Identification of essential amino acid residues in the hydrophilic loop regions of the CMP-sialic acid transporter and UDP-galactose transporter

Kah Fai Chan, Peiqing Zhang, and Zhiwei Song

Bioprocessing Technology Institute, Agency for Science, Technology and Research (A*STAR), 20 Biopolis Way, #06-01 Centros, Singapore 138668, Singapore

Received on December 11, 2009; revised on February 1, 2010; accepted on February 2, 2010

The Golgi CMP-sialic acid transporter (CST) is a type III transmembrane protein with 10 transmembrane domains that are linked by eight hydrophilic loops. To investigate the function of these hydrophilic loops, the green fluorescent protein (GFP) was inserted into each loop of the transporter. Expression and localization of the resulting CST-GFP fusion proteins were confirmed by analyzing the fluorescence of GFP. The transport activity of the CST-GFP proteins was analyzed by a previously described erythropoietin/isolectric focusing assay in CST-deficient MAR-11 cells. Interruption of the second and fourth luminal loops and the fourth cytosolic loop of CST with GFP resulted in complete or partial loss of transport activity. Regions in these loops that play crucial roles in CST’s activity were identified by Gly substitutions. Single amino acid substitution experiments revealed that Lys272 of the fourth loop on the cytosolic side of CST is essential for transport activity. Mutation of the conserved Lys residue (Lys272) in the UDP-galactose transporter (UGT) also resulted in a complete loss of its activity. Point mutations of highly conserved amino acid residues in the loop regions identified Leu136 of CST as essential for its activity. However, mutation of the conserved Leu residue in UGT (Leu160) did not affect the transport activity of UGT. Finally, mutation of Leu224 in UGT completely inactivated the activity of UGT, although mutation of its conserved counterpart in CST, Leu199, did not have any effect on CST. This study provides a structure–function analysis of the loop regions in CST and UGT.

Keywords: CMP-sialic acid transporter/hydrophilic loops/isolectric focusing (IEF)/recombinant erythropoietin (EPO)/structure–function relationship

Introduction

The process of protein N-glycosylation can be divided into two major steps, the assembly and transfer of pre-assembled core oligosaccharides from precursor dolichol pyrophosphate to Asn residues of nascent proteins in the endoplasmic reticulum (ER) followed by the subsequent processing of N-glycans in the Golgi apparatus. Within the Golgi lumen, the N-glycan undergoes further modifications by a series of glycosyltransferases to yield mature complex type N-glycans (Kornfeld and Kornfeld 1985). The sugar substrates added onto the growing glycans are in the form of nucleotide sugars which are synthesized in the cytoplasm or nucleus. These substrates are then transported into the Golgi apparatus by a set of nucleotide-sugar transporters (NSTs).

NSTs are highly conserved type III transmembrane proteins. They usually contain six to 10 transmembrane domains linked by hydrophilic loops on both sides of the Golgi membrane (for reviews, see Hirschberg et al. 1998; Gerardy-Schahn et al. 2001; Handford et al. 2006). CMP-sialic acid transporter (CST) represents one of the best-studied NSTs and is localized in the medial-trans Golgi (Zhao et al. 2006). It contains 10 transmembrane domains with both the amino and carboxyl termini facing the cytosolic side of the Golgi membrane (Eckhardt et al. 1999). The CST functions as an antiporter that couples the import of CMP-sialic acid into the Golgi lumen and the export of CMP back to the cytosol (Beijon et al. 1997; Hirschberg et al. 1998). Early structure–function studies on CST involved the expression of murine CST in yeast followed by the analysis of the transport activity using Golgi-enriched vesicles prepared from the yeast (Berninsone et al. 1997).

The roles of the N- and C-terminal cytosolic tails in CST localization and trafficking have been analyzed. The C-terminal tail of CST was found to be essential for its ER export and Golgi localization. Deletion of the last four amino acids (IIGV) of the C-terminal tail eliminated the export signals and prevented ER-to-Golgi translocation (Zhao et al. 2006). The importance of highly conserved Gly residues within the transmembrane domains was revealed through the mutation of a Gly residue to different amino acids (Eckhardt et al. 1998). Introduction of a G281D point mutation identified in UDP-galactose transporter (UGT)-deficient Lec8 CHO mutant into the homologous region of the CST also led to the inactivation of the transporter (Oelmann et al. 2001). The significance of the Gly residues in the transmembrane domains of the CST was further confirmed using an erythropoietin/isolectric focusing (EPO/IEF) assay system (Lim et al. 2008). Submolecular regions re-

© The Author 2010. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org
Fig. 1. Insertion of GFP at three loops in CST affects its transport activity. (A) Thirteen CST-GFP fusion constructs were generated by overlap PCR. In each CST-GFP fusion construct, one GFP open reading frame with two linker peptides linked to its N- and C-termini was inserted in a different position of CST. The insert in CST-GFP1 is shown in detail as an example for how other CST-GFP fusion constructs were generated. Detailed insertion sites for all constructs are shown in Supplementary Table 1. (B) Wild-type CST (CST-WT) and different CST-GFP fusion constructs were co-transfected with an EPO expressing vector into MAR-11 cells. Two days after transfection, recombinant EPO produced in the conditioned media were analyzed by the EPO/IEF assay. CST-GFP4 and CST-GFP8 did not show any transport activity, while CST-GFP10 showed partial loss of activity. (C) Twenty-four hours after transfection, the Golgi complex was visualized using an anti-giantin antibody (red). The green fluorescence of GFP in CST-GFP fusions was observed under UV light. Top panel, HA-CST was used as a control, and anti-HA antibody staining is shown in green. Except for CST-GFP10 which did not show any green fluorescence of GFP, all other CST-GFP fusion proteins are colocalized with the Golgi marker, giantin.
sponsible for the determination of substrate specificity were identified using a panel of CST-UGT chimera (Aoki et al. 2001). Certain chimeric transporters with sequences obtained from both CST and UGT were able to transport both UDP-galactose and CMP-sialic acid. It was further demonstrated that transmembrane domain 7 of the CST is required for substrate specificity, whereas transmembrane domains 2 and 3 affect the efficiency of nucleotide-sugar transport (Aoki et al. 2003).

Recently, we have developed a simple EPO/IEF assay for more sensitive structure-function analyses of CST. Recombinant EPO expressed in CST-deficient CHO MAR-11 cells completely lack sialic acid due to dysfunctional CST in MAR-11 cells. Co-expressing normal CST in MAR-11 cells restores the EPO sialylation pattern that is similar to the EPO produced by wild-type CHO cells. The mutant version of CST with defective transport activity was unable or only partially able to sialylate recombinant EPO as shown on the IEF gel (Lim et al. 2008). This assay seems to be more sensitive than the lectin-based cell surface staining assay used in the assessment of CST and UGT activities in Lec2 and Lec8 cells (Aoki et al. 2001).

In this study, we investigated the function of the hydrophilic loops of CST. Green fluorescent protein (GFP) was inserted into different loops of CST to study the function of each loop in CST. With this “GFP-insertion scan” we identified regions on three distinct loops of CST, the second and fourth luminal and fourth cytosolic loops, which are essential for transport activity. A detailed mutagenesis analysis identified several highly conserved amino acids that are crucial for the transport activity of CST and UGT.

Results

In this report, an indirect measurement of the transport activity was employed. The activity of each mutant CST is assessed by measuring the incorporation of sialic acid into the EPO molecule produced in CST-null MAR-11 cells. With the EPO/IEF assay, we have clearly demonstrated a positive correlation between EPO sialylation and the amount of plasmid DNA that was produced in CST-null MAR-11 cells. With the EPO/IEF measuring the incorporation of sialic acid into the EPO molecule, we employed. The activity of each mutant CST is assessed by Western blot analysis (Lim et al. 2008). A construct expressing EPO was co-transfected with each of the 13 CST-GFP fusion constructs into MAR-11 cells. Conditioned media from each transfection were harvested 48 h after transfection, and the sialylation patterns of recombinant EPO were analyzed by IEF. The results are shown in Figure 1B. The lane on the farthest left shows EPO produced in MAR-11 cells. Due to the lack of sialylation, the EPO was not resolved by IEF. The several bands observed in this lane could be the result of aggregation between unsialylated EPO. The lane under “CST-WT” shows the EPO produced by MAR-11 cells that were co-transfected with a construct encoding normal CST. The rest of the lanes show the sialylation patterns of EPO produced in MAR-11 cells that were co-transfected with respective CST-GFP fusion constructs as indicated. EPO molecules produced in MAR-11 cells co-transfected with CST-GFP4 and CST-GFP8 completely lack sialylation, suggesting that the GFP insertion in these two constructs has completely inactivated the CST (Figure 1B). One GFP insertion, CST-GFP10, has partially lost transport activity of CST. All other insertions did not affect the activity of CST. Given that these experiments were all carried out in the same MAR-11 mutant cells, all other factors that can affect the sialylation of recombinant EPO, such as galactosyltransferases and sialyltransferases, remain unchanged in all experiments. The only difference in each transfection lies in the CST-GFP constructs. Hence, the three loops that were interrupted by GFP in CST-GFP4, CST-GFP8 and CST-GFP10 play important roles in transporting CMP-sialic acid.

GFP-CST fusion proteins are localized to the Golgi apparatus

Cells that were transfected with CST-GFP fusion constructs were analyzed under a fluorescent microscope. Cells transfected with all CST-GFP constructs, except CST-GFP10, clearly expressed the typical green fluorescence of GFP. The expressed CST-GFP fusion proteins are all localized to the Golgi as the green fluorescence overlaps with the Golgi marker, giantin (see Supplementary Figure 1 online). For a closer look at the cellular localization of the CST-GFPs, several images are enlarged and shown in Figure 1C. As a control, cells transfected with a HA-tagged CST (HA-CST) construct were stained with an anti-HA antibody. HA-CST is clearly localized to the Golgi as it is colorized with a Golgi marker, giantin (top panel, Figure 1C). CST-GFP2, CST-GFP4 and CST-GFP8 are expressed and clearly localized to the Golgi (Figure 1C), although the latter two failed to restore the sialylation of EPO in MAR-11 cells (Figure 1B). Among 13 CST-GFPs, only CST-GFP10 failed to show any fluorescence, and the reason is unknown. Overall, these data suggest that insertion of GFP into these loop regions of CST does not affect its Golgi localization. In addition, the intensity of the green fluo-
Fig. 2. GS in the three loops of CST-GFP12 inactivates CST. (A) Tetra-Gly substitution-1 (TGS-1): the tetrapeptide (111NLDA114) in CST-GFP12 were mutated to four Gly residues (GGGG); TGS-2: the tetrapeptide (232EIKE235) in CST-GFP12 were mutated to GGGG; TGS-3: 236KGFF239 were changed to GGGG; TGS-4: 267TDNF270 were mutated to GGGG; TGS-5: 269NIMK272 were mutated to GGGG; TGS-6: 271MKGF274 in CST-GFP12 were mutated to GGGG. (B) Twenty-four hours after transfection, the Golgi complex was visualized by anti-giantin staining (red). The green fluorescence of GFP in all CST-GFP12-TGS mutants was observed under UV light. Results show that all CST-GFP12-TGS mutant proteins are localized in the Golgi. (C) CST-GFP12 and different CST-GFP12-TGS mutants were co-transfected with an EPO expressing vector into MAR-11 cells. The recombinant EPO produced in each transfection were analyzed by the EPO/IEF assay. Except TGS-2, all other TGS mutations of CST-GFP12 demonstrate loss of transport activity. (D) CST-GFP12 and three different di-Gly substitution (DGS) mutants were co-expressed with EPO in MAR-11 cells. EPO/IEF assay reveals that mutating 112LD113 to two Gly residues (DGS-1) inactivated the CST. However, changing amino acids 238FF239 in GFP-CST12 to two Gly residues (DGS-3) or changing 260DN269 to two Gly residues in GFP-CST12 (DGS-4) did not affect the activity of CST-GFP12.
rescene seems to be similar whether the GFP is inserted into the CST on the cytosolic side or on the Golgi lumen side. In two constructs, GFP was linked to the C-terminus of CST. In CST-GFP11, the N-terminus of GFP was directly fused to the C-terminus of CST. In CST-GFP12, a peptide linker, SGGGGS, was placed in between. The results suggest that linking GFP to the C-terminus of CST with (CST-GFP12) or without (CST-GFP11) a linker does not affect its Golgi localization (Figure 1C). CST-GFP12 was chosen to generate all the mutations in following experiments. As such, the expression and localization of CST-GFP can be easily observed.

Mutating the amino acids in the three loops in CST-GFP12 resulted in losing transport activity

To confirm the results observed in the GFP insertion experiments, amino acids in the three loops, the second and fourth lumenal loops and the fourth cytosolic loop, were mutated in CST-GFP12 construct. Six tetra-Gly substitution (TGS) mutants were generated. In each of these constructs, four consecutive amino acids in a loop were substituted with four Gly residues. For example, in tetra-Gly substitution-1 (TGS-1), four amino acids in the second lumenal loop in CST-GFP12, 111NLDA114 (mutated in TGS-1) were substituted with four Gly residues. These constructs are TGS-2 (232EIKE235), TGS-3 (233KGFF239), TGS-4 (267TDN1270), TGS-5 (269NIM272) and TGS-6 (271MKGF274). The amino acid residues in the parentheses indicate the substituted amino acids in each construct (Supplemental Table 1). To ensure that these six CST-GFP12 mutant proteins (TGS-1 to TGS-6) were stably expressed and localized to the Golgi apparatus, they were transiently expressed in MAR-11 cells. The Golgi apparatus were visualized with anti-giantin. As shown in Figure 2A, all six TGS mutants were expressed and colocalized with the Golgi marker giantin. The transport activity of these TGS mutants was analyzed with the EPO/IEF assay, and the results are shown in Figure 2C. The lane on the farthest left shows the unsialylated EPO produced by MAR-11 cells. The second lane shows EPO produced by MAR-11 cells that were co-transfected with CST-GFP12. Six other lanes show EPO produced by MAR-11 cells that were co-transfected with each of the six TGS mutants. With the exception of TGS-2, all TGS mutants showed no transport activity. The fact that mutant TGS-2 retained most of its transport activity explained why the transport activity was only partially lost in CST-GFP10. In CST-GFP10, the GFP was inserted into CST between I233 and K234 which is localized in the middle of tetrapeptide 233EIKE235 that was mutated in TGS-2. Therefore, GFP was not inserted into an important region in CST-GFP10. However, the GFP was inserted at a site that is close to a crucial region for activity (236KGFF239) that was mutated in TGS-3. All other TGS mutants completely abolished the transport activity in CST-GFP12 (Figure 2C). These results confirmed that these three loops are critical for the transport activity of CST.

To further narrow down the essential amino acids identified in the TGS mutants, we next replaced only two amino acids with Gly residues within the regions covered by the TGS mutants. For example, in di-Gly substitution-1 (DGS-1), two amino acids 112LD113 that are in the middle of the tetrapeptide 111NLDA114 (mutated in TGS-1) were substituted with two Gly residues. EPO/IEF results shown in Figure 2D suggest that mutating 112LD113 was enough to inactivate the transport activity of CST, suggesting that these two amino acids are essential for CST’s activity. However, mutating only two amino acids in DGS-3 (238FF239) and DGS-4 (268DN269) failed to inactivate the transporter (Figure 2D), suggesting that in these two regions mutating four (maybe three) amino acids is required in order to inactivate the transporter.

Lys272 and Ile270 in the fourth loop on the cytosolic side are crucial for CST activity

To investigate the role of single amino acids in these glycine substitution (GS) regions, amino acid residues covered by the TGS mutants were each mutated to an Ala residue (Supplemental Table 1). Analyzing the IEF patterns of EPO produced by MAR-11 cells co-expressing the Ala mutant constructs reveals that mutation of Lys272 to Ala caused CST-GFP12 to lose its activity completely (Figure 3A). All other Ala substitution mutants remained active. In a similar fashion, each of these amino acids was individually changed to a Gly residue to further analyze the impact of the side chain of each amino acid. Again, Lys272 substitution with Gly also completely inactivated the transporter (Figure 3B). CST-GFP12 with its Lys272 changed to Ala or Gly did not affect its stability or its Golgi localization (Figure 3C), suggesting that these two CST-GFP12 substitution mutants were properly expressed and localized to the Golgi. These results demonstrate that Lys272 is essential for CST function. A closer analysis on Figure 3B revealed that changing Ile270 to a Gly may have partially reduced the transport activity, although changing the same amino acid to Ala did not seem to obviously affect its activity (Figure 3A).

To further investigate the role of Lys272, it was also mutated to two other amino acids, Arg and His. CST-GFP12. It was changed to Arg because Arg also has a positively charged side chain. Lys was changed to His because His is the amino acid found in the homologous position in human GDP-fucose transporter. The activities of these mutants were analyzed with the EPO/IEF assay. As shown in Figure 4A, CST-GFP12 with a Lys272Arg mutation is as active as the wild-type transporter. However, when Lys272 was changed to a His residue, it completely inactivated the transporter. As shown in Figure 3B, the transporter lost partial activity when Ile270 was changed to Gly. However, it did not seem to affect the transport activity when Ile270 was changed to Ala (Figure 3A). These results suggest that the size of the side chain in Ile270 may affect CST’s activity. To confirm this hypothesis, Ile270 was mutated to Thr, which has a hydroxyl group, but the size of the side chain is similar to that in Ile. Ile270 was also mutated to Trp to enlarge the hydrophobic side chain. The transport activities of these mutants were analyzed using the EPO/IEF assay, and the results are shown in Figure 4B. Indeed, Ile270Thr mutant showed similar activity as the wild-type CST. The relative transport activity of these mutant constructs compared to the wild type are in the order of Ile270Trp < Ile270Gly < Ile270Ala < Ile270Thr < wild-type control. Therefore, the size of the side chain at this position does seem to have an impact on the activity of the transporter. These results further illustrate the sensitivity of the EPO/IEF assay.
The conserved Lys residue in UGT is also essential for its transport activity

CST and UGT are highly related NSTs. The Lys272 in CST is conserved in UGT (Lys297). To investigate the function of Lys297 in UGT, it was mutated to Arg, His, Ala and Gly in HA-tagged UGT (HA-UGT). The wild-type and the mutated UGTs were transiently transfected into the UGT-deficient Lec8 cells. The results from the EPO/IEF assay are shown in Figure 5. Similar to CST, Lys297 in UGT is indeed essential for its function. Unlike CST, UGT with a Lys297Arg mutation lost most of its activity. Similar to CST, substituting Lys297 in UGT with either His, Ala or Gly completely inactivates the transporter. The subcellular localization of the wild-type and mutant UGT was analyzed with an anti-HA antibody. The results showed that they are all expressed and shared similar localization (data not shown).

The conserved Lys residue in UGT is also essential for its transport activity

Since the two amino acids (Ile270 and Lys272) in CST and Lys297 in UGT are crucial for the transporter’s activities and they are localized on a loop facing the cytosol, we examined whether this loop is involved in specific substrate recognition. Figure 6A shows the sequence alignment between Chinese hamster CST and UGT at the region flanking the critical Lys residue. It was observed that CSTs across different mammalian species shared the same conserved amino acid sequence (underlined) compared to the conserved sequence in UGTs. A chimeric transporter was made by replacing the underlined sequence of CST with the underlined sequence of UGT to give rise to a mutated CST (CSTmt). Similarly, a mutated UGT (UGTmt) was generated by replacing the underlined sequence of UGT with that of CST. Chimeric constructs were expressed in MAR-11 and Lec8 cells by transient expression. EPO/IEF assay subsequently showed that CSTmt successfully restored EPO sialylation in MAR-11 cells, suggesting that replacing the underlined region in CST with that from UGT did not affect the activity of CST (Figure 6B). However, the mutated UGT, UGTmt, has lost most of its transport activity as UGTmt failed to complement the genetic defect in Lec8 cells as shown in Figure 6C. These results suggest that the amino acids flanking the conserved Lys residue (Lys297) in UGT are critical for its activity.

The role of highly conserved amino acids in the loop regions of CST and UGT on their transport activities

Sequence alignment of CSTs, UGTs, UDP-GlcNAc transporters and GDP-fucose transporters from different mammalian species revealed that some amino acids are highly conserved, and several of them are located in the loop regions. These conserved amino acids in CST are shown in Figure 7A. To investigate whether any of these conserved amino acids are crucial for the activities of CST and UGT, each of these 10 amino acids was mutated to Gly in both CST and UGT.
EPO/IEF results showed that replacing most of these amino acids with Gly did not affect transport activity with two exceptions. As shown in Figure 7B, replacing Leu136 in CST with Gly (L136G) completely inactivated the CST as it failed to sialylate EPO in MAR-11 cells. However, as shown in Figure 7C, UGT lost its activity when a different Leu (Leu224), corresponding to Leu199 in CST, was mutated to Gly as it failed to sialylate EPO in Lec8 cells. Partial loss of UGT activity was also observed when Ile221, corresponding to Val198 in CST, was replaced with Gly.

The results presented in this report are summarized in Figure 8. We have identified Lys272 and Leu136 as the single amino acid residues on the hydrophilic loop regions required for transport activity of CST. Mutating either of these residues leads to a complete loss in CST activity as indicated by the subsequent loss of EPO sialylation on the EPO/IEF assay. On the lumenal side, two regions are also essential for CST's activity. Mutating 112LD113 to two Gly residues or mutating 267TDNI270 or 236KGFF239 to four Gly residues completely inactivated the transporter. In UGT, two amino acids, Lys297 and Leu224, both on the cytosolic side, were identified as essential for its activity.

**Discussion**

It is essential to maintain sufficient amount of CMP-sialic acid in the Golgi lumen for proper sialylation of N- and O-linked glycans to occur. The degree of sialylation on recombinant EPO is positively correlated with the activity of CST in our IEF assay. While many factors, such as galactose transferase and sialyltransferase, can affect the overall sialylation of recombinant EPO, these were kept constant in our experiments as all these analyses were carried out in the same MAR-11 mutant cells. The only real difference in each transfection lies in the CST construct used. Therefore, the degree of sialylation on recombinant EPO directly reflects the activity of each CST construct.

Previous efforts to understand roles of single amino acid residues to the function of CST and UGT and their localization include analysis of point mutations found in CHO mutants Lec2 (Eckhardt et al. 1998) and Lec8 (Oelmann et al. 2001). With the 10 transmembrane domain model (Eckhardt et al. 1999) as a good estimate of the transporter topology (Caffaro and Hirschberg 2006), importance of Gly residues in trans-
membrane domains has been studied. (Eckhardt et al. 1998; Lim et al. 2008). Here, we have focused on the functional information of the hydrophilic loops of the CST and identified several residues crucial for transporter activity. A more detailed study will involve heterologous expression of these mutant CST and UGT constructs and reconstitution of transporters into proteoliposomes (Caffaro and Hirschberg 2006) to understand how these residues affect the transporter at the mechanisms level.

In this study, we seek to understand the function of different hydrophilic loops of the CST. The strategy employed in this work was to insert a GFP molecule into different loops to generate a panel of CST-GFP fusion proteins. The inserted GFP protein serves a dual purpose. First, the expression and localization of these CST-GFP fusion proteins in the cell can be easily analyzed by examining the green fluorescence of GFP. Second, insertion of GFP into a loop disrupts the structure of that particular loop which may result in loss of transport activity if the loop is crucial for its activity. The fusion protein constructs were generated based on the 10 transmembrane domain model of CST (Eckhardt et al. 1999). GFP has been used numerous times in literature to follow the localization of a protein of interest mostly by fusing GFP either to the N-terminus or the C-terminus of the protein. Inserting a GFP molecule that contains 239 amino acids into a loop of a CST molecule that contains only 336 amino acids may have a tremendous impact on the function of CST by possibly disrupting the topology of the CST. Immunofluorescence studies of the CST-GFP fusion proteins were carried out, and the activity of the CST-GFP was assayed through EPO/IEF assay in MAR-11 cells. Surprisingly, the insertion of GFP generally does not affect the localization and activity. Immunofluorescence staining results indicate that the fusion proteins are still localized to the Golgi (Supplementary Figure 1 and Figure 1C). Our functional analysis using the EPO/IEF assay revealed that insertion of GFP at three loops affected the activity of CST. These are the second and fourth lumenal loops and the fourth cytosolic loop of CST. Among the 13 CST-GFP fusion proteins, only CST-GFP10 failed to show any fluorescence, and the reason for this remains unknown. However, EPO expressed in MAR-11 cells that were co-transfected with this construct was partially sialylated, suggesting that inserting GFP into the fourth lumenal loop partially inactivated the transporter (Figure 1B). Our results are in good agreement with a previous report which suggests that the second lumenal and fourth cytosolic loops of CST play crucial roles because epitope insertions into these loops inactivated transport activity (Eckhardt et al. 1999).

Three loops of the CST, second and fourth lumenal loops and fourth cytosolic loop, were confirmed to be essential for its activity by performing Gly substitution experiments. Single amino acid substitution experiments showed that Lys272 on the fourth

Fig. 6. The regions flanking Lys272 in CST and Lys297 in UGT are not involved in substrate recognition. (A) The sequence alignment between CST and UGT at the region flanking the critical Lys residue is shown. To investigate the role of this region for substrate recognition, the underlined regions were exchanged between CST and UGT to generate CSTmt (CST containing the underlined fragment exchanged from UGT) and UGTmt (UGT containing the underlined fragment exchanged from CST). The arrow indicates the conserved Lys residue. (B) Constructs expressing CST, UGT, CSTmt or UGTmt were co-transfected with an EPO-expressing vector into MAR-11 cells. Recombinant EPO produced in each transfection was analyzed by IEF. Results show that CSTmt was able to restore the defect in MAR-11 cells. Exchanging the underlined region with UGT did not affect the activity of CST. (C) The same constructs shown in (B) were transfected into Lec8 cells. The results show that UGTmt, which is UGT exchanged for the underlined region with CST, lost most of its activity.
Fig. 7. The role of highly conserved amino acid residues in the loop regions in CST and UGT. (A) Ten amino acids in the loop regions of CST were found to be highly conserved among CSTs, UGTs, UDP-GlcNAc transporters and GDP-fucose transporters from different mammalian species. (B) Ten CST-GFP12 mutants were generated by replacing each of the conserved amino acids with a Gly residue. Each of the CST-GFP12 mutants was co-transfected with an EPO vector into MAR-11 cells. Recombinant EPO produced from each transfection was analyzed by IEF. The results show that mutation of Leu136 (L136G) in CST results in complete loss of its activity. Mutations in any other conserved amino acids did not affect the activity of the CST. (C) Similarly, 10 HA-UGT mutants were generated by replacing each of the conserved amino acids with a Gly residue. Transport activity of each UGT mutant was analyzed in Lec8 cells, and the results show that mutating Leu224 to a Gly (L224G) in UGT completely inactivated the transporter. Ile223 to Gly mutation (I223G) in UGT results in a partial loss of activity. Mutating any other conserved amino acids in UGT did not affect its activity.
cytosolic loop of CST was absolutely required for its function. As CST and UGT are highly homologous NSTs, the conserved Lys residue in UGT (Lys297) was also mutated by site-directed mutagenesis. Indeed, mutating Lys297 in UGT completely inactivated the transporter. Single Gly substitution experiments also revealed that Ile270 in CST may play an important role as Ile270Gly mutation resulted in a partial loss of its activity.

Sequence alignment analysis of CSTs, UGTs, GlcNAc transporters and GDP-fucose transporters from different mammalian species revealed 10 highly conserved amino acids in the loop regions. Mutagenesis studies revealed that substituting Leu136 in CST with Gly completely inactivated the transporter. However, mutating only two amino acids within these regions (CST-FF239 and 267TDNI270) to four Gly residues also inactivated the CST. However, mutating only two amino acids within these regions (CST-FF239 and 267TDNI270) did not affect CST’s activity.

Fig. 8. Summary of amino acid residues crucial for transport activity of CST identified in this study. Single amino acid substitution to replace either Leu136 or Lys272 with Ala or Gly completely inactivated the CST. Mutating Ile132Leu133 to two Gly residues also inactivated CST, whereas mutating either one of these two amino acids did not affect its activity. Mutating two tetrapeptide regions in the CST (CST-KGFF239 and 267TDNI270) to four Gly residues also inactivated the CST. However, mutating only two amino acids within these regions (238FF239 and 268DN269) did not affect CST’s activity.

In previous studies, the assessment of UGT activity was performed in Lec8 or Lec2 background, it is difficult to compare the exact insertion site of GFP in CST-GFP2 is between the N-terminus and C-terminus of the GFP, acting as two spacers between the CST and the GFP. A total of 13 CST-GFP fusion constructs were generated, each with a GFP flankned by two linker peptides inserted into different positions of the CST (Figure 1A). Please refer to Supplementary Table 1 for exact insertion sites. The constructs were generated by overlap polymerase chain reaction (PCR). Different mutations were later introduced into one of the CST-GFP fusion proteins, CST-GFP12, either by overlap PCR or by site-directed mutagenesis using the Site-Directed Mutagenesis Kit (Stratagene, Cedar Creek, TX). All the CST-GFP fusion constructs and mutated CST-GFP12 were cloned into the pcDNA3.1 vector between the EcoR I and Xba I sites. In order to ensure same translation efficiency, a Kozak sequence (GCCACCATG) was inserted into the fusion constructs.

In summary, using the GFP insertion strategy, we discovered that three specific hydrophilic loops of CST play crucial roles for its activity. Mutagenesis experiments identified the responsible amino acids in each of the three loops. Mutating Lys272 alone completely inactivates CST. Mutating the conserved Lys residue in UGT results in a complete inactivation of UGT. A sequence alignment revealed several highly conserved amino acids in the loop region. Mutagenesis experiments showed that Leu136 in CST is essential for its activity. Interestingly, mutating a different conserved amino acid in UGT (Leu224) results in a complete inactivation of UGT. These results provide new insights into the structural and functional relationship of CST and UGT.

Materials and methods
Cell culture
CHO MAR-11 cells were cultured in Dulbecco’s modified Eagle’s medium (DMEM, Invitrogen, Grand Island, NY) supplemented with 10% fetal bovine serum (FBS) (Invitrogen, Grand Island, N.Y.), at 37°C with 5% CO2. CHO Lec8 cells were a kind gift from Dr. Pamela Stanley (Albert Einstein College of Medicine, NY) and were cultured in α-MEM supplemented with proline (40 mg/L) (Invitrogen, Grand Island, NY) containing 10% FBS, at 37°C with 5% CO2.

Constructs to express CST-GFP fusion proteins
To investigate the role of each hydrophilic loop in the CST, GFP was inserted into different loops of the Chinese hamster CST (GenBank accession number: Y12074) as shown in Figure 1A. A linker peptide, SGGGGS, was placed at both the N-terminus and C-terminus of the GFP, acting as two spacers between the CST and the GFP. A total of 13 CST-GFP fusion constructs were generated, each with a GFP flankned by two linker peptides inserted into different positions of the CST (Figure 1A). Please refer to Supplementary Table 1 for exact insertion sites. The constructs were generated by overlap polymerase chain reaction (PCR). Different mutations were later introduced into one of the CST-GFP fusion proteins, CST-GFP12, either by overlap PCR or by site-directed mutagenesis using the Site-Directed Mutagenesis Kit (Stratagene, Cedar Creek, TX). All the CST-GFP fusion constructs and mutated CST-GFP12 were cloned into the pcDNA3.1 vector between the EcoR I and Xba I sites. In order to ensure same translation efficiency, a Kozak sequence (GCCACCATG) was inserted into the fusion constructs.
added upstream of the translation start codon in each construct. A construct that expresses HA-CST by adding a HA-tag to the N-terminal of CST was also cloned into the pcDNA3.1 vector.

Construction of Chinese hamster UGT expression vector
The open reading frame of Chinese hamster UGT (GenBank accession number: AF299335) was cloned into the pcDNA3.1 vector. This version of UGT is localized to the ER (Kabuss et al. 2005). A HA-tag was added to the N-terminus of the UGT for cellular localization. Different UGT mutants were created by mutagenizing this HA-UGT construct using the Site-Directed Mutagenesis Kit (Stratagene). All the constructs generated in this work are summarized in Supplementary Table 1.

Transient expression of human EPO in CHO MAR-11 cells
cDNA encoding the open reading frame of human EPO was cloned into pcDNA3.1. Before transfection, 5 × 10⁵ CHO MAR-11 cells were seeded into each well of six-well plates and cultured overnight. The EPO expression construct was transiently transfected into mutant MAR-11 cells using Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA) according to manufacturer’s protocol. In each transfection, 4 μg plasmid DNA and 10 μl Lipofectamine 2000 reagent were added to each well containing 2 ml of medium (DMEM, 10% fetal bovine serum). Forty-eight hours after the transfection, the media was collected and the EPO protein in the media analyzed by IEF/Western blot as described previously (Lim et al. 2008). To analyze the activity of CST-GFP fusion proteins, 2 μg plasmid DNA encoding human EPO and 2 μg plasmid DNA encoding a CST-GFP fusion protein or a mutant construct of CST-GFP12, were mixed and transfected into MAR-11 cells. For analysis of the UGT protein activity, 2 μg plasmid DNA encoding human EPO and 2 μg plasmid DNA encoding UGT protein were mixed and transfected into Lec8 cells.

IEF analysis of recombinant EPO
The sialylation patterns of EPO in different samples were analyzed by IEF followed by Western blot modified from previously published protocols (Lasne et al. 2002; Schriebel et al. 2006). The pH range for IEF was from 3 to 10. To ensure equal loading for IEF, the concentrations of EPO in different samples were predetermined by enzyme-linked immunosorbent assay using an EPO ELISA kit (Roche, Mannheim, Germany).

Immunofluorescence staining
MAR-11 cells were plated on glass coverslips and grown overnight before transfection. Constructs expressing different CST-GFP fusion proteins or CST-GFP12 with different mutations were transfected into MAR-11 cells. Immunofluorescent staining of Golgi apparatus with anti-giantin was performed according to previously published methods (Lim et al. 2008). Briefly, 24 h after transfection, cells were fixed in 4% paraformaldehyde, permeabilized with 0.1% Triton X-100 (Sigma, St. Louis, MO), and blocked in 10% normal goat serum (Invitrogen, Carlsbad, CA) for 30 min. Cells were then incubated with a 1:100 dilution of anti-HA monoclonal antibody (Sigma) and a 1:200 dilution of rabbit anti-giantin serum (Abcam, Cambridge, UK) in phosphate-buffered saline (PBS) with 1% bovine serum albumin (BSA) for 1 h. Cells were washed three times with PBS and subsequently incubated with goat anti-mouse antibody conjugated with Alexa Fluor 594 and goat anti-rabbit antibody conjugated with Alexa Fluor 488 (Invitrogen) in PBS with 1% BSA for 1 h. Cells were washed three times with PBS before mounting on glass slides with Prolong Gold antifade mounting medium (Invitrogen, Eugene, OR). Fluorescence images were obtained using a Carl Zeiss META confocal microscope. For cells that were transfected with CST-GFP fusion constructs, staining step with the anti-HA monoclonal antibody was omitted.

Supplementary data
Supplementary data for this article is available online at http://glycob.oxfordjournals.org/.

Acknowledgements
We would like to thank Dr. May May Lee and Ms. Natasha Ann Pereira for critical review of the manuscript. This work was funded by The Agency for Science, Technology and Research (A*STAR), Singapore.

Abbreviations
BSA, bovine serum albumin; CST, CMP-sialic acid transporter; CSTmt, mutated CST; DGS, di-Gly substitution; DMEM, Dulbecco’s modified Eagle’s medium; EPO/IEF, erythropoietin/isoelectric focusing; ER, endoplasmic reticulum; FBS, fetal bovine serum; GFP, green fluorescent protein; GS, glycine substitution; HA-tag, hemagglutinin-tag; NSTs, nucleotide-sugar transporters; PBS, phosphate-buffered saline; PCR, polymerase chain reaction; PNA, peanut agglutinin; TGS, tetra-Gly substitution; UGT, UDP-galactose transporter; UGTmt, mutated UGT.

References
Hellbusch CC, Sperandio M, Frommhold D, Yakubenia S, Wild MK, Popovici D, Vestweber D, Gröne HJ, von Figure K, Lübbe T, et al. 2007. Golgi GDP-


