFT remain poorly understood [79]. Although overexpression of SLC35C2 (a putative GFT) also shows slight competition with the GFT in the O-fucosylation of Notch, GFT is essential for activity and localisation of the GFT remain poorly understood [79]. As with all NSTs, elucidating the structure–function relationship remains elusive due to the lack of a crystal structure. Studying the GFT has had an added challenge due to the lack of an appropriate mutant cell line. Recently, a novel Chinese hamster ovary (CHO) mutant (CHO-gmt5) was established that harboured double genetic defects in both the CST and GFT producing N-glycans deficient in both sialic acid (Sia) and fucose (Fuc) [79]. Studies using this mutant found that the C-terminal tail of the GFT was critical for its activity (Fuc-binding lectin recognition) but not localisation to the Golgi, in contrast to the murine CST [80] and several other transporters. This latest CHO-gmt5 study highlights several new structure/function relationships for this transporter [79].
5. SLC35D2: UDP-N-acetylglucosamine/UDP-glucose transporter (HFRC1) also known as GDP-mannose transporter in non-humans (GMT)

The GDP-mannose transporter (GMT) from *L. donovani* and *Saccharomyces cerevisiae* was originally identified and characterised in 1997 [81, 82]. In 2003 a novel human nucleotide sugar transporter gene, *hfrc1*, was cloned and characterised as being a multi-substrate specific NST homologous to *Drosophila melanogaster* fringe connection, *C. elegans* sqv-7 and human UGTrel7. In yeast, the heterologous expression of HFRC1 revealed the multi-substrate transport of UDP-GlcNAc, UDP-Glc and GDP-Man. Interestingly, and importantly, when expressed in mammalian cells, UDP-GlcNAc and UDP-Glc were transported but GDP-Man was not [83]. HFRC1 was subsequently identified and characterised as a member of SLC35D2 by the same group that had previously identified and characterised the murine and human SLC35D1 [84,85]. The transporter encoded by *hfrc1* is localised in the Golgi apparatus and exhibits approximately 50% identity with the human SLC35D1 [84]. It should be stressed that confirming GMT function in yeast is problematic as yeast possesses a high level of endogenous GDP-Man transport that can potentially interfere with the detection of heterologous expressed GMT activity [86].

The GMT is fundamentally essential for pathogens such as *A. fumigatus* whose cell wall is comprised predominantly of galactomannan, the main

---

**Fig. 2.** The direct analysis of Aspergillus GMT interaction with GDP-Man, GDP and GMP using STD NMR spectroscopy. $^1$H (a) and competition STD NMR spectra of Aspergillus Golgi-enriched fractions complexed with GDP-Man (b) followed by the addition of equimolar amounts of GMP (c) and GDP (d). Some STD signals were found to increase due to overlapping chemical shifts (e.g. the H1 ribose signal at 5.65 ppm), however the H8 guanine signal of the three ligands does not have the same chemical shift and therefore could be used to monitor the interaction of the GMT with GDP-Man, GDP and GMP. The H8 GuaGDP-Man signal (b) is reduced following the addition of GMP and GDP (c and d, respectively) with a corresponding appearance of H8 guanine signals associated with GMP and GDP. Specific mannose signals were reduced by $\sim$50% following the addition of equimolar GMP (c), and the signals after addition of GDP (d) showed a further reduction of $\sim$50% compared to (c).
defence against the host immune system [87]. The key role played by the GMT in the biosynthesis of the fungal galactomannan cell wall was recently highlighted by the absence of galactomannan synthesis following the targeted deletion of the GMT [88]. This suggests that the GMT may be an attractive target for drug discovery, particularly given the lack of a human GMT activity [83]. Similarly, the protozoa Leishmania is protected by a glycolycox composed mainly of Gal- and Man-glycoconjugates. This protective coat is a virulence factor that shields the parasite from hostile environments and supports its development and method of invasion. As with the cell wall of A. fumigatus, interruption of the corresponding essential transporter GMT in L. donovani had a severe impact on its pathogenicity [89].

We have recently utilised a Saturation Transfer Difference NMR spectroscopy (STD NMR) [90,91] approach to complement functional used assays to monitor the interaction of nucleotides and nucleotide sugars with the NSTs present in isolated Golgi-enriched fractions [92–94]. STD NMR is based on saturating the protein resonances with a cascade of selective pulses (on-resonance spectrum). The magnetization is rapidly transferred through the entire protein mediated by spin diffusion. If a ligand is in fast exchange with the protein-binding site then the saturation can be transferred to the binding ligand. Ligand protons that are in close contact with the protein will receive a higher degree of saturation than ligand protons that are more solvent exposed. A spectrum without any saturation (off-resonance spectrum) is simultaneously acquired and subtraction of the on-resonance spectrum and the off-resonance spectrum results in the final difference spectrum (STD) showing only signals from binding ligands. Additionally, protons in close proximity to the protein surface will show stronger STD NMR signals compared to ligand protons that are solvent exposed. Non-binding ligands will not show any STD NMR signals at all [95].

Utilising our STD NMR spectroscopy approach we directly investigated the binding of GDP-Man, GDP, GMP and Man to the Golgi-enriched fractions isolated from Aspergillus ( [93] and Fig. 2). We showed through STD NMR competition experiments that GDP binds tighter to the Aspergillus GMT than GMP and GDP-Man, with Man binding only weakly. Based on these experiments the relative importance/affinity of individual ligand moieties that bind the Aspergillus GMT were summarised as follows; GDP $\geq$ GMP $\geq$ GDP-Man $\geq$ Man. The natural antiporter substrates for the GMT are GDP-Man and GMP. However, the observation that GDP binds the GMT with higher affinity than the natural substrates (GMP-Man and GMP) can now be exploited in the design of novel GMT inhibitors [93]. (See Fig. 2.)

6. SLC35A1 CMP-sialic acid transporter (CST)

The CST is located exclusively in the Golgi apparatus where it co-localises with ST6Gal in the medial and trans Golgi, and translocates CMP-sialic acid (CMP-Sia) from the cytosol into the Golgi lumen in exchange for CMP in an antiporter mechanism (see Fig. 1) [14]. The cDNA of the murine CST was first isolated in 1996 by complementation cloning. Chinese hamster ovary mutants of the complementation group Lec2 (CHO 6B2) express a strong reduction of sialylated glycoconjugates due to a defect in the CST. By expression cloning, a cDNA encoding the mCST was identified that complemented the Lec2 phenotype [96]. It was shown to encode a highly hydrophobic, multiple membrane spanning protein of 36.4 kDa. Using the same cloning strategy the cDNA encoding the hamster CST was also isolated with the amino acid sequence showing a 95% identity with the mCST [97]. Related cDNAs from human, S. cerevisiae, and C. elegans were also identified by homology searches of gene databases. Expression of the murine CST in yeast was used to confirm the ability of the cloned CST to translocate CMP-Sia [13].

Due to the fact that yeast does not express Sia, it represented an ideal background free model to study the CST and to demonstrate that the cDNA identified by complementation cloning encoded an active transporter and not just an accessory protein required for CMP-Sia transport/translocation.

Subsequent to cloning and expression of the CST, independent groups began structure–function relationship investigations. Initially, the CST derived from five independent clones of the Lec2 complementation group were analysed to determine the molecular defects leading to the inactivation of the CST. One of these defects was observed to be a single missense mutation, Gly189Glu. It was shown that the mutant CST mRNA expression level was the same as that of the wild-type, and the mutant was also correctly targeted to the Golgi apparatus. This indicated that the Gly189Glu mutation was directly responsible for the inactivation of CST transport activity. Exchanging Gly189 to Ala

![Diagram representing the membrane topology of CST as proposed by independent studies.](image)
Table 2  
CST and CST/UGT chimeric mutations that altered transport and/or substrate recognition. CST and CST/UGT chimeric mutants shown to affect the transport and/or substrate recognition have been summarised. A complete list of all CST and CST/UGT mutants (based on the available literature) assessed including those that had no effect on transport and/or substrate recognition has been included in the supplementary data Table A.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Location</th>
<th>Background</th>
<th>Effect</th>
<th>Experiment</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>C16A</td>
<td>TMD 1</td>
<td>Site-directed mutagenesis of mouse CST. Generated and expressed in P. pastoris.</td>
<td>Essential for CST substrate specificity</td>
<td>STD NMR spectroscopy</td>
<td>Highly conserved in CSTs, but not present in Mn &amp; Hn UGT [94].</td>
</tr>
<tr>
<td>K65A</td>
<td>2nd loop</td>
<td>Site-directed mutagenesis of mouse CST. Generated and expressed in P. pastoris.</td>
<td>Significant effect on CMP-Neu5Ac recognition. Suggests binding the Sia moiety and not the nucleotide moiety.</td>
<td>STD NMR spectroscopy</td>
<td>Highly conserved in CSTs, but not present in Mn &amp; Hn UGT [94].</td>
</tr>
<tr>
<td>Q101H</td>
<td>TMD 3</td>
<td>Naturally occurring mutation in patient with intellectual disability and bleeding diathesis</td>
<td>50% reduction in CST activity</td>
<td>Functional assay</td>
<td>[16]</td>
</tr>
<tr>
<td>L112G &amp; D113G</td>
<td>3rd loop</td>
<td>CST-GFP12 mutant. Co-transfected with an EPO vector into CHO MAR-11 cells</td>
<td>Complete inactivation of CST</td>
<td>EPO/IEF assay</td>
<td>[101]</td>
</tr>
<tr>
<td>L136G</td>
<td>4th loop</td>
<td>MAR-11 cells</td>
<td>Essential for CST activity</td>
<td>EPO/IEF assay</td>
<td>Highly conserved in CST. UGT conserved counterpart, L160G did not affect UGT [101].</td>
</tr>
<tr>
<td>G153A &amp; G154A</td>
<td>TMD 5</td>
<td>Cotransfection of constructs that express CST into CHO CST-deficient MAR-11 cells.</td>
<td>Slightly reduced activity</td>
<td>EPO/IEF assay</td>
<td>[99]</td>
</tr>
<tr>
<td>G153I &amp; G154I</td>
<td>Double mutant</td>
<td>Cotransfection of constructs that express CST into CHO CST-deficient MAR-11 cells.</td>
<td>Clear reduction in activity</td>
<td>EPO/IEF assay</td>
<td>[99]</td>
</tr>
<tr>
<td>G177A &amp; G179A</td>
<td>TMD 6</td>
<td>Cotransfection of constructs that express CST into CHO CST-deficient MAR-11 cells.</td>
<td>Slightly reduced activity</td>
<td>EPO/IEF assay</td>
<td>[99]</td>
</tr>
<tr>
<td>G177I &amp; G179I</td>
<td>Double mutant</td>
<td>Cotransfection of constructs that express CST into CHO CST-deficient MAR-11 cells.</td>
<td>Clear reduction in activity</td>
<td>EPO/IEF assay</td>
<td>[99]</td>
</tr>
<tr>
<td>G189A &amp; G192A</td>
<td>TMD 6</td>
<td>Cotransfection of constructs that express CST into CHO CST-deficient MAR-11 cells.</td>
<td>Clear reduction in activity</td>
<td>EPO/IEF assay</td>
<td>[99]</td>
</tr>
<tr>
<td>G198E</td>
<td>TMD 6</td>
<td>Site-directed mutagenesis. Expressed in CHO-WT and mutant Lec 2</td>
<td>Clear reduction in activity</td>
<td>Complementation analysis</td>
<td>Gly189 in highly conserved CST region. Indicates size of aa is critical for activity — not charge [98]. As above [98].</td>
</tr>
<tr>
<td>G189Q</td>
<td>TMD 6</td>
<td>Site-directed mutagenesis. Expressed in CHO-WT and mutant Lec 2</td>
<td>Clear reduction in activity</td>
<td>Complementation analysis</td>
<td>As above [98].</td>
</tr>
<tr>
<td>G189S</td>
<td>TMD 6</td>
<td>Site-directed mutagenesis. Expressed in CHO-WT and mutant Lec 2</td>
<td>Clear reduction in activity</td>
<td>Complementation analysis</td>
<td>As above [98].</td>
</tr>
<tr>
<td>Y214A</td>
<td>TMD 7</td>
<td>Site-directed mutagenesis of Mouse CST. Generated and expressed in P. pastoris.</td>
<td>Dramatic effect on CMP-Neu5Ac recognition. Suggests binding of the Sia moiety and not the nucleotide moiety.</td>
<td>STD NMR spectroscopy</td>
<td>Highly conserved in CSTs, but not present in Mn &amp; hUGT [98]</td>
</tr>
<tr>
<td>Y214G</td>
<td>TMD 7</td>
<td>hUGT:hCST chimera</td>
<td>Loss of CST activity (retained UGT activity)</td>
<td>Complementation analysis</td>
<td>[70]</td>
</tr>
<tr>
<td>S216F</td>
<td>TMD 7</td>
<td>hUGT:hCST chimera</td>
<td>Loss of CST activity (retained UGT activity)</td>
<td>Complementation analysis</td>
<td>[70]</td>
</tr>
<tr>
<td>23GTPGG6P to 38GGCC7m</td>
<td>8th loop</td>
<td>CST-GFP12 mutant. Co-transfected with an EPO vector into CHO MAR-11 cells</td>
<td>Complete inactivation of CST</td>
<td>EPO/IEF assay</td>
<td>[101]</td>
</tr>
<tr>
<td>G256A &amp; G257A</td>
<td>TMD 8</td>
<td>Cotransfection of constructs that express CST into CHO MAR-11 cells.</td>
<td>Clear reduction in activity</td>
<td>EPO/IEF assay</td>
<td>[99]</td>
</tr>
<tr>
<td>G256G &amp; G257I</td>
<td>Double mutants</td>
<td>Cotransfection of constructs that express CST into CHO-deficient MAR-11 cells.</td>
<td>Clear reduction in activity</td>
<td>EPO/IEF assay</td>
<td>[99]</td>
</tr>
<tr>
<td>G256A &amp; G257A</td>
<td>TMD 8</td>
<td>Cotransfection of constructs that express CST into CHO MAR-11 cells.</td>
<td>Complete inactivation of CST</td>
<td>EPO/IEF assay</td>
<td>[101]</td>
</tr>
<tr>
<td>I270G</td>
<td>8th loop</td>
<td>CST-GFP12 mutant. Co-transfected with an EPO vector into CHO MAR-11 cells</td>
<td>Partially reduced transport activity</td>
<td>EPO/IEF assay</td>
<td>[101]</td>
</tr>
<tr>
<td>I270W</td>
<td>8th loop</td>
<td>CST-GFP12 mutant. Co-transfected with an EPO vector into CHO MAR-11 cells</td>
<td>Partially reduced transport activity</td>
<td>EPO/IEF assay</td>
<td>[101]</td>
</tr>
<tr>
<td>K972A</td>
<td>8th loop</td>
<td>CST-GFP12 mutant. Co-transfected with an EPO vector into CHO MAR-11 cells</td>
<td>Complete inactivation of CST</td>
<td>EPO/IEF assay</td>
<td>[101]</td>
</tr>
<tr>
<td>K972G</td>
<td>8th loop</td>
<td>CST-GFP12 mutant. Co-transfected with an EPO vector into CHO MAR-11 cells</td>
<td>Complete inactivation of CST. Essential for CST transport activity</td>
<td>EPO/IEF assay</td>
<td>[101]</td>
</tr>
<tr>
<td>K972H</td>
<td>8th loop</td>
<td>CST-GFP12 mutant. Co-transfected with an EPO vector into CHO MAR-11 cells</td>
<td>Complete inactivation of CST. Essential for CST transport activity</td>
<td>EPO/IEF assay</td>
<td>Changed to His as His is in hGDP-fucose transporter. [101]</td>
</tr>
<tr>
<td>Deletion of last 4 amino acids IICV</td>
<td>C-terminal tail</td>
<td>hUGT-hCST chimera</td>
<td>Eliminated the export signals and prevented ER-to-Golgi translocation</td>
<td>Affects the efficiency of CST</td>
<td>[80]</td>
</tr>
<tr>
<td>TMD 2 &amp; 3</td>
<td>TMD 2</td>
<td>hUGT-hCST chimera</td>
<td>Required for substrate specificity</td>
<td>Complete inactivation of CST</td>
<td>[100]</td>
</tr>
<tr>
<td>TMD 7</td>
<td>TMD 3</td>
<td>hUGT-hCST chimera</td>
<td>Required for substrate specificity</td>
<td>Complete inactivation of CST</td>
<td>[100]</td>
</tr>
<tr>
<td>CST-GFP4 (GFP loop interruption)</td>
<td>3rd loop</td>
<td>CST-GFP12 mutant. Co-transfected with an EPO vector into CHO MAR-11 cells</td>
<td>Complete inactivation of CST</td>
<td>EPO/IEF assay</td>
<td>[52,101]</td>
</tr>
</tbody>
</table>

(continued on next page)
did not affect transport, though Gly189Gln and Gly189Ile mutants resembled the inactivates Gly189Glu mutant (Table 2 and Supplementary data Table A). This suggested that the insertion of a large amino acid at position 189 rather than the charge associated with Glu rendered the CST inactive. The exchange Gly189 to Glu occurs in a region that is highly conserved in both the mammalian CST and UGT, as well as in Schizosaccharomyces pombe and C. elegans, suggesting this region is essential for a functional transporter [98].

The initial topology model established by Eckhardt et al. (1999) predicted 10 TM domains with both the N- and C-terminals facing the cytosolic side of the Golgi membrane (Fig. 3). This model was deduced by epitope-tagging studies, site-directed mutagenesis and hydrophobicity plots [52], and established that Gly189 is one of 10 Gly residues spread across four TM domains (TM 5–8) that were presumed to form a putative hydrophobic channel [98]. The creation of eight double mutants exchanging each Gly pair with Ala and Ile confirmed the importance of these Gly residues [99] (Fig. 3 and Table 2). In order to assess these and other mutants Lim et al. (2008) established the EPO/IEF assay to assess CST activity. Briefly, recombinant human erythropoietin (EPO) is a heavily glycosylated molecule and a simple analysis of the sialylation pattern using isoelectric focusing (IEF) can identify any changes in these patterns. MAR-11 CHO cells lack a functional CST and these cells were used as the host to analyse the relative activities of different mutant transporters. Using this assay system Lim et al. (2008) concluded that there was a direct correlation between the increased steric hindrance associated with the exchange of Gly with either Ala or Ile and the reduction of CST substrate translocation [99].

The suggestion of a Gly-rich hydrophilic channel or pore through which CMP-Sia can pass is contradictory to the hypothesis that the CST is a simple solute carrier [12] and to the concept of an antiporter mechanism [44]. This was based on experimental evidence that the transporter has the ability to alternatively expose its CMP/CMP-Sia binding site from either the cytosolic or luminal side of the Golgi membrane [12]. However, the two hypotheses can be reconciled in that the binding of CMP or CMP-Sia to the CST may permit a conformational change that allows the formation of a Gly-rich hydrophilic channel, enabling the selective translocation of the bound molecule. This unfortunately can only be conclusively demonstrated through the elucidation of the CST crystal structure.

Much of what we know regarding the CST (and UGT) structure–function relationship comes from a series of elegant studies evaluating the function of an array of UGT/CST chimeric transporters in both LeC2 and LeC8 CHO complementation groups [46,70,100]. Initial studies showed that substitution of CST helix 7 into the UGT chimera was enough to elicit CST activity, with the addition of helices 2 and/or 3 greatly enhancing the efficiency, suggesting that this chimeric CST/UGT now had the ability to recognise and transport both UDP-Gal and CMP-Sia [46,100] (Table 2). More recently further analysis of UGT/CST chimeras has defined a sub-molecular region that is necessary for CMP-Sia recognition [70]. Analysis of chimeras indicated that the Val208–Gly217 stretch in the CST (located in the helix 7) was essential for CST activity. Two of the amino acids located in this stretch, Tyr214 and Ser216 were subsequently identified by site-directed mutagenesis (both single and double mutants) to be important for CMP-Sia recognition, with Tyr214 found to be critical for substrate recognition (Table 2). The authors postulated that hydrogen bond formation involving the hydroxyl side-chains of these two amino acids may make specific interaction with the Sia moiety of the substrate [70].

Utilising STD NMR spectroscopy we were able to confirm the importance of Tyr214 in CMP-Sia recognition, specifically that it is intimately involved in the recognition and binding to the Sia moiety of CMP-Sia [94]. The generation of a Tyr214Ala CST mutant leads to the complete loss of STD signals associated with Sia, even though significant binding was observed to the CMP moiety. In addition to Tyr214, we were also able to identify another CST mutant, Lys65Ala that leads to a significant reduction in CMP-Sia recognition [94]. The latter residue in particular was identified using a bio-informatic approach where a sequence alignment of the mouse and human UGT and seven evolutionary diverse CSTs (H. sapiens, Mus musculus, Gallus gallus, Xenopus laevis, Takifugu rubripes, and Danio rerio) was performed. As shown in Supplementary Fig. A, a number of structural elements including the 10 Gly residues stretching TM 5–8 that have been implicated in the formation of a transporter channel [99], and Glu101, Leu136, and Lys272 and residues in the human CST loop regions shown to be essential for transport activity [101] appear not to confer substrate specificity as they are conserved in both the CST and UGT. However, Ser216 and Tyr214, independently identified by Maggioni et al. [94] and Takeshima-Futagami et al. [70] as being important for CMP-Sia recognition are highly conserved in evolutionary distant CST, but are not found in either the mouse or human UGT. Supplementary Fig. A also reveals the presence of seven Cys residues absolutely conserved in all CSTs that are not present in the UGT. Three Cys residues are present within the UGT, but their locations differ to those present in the CST sequence. To further explore the role of these Cys residues in disulphide bond formation, a web-based disulphide bond prediction algorithm (DIANN) [102] was used to analyse the CST and UGT sequence. Of the seven Cys residues in the CST six were predicted to form disulphide bonds, giving both intra- and inter-TMD connections. The Cys putatively involved in disulphide bond formation are Cys16–Cys49 (TM1–TM2D), Cys127–Cys131 (TM4–TM4D) and Cys152–Cys307 (TM5–TM9) (highlighted in Fig. A2). Of the three Cys residues in the UGT none were predicted to be involved in disulphide bond formation [94]. We therefore explored how the disruption of two of these putative disulphide bonds in the CST (C16A and C152A, Supplementary Fig. A) affected CST substrate recognition. However only Cys16Ala had any effect on CMP-Sia binding as assessed by STD NMR spectroscopy, this is despite the fact that alkylating and reducing agents completely abolished CST–CMP-Sia interaction [94]. This is interesting in so far as sialyltransferases contain two to four invariant Cys residues that are all involved in CMP-Neu5Ac substrate binding. These Cys residues are indispensable to the structural and functional integrity of sialyltransferases, with complete loss of catalytic activity observed following the mutation of either invariant Cys to Ala or Ser [103]. However, unlike sialyltransferases where the mutation of single Cys residues abolishes activity, it would appear that in the CST disruption of multiple Cys is required to achieve a similar outcome.

In addition to the importance of specific TM domains in CST and UGT activities, the function of the CST hydrophilic loops through green fluorescent protein (GFP) insertion experiments has also been assessed [101]. Three distinct loops that congregate around TM3D and TM7D, as well as several highly conserved amino acids were found to be crucial for the transport activity of both the CST and UGT (Table 2 and highlighted in orange in Fig. 3).

### Table 2 (continued)

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Location</th>
<th>Background</th>
<th>Effect</th>
<th>Experiment</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>CST-GFP8 (GFP loop interruption)</td>
<td>7th loop</td>
<td>CST-GFP12 mutant, Co-transfected with an EPO vector into CHO MAR-11 cells</td>
<td>Complete inactivation of CST</td>
<td>EPO/IEF assay a</td>
<td>[52,101]</td>
</tr>
<tr>
<td>CST-GFP10 (GFP loop interruption)</td>
<td>8th loop</td>
<td>CST-GFP12 mutant, Co-transfected with an EPO vector into CHO MAR-11 cells</td>
<td>Partial inactivation of CST</td>
<td>EPO/IEF assay a</td>
<td>[52,101]</td>
</tr>
</tbody>
</table>

a EPO/IEF assay: Recombinant human erythropoietin/isolectric focusing assay.
7. Summary and outlook

Significant progress has been made over the past decade towards not only the elucidation of NST structure-function relationship, but also better understanding the role of NSTs in various disease states including CDGs and microbial pathogenesis (e.g. A. fumigatus and Leishmania). These data have been generated using a multidisciplinary approach employing techniques ranging from site directed mutagenesis and complementation analyses to STD NMR spectroscopy and transport assays. The studies covered in this review have provided a fundamental understanding of several important NSTs. The continued use of multidisciplinary approaches towards understanding NST structure and function will provide further important advances in the field. However, only with the elucidation of the 3-dimensional structure of an NST will a full understanding of the structure-function relationship of this important class of transporter be realised.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.csbj.2014.05.003.

References


