2 Donor substrates

Monosaccharides are chemically inert molecules, thus they need to be activated to high-energy donor substrates for glycosyltransferase reactions. Typically, phosphate is used as reactive group, either alone or in the context of a nucleotide. Phosphate (P) acts as a versatile leaving group in the nucleophilic attacks involved in glycosylation reactions. At the endoplasmic reticulum (ER) membrane of eukaryotic cells, Man and Glc are also linked to the polyprenol dolichol-P (Dol-P). These Dol-P linked sugars are used as donor substrates for hydrophobic glycosyltransferases that are embedded in the ER membrane.

The biosynthesis of donor substrates usually begins by the phosphorylation of hexoses at C1 and C6 mediated by hexokinase and phosphomutase enzymes. Hexokinase is an ubiquitous activity found from bacteria to mammals. By phosphorylating Glc at C6, hexokinase prevents the diffusion of Glc out the cytosol, where it is further processed. Hexokinase is not specific for Glc, since Man is also phosphorylated at C6 by the enzyme. The liver has an additional enzyme, glucokinase, which is specific for Glc and is important for glycogen formation. The sugar Gal is directly phosphorylated at C1 by the galactokinase enzyme. Deficiency of galactokinase leads to a form of galactosemia, which is characterized by the development of cataract due to the accumulation of galactitol in the lens.

Phosphoglucomutase catalyzes the transition of phosphate in both C1 to C6 and C6 to C1 ways. The enzyme is important for the breakdown of glycogen where Glc-1-P is converted to Glc-6-P. In the liver, a Glc-6-phosphatase leads to the release of Glc, which drives the equilibrium of the phosphoglucomutase reaction towards Glc-6-P. This phosphatase is only found in the liver, which is the main organ regulating the levels of blood sugar. The deficiency of the main phosphoglucomutase isoform (PGM1) leads to a rare form of glycogen storage disease (GSD type XIV). Defects of glycogen breakdown mainly lead to hypoglycemia and muscle weakness, whereas some forms are severe and lead to death in infancy.

The phosphoglucomutase isoform PGM3 also acts on GlcNAc-6-P and a dedicated phosphomannomutase (PMM2) catalyzes the formation of Man-1-P from Man-6-P. Mutations in the PMM2 gene are the main cause of congenital disorders of glycosylation (CDG).

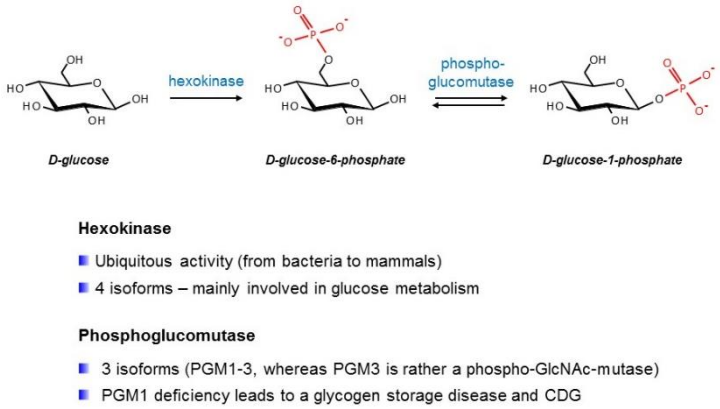


Figure 1 – Phosphorylation of Glc through hexokinase and phosphoglucomutase activity in the cytoplasm.

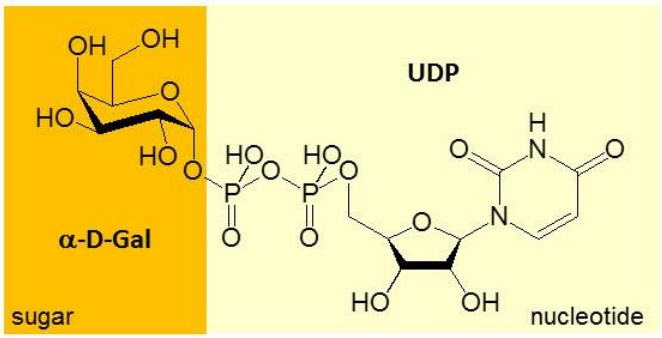
Nucleotide-activated sugars

The first nucleotide-activated sugar was identified by Leloir, who established UDP-Glc as substrate for the polymerization of Glc to glycogen. UDP is the most common nucleotide used to activate sugars in animals, as it is found linked to Glc, Gal, GlcNAc, GalNAc, GlcA and Xyl. Only two additional nucleotides are used in animals, GDP that is linked to Man and Fuc, and CMP that is linked to Sia. Whereas animal cells assemble their glycans from nine nucleotide-activated sugars, plants and in bacteria use many more substrates. More than 70 different nucleotide-activated sugars have been identified in bacteria to date.

The profusion of nucleotide-activated sugars in plants and bacteria is puzzling. Why is a single form of activated Glc not sufficient in plants, which use UDP-Glc as substrate for cellulose biosynthesis and ADP-Glc as substrate for starch formation? Similar apparent redundancies are observed in bacteria. Could such a fragmentation of nucleotide-activated monosaccharides be related to a defense strategy against parasites, viruses and bacteriophages? Such pathogens divert the host metabolism for their need. An increased complexity in metabolic pathways may enable host cells to cope with such pathogen-driven diversions. Accordingly, the proliferation of nucleotide-activated sugars may represent a form of metabolic immunity (see glycans and evolution).

Nucleotide-activated sugars are involved in the biosynthesis of large molecules like glycans, oligosaccharides and polysaccharides. Furthermore, the nucleotide-activated sugar UDP-GlcA is also widely used for the detoxification of xenobiotics by UDP-glucuronosyltransferases (UGT), which are mainly expressed in the liver in humans. The addition of GlcA to apolar molecules increases their solubility and thereby their elimination from the body.

Nucleotide-activated sugars are also important for the replication of some viruses. For example, the bacteriophages T2, T4 and T6 express glucosyltransferases that add Glc from UDP-Glc to hydroxymethylcytosine, thereby protecting the phage DNA from bacterial restriction enzymes. Similarly, baculoviruses express glucosyltransferases that transfer Glc from UDP-Glc to the molting hormone ecdysone, which inactivates the hormone and maintains the infected insect at a prepupal stage. Also, small molecules like the antibiotics aminoglycosides, e.g. neomycin, kanamycin and streptomycin, are synthesized from nucleotide-activated sugars.



UPD-Glc	GDP-Man	CMP-Sia
UDP-Gal	GDP-Fuc	
UDP-GlcNAc		
UDP-GalNAc		
UDP-GlcA		
UDP-Xyl		

Figure 2 – Structure of UDP-α-D-Gal and list of nucleotide-activated sugars found in animal cells.

In eukaryotes, nucleotide-activated sugars are synthesized in the cytosol with the exception of CMP-Sia, which is synthesized in the nucleus. The NDP-pyrophosphorylase reaction is readily reversible. However, pyrophosphate is rapidly hydrolyzed *in vivo* to orthophosphate by an inorganic pyrophosphatase. Therefore, the nearly irreversible hydrolysis of pyrophosphate drives the reaction towards the synthesis of the nucleotide-activated sugar product. The substrates for the synthesis of nucleotide-activated sugars are sugar-1-P and NDP. Here again, Sia does not follow the rule, since Sia itself and CMP are the substrates for the formation of CMP-Sia by the N-acylneuraminate cytidylyltransferase enzyme.

Not all nucleotide-activated sugars are built from monosaccharides units. For example, UDP-GlcA and UDP-Xyl are not synthesized from GlcA and Xyl, respectively, but from UDP-Glc, which is first oxidized to UDP-GlcA and then decarboxylated to UDP-Xyl. The so-called interconversion pathway plays a major role in maintaining stable intracellular levels of nucleotide-activated sugars under conditions where the supply of specific monosaccharides may be limited. In addition to the interconversion pathway, nucleotide-activated sugars can be

synthesized through the salvage pathway, i.e. through the breakdown of oligosaccharides in lysosomes and subsequent transfer of monosaccharides to the cytosol.

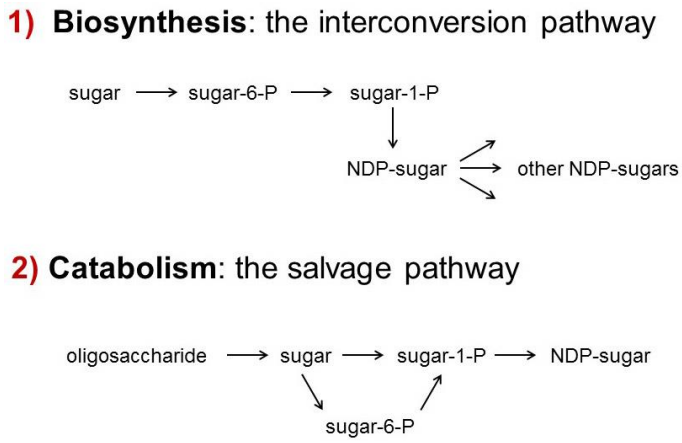


Figure 3 – Biosynthesis of nucleotide-activated sugars through the interconversion and salvage pathways.

The interconversion of nucleotide-activated sugars is achieved by multiple enzymatic activities, such as 4,6-dehydratase and 4-epimerase/reductase (involved in conversion of GDP-Man to GDP-Fuc) or 2-epimerase (involved in conversion of UDP-GlcNAc to UDP-ManNAc). Bacteria and plants express a large amount of interconverting enzymes in contrast to animals, which have fewer nucleotide-activated sugars.

Despite the limited set of interconverting activities in animals, these reactions are essential to the biosynthesis of nucleotide-activated sugars and thereby to the formation of glycoconjugates. For example, UDP-Glc can be formed from Glc-1-P and from the interconversion of UDP-Gal (through 4-epimerisation). In fact, the interconversion from UDP-Gal represents a major source of UDP-Glc in animal cells, as shown by the incorporation of [^{3}H]Glc in glycans after metabolic labeling of cells with [^{3}H]Gal. The interconversion of nucleotide-activated sugars in animal cells is controlled by rate-limiting enzymes, which are inhibited by the end-products of their respective pathways. Along this line, UDP-Glc dehydrogenase (UGD) is inhibited by UDP-Xyl, glutamine:fructose-6-P acetyltransferase (GFAT) by UDP-GlcNAc, UDP-GlcNAc 2-epimerase/kinase (GNE) by Sia and GDP-Man 4,6-dehydratase (GMDS) by GDP-Fuc.

Deficiency of interconverting enzymes often leads to metabolic disorders characterized by the accumulation of toxic intermediates and by defects of glycosylation. For example, mutations in the UDP-Gal 4-epimerase causes a mild form of galactosemia, a disease where patients have to lower their dietary galactose intake to reduce the accumulation of toxic intermediates such as galactitol.

Another defect of nucleotide-activated sugar interconversion is given by the deficiency of the rate-limiting enzyme UDP-GlcNAc 2-epimerase (GNE). The GNE protein includes two catalytic domains, the first one with the 2-epimerase

activity and the second one with a ManNAc kinase activity. Mutations in both domains lead to a reduced activity and thus to low biosynthesis of Sia. Surprisingly, the resulting disease, called Hereditary Inclusion Body Myopathy or Nonaka Myopathy, is rather mild considering the broad usage of Sia in glycoconjugates. The reason of this unexpected mild phenotype is probably related to the important contribution of the salvage pathway to generate enough CMP-Sia for incorporation into glycans. However, it is worth noting that the inactivation of the GNE gene in mice leads to embryonic lethality, which demonstrates that the importance of Sia biosynthesis for early embryonic development.

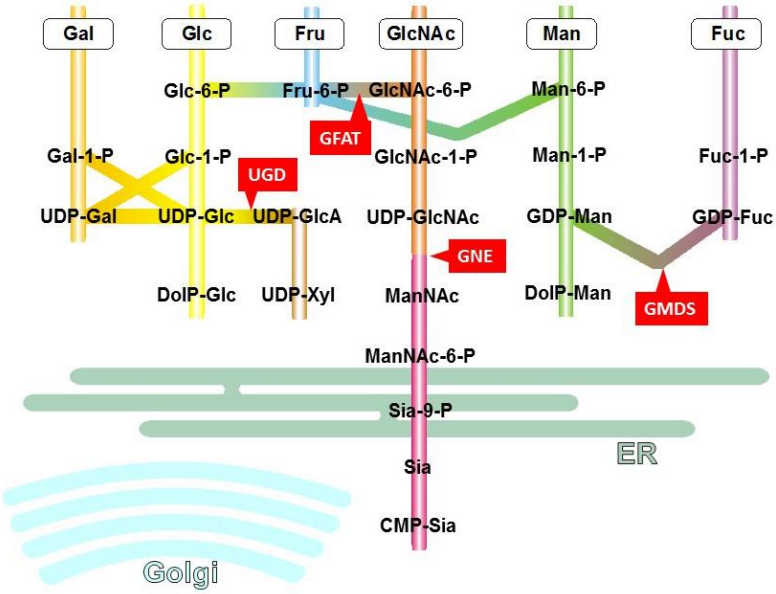
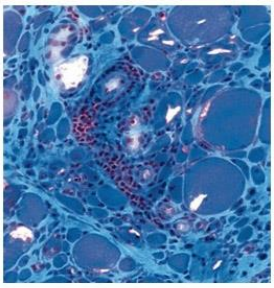


Figure 4 – NDP-sugar biosynthesis pathways. All reactions take place in the cytoplasma with the exception of the last step of CMP-Sia biosynthesis occuring in the nucleus. Rate-limiting enzymes are marked in red boxes. UGD: UDP-Glc dehydrogenase, GFAT: glutamine:fructose-6-P acetyltransferase, GNE: UDP-GlcNAc 2-epimerase/kinase, GMDS: GDP-Man 4,6-dehydratase.

As mentioned previously, GNE is a rate-limiting activity in the synthesis of Sia since the enzyme is inhibited by CMP-Sia. Some mutations in the GNE gene in fact do not inactivate the enzymatic activity itself but abolish the allosteric effect of CMP-Sia, thereby opening the gate to a production of Sia *ad libitum*. Needless to say that such an excess production of Sia is not healthy and leads to the disease sialuria, which is characterized by increased presence of Sia in urine (hence the name sial-uria), hepatosplenomegaly and airway infections.

Plants and bacteria produce a multitude of nucleotide-activated sugars and thereby rely extensively on interconverting enzymes to synthesize these donor substrates. These nucleotide-activated sugars are mainly involved in the biosynthesis of complex cell wall polysaccharides. Note that in plants and bacteria, a specific carbohydrate can be linked to different nucleotide. For example, in bacteria Glc is linked to UDP, GDP and TDP. In animals, carbohydrates are only found linked to a single type of nucleotides.



Hereditary inclusion body myopathy

- adult-onset muscular weakness
- affects mainly leg muscles (without quadriceps)
- origin and constitution of vacuoles unknown

Sialuria

- accumulation and excretion of free sialic acid
- hepatosplenomegaly
- frequent upper respiratory infections in infancy

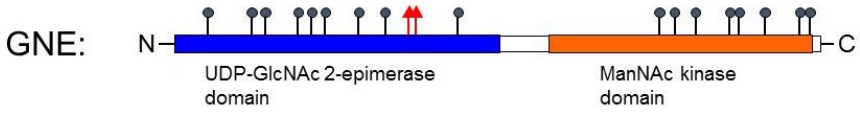


Figure 5 – GNE mutations are associated with two diseases characterized by different symptoms. Common mutations detected in hereditary inclusion body myopathy are marked with black circle on the schematic representation of the GNE protein. Mutations detected in sialuria are marked with red triangles.

A significant portion of carbohydrates are not synthesized *de novo* but recycled through salvage pathways. Monosaccharides like Glc and Gal are taken up in the gastrointestinal tract by dedicated transporters. By contrast, complex oligosaccharides, such as those found in maternal milk, are not digested and absorbed in the intestine. Within cells, glycoproteins are glycolipids are continuously broken down in lysosomes, thereby generating new building blocks for the synthesis of new molecules. The biological importance of glycan recycling in lysosomes is illustrated by lysosomal storage disorders such as Fabry's and Gaucher's diseases, in which deficiency of galactosidase and glucosidase enzymes lead to the toxic accumulation of glycolipids in lysosomes.

As mentioned previously in the context of Hereditary Inclusion Body Myopathy, Sia is also extensively recycled in humans through the salvage pathway. Furthermore, Sia appears to be taken up from food products in the intestine and incorporated in glycans. It has been shown that NeuGc, which cannot be synthesized in humans, is indeed incorporated in the glycans of multiple human cell types after ingestion of dietary products rich in NeuGc, such as red meat. The incorporation of NeuGc in human glycans modifies the susceptibility to bacterial toxins, such as the subtilase cytotoxin SubAB produced by Shiga toxigenic *Escherichia coli*, which rely on this type of Sia to target host cells.

Polyprenol-P-activated sugars

Not all donor substrates of glycosyltransferase reactions are nucleotide-activated sugars. Polyprenol-based lipids are commonly used in bacteria, archaea and eukaryotes to build donor substrates. The glycosyltransferases using polyprenol-based substrates are mainly hydrophobic proteins with multiple trans-membrane

domains. The number of isoprene repeats varies according to the organisms, prokaryotes having 11-12 isoprene units, yeasts 14-16 units, animals 16-21 units and plants 20-24 units. In Dol, the last isoprene unit is reduced. This reduction is required for the subsequent phosphorylation step. All polyprenol-based substrates are phosphorylated with the exception of the undecaprenol substrates involved in the elongation of the LPS outer core, which carry a pyrophosphate group.

Decaprenyl (C_{50})-P-sugar	Bacteria	Lipoglycan biosynthesis in mycobacteria
Undecaprenyl (C_{55})-P-sugar	Bacteria	O-antigen biosynthesis (LPS)
Dodecaprenyl (C_{60})-P-GalA	Bacteria	Lipid A modification in *Rhizobium leguminosarum*
Dolichol (C_{55-60})-P-sugars	Archaea	N-glycosylation
Dolichol (C_{70-120})-P-Man/Glc	Eukaryotes	N-glycosylation, GPI biosynthesis, O-mannosylation, C-mannosylation

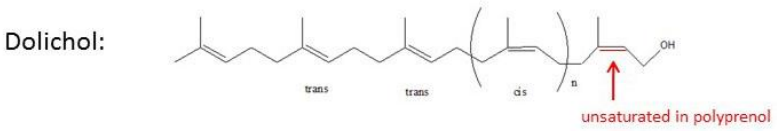


Figure 6 – Polyprenol-linked substrates in all domains of life. Dolichol has a saturated α-isoprene unit whereas prokaryotic polyprenol have an unsaturated bond.

Polyprenol and Dol share a common biosynthesis pathway with cholesterol up to the formation of farnesyl-PP. The basic isoprene unit is synthesized from acetyl-CoA, which is condensed to 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) and further reduced to mevalonate. The HMG-CoA reductase activity is rate limiting in the pathway, as it is regulated at the transcription and protein levels by sterol concentrations. The HMG-CoA reductase enzyme is also the target of inhibitors called statins, which are commonly prescribed as cholesterol-lowering drugs.

After reaching farnesyl-PP, the growing polyprenols are anchored in membranes and elongated by cis-prenyltransferase enzymes, also called dehydrodolichyl diphosphate synthase (DHDDS). The DHDDS enzymes are responsible for the variable chain length of polyprenols observed between prokaryotes and eukaryotes. Before receiving a monosaccharide unit, Dol undergo three modifications. First, the pyrophosphate group is removed by phosphatases, then the last isoprene unit is reduced and finally a phosphate group is added again by a dedicated kinase. The monosaccharides Man and Glc are attached to Dol-P in eukaryotes and various exotic sugars like GalA and Ara4N to polyprenols in bacteria. Dol-P-Man and –Glc are substrates in the final assembly steps of lipid-linked

oligosaccharides in the N-linked glycosylation pathway, in the biosynthesis of the glycosylphosphatidylinositol (GPI) anchor, in O-mannosylation and in C-mannosylation. In bacteria, polyprenol-P linked monosaccharides are substrates for the modification of the lipid A core and for the synthesis of arabinogalactan and lipoarabinomannan in mycobacteria. Deficiencies in cis-prenyltransferase, Dol reductase, Dol kinase and Dol-P-Man synthase lead to a shortage of Dol-P linked substrates and thus to disorders of glycosylation.

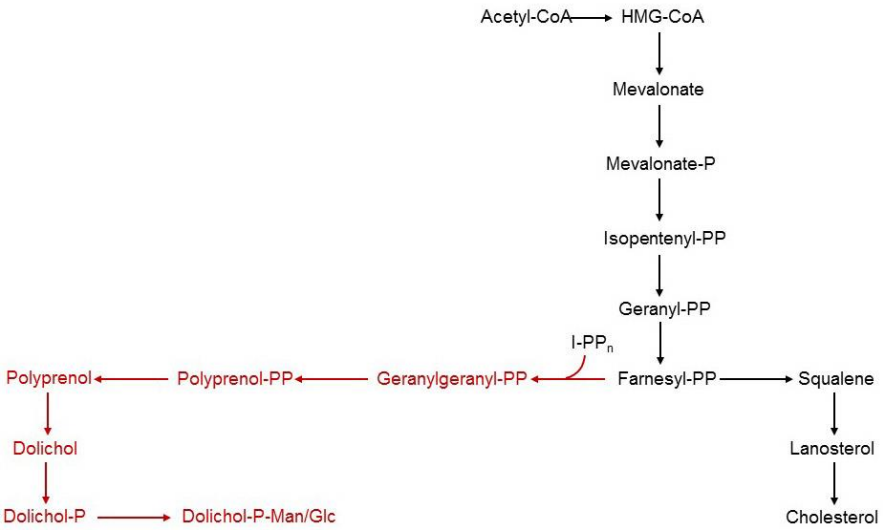
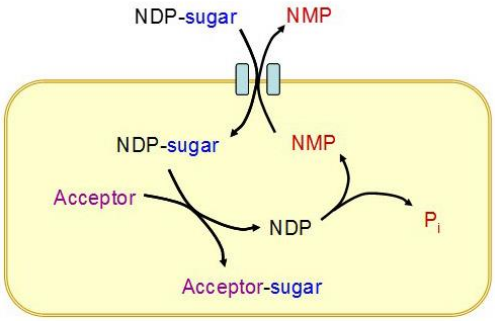


Figure 7 – Mevalonate and dolichol biosynthesis pathway.

Transporters

The transporters for nucleotide-activated sugars are actually antiporters, which import nucleotide-activated sugars and export the corresponding nucleotide-monophosphates into the cytosol. Most glycosyltransferase reactions yield a nucleotide-diphosphate, meaning that a phosphatase reaction is required to generate nucleotide-monophosphates. Most transporters are specific for a given nucleotide-activated sugar but some transporters enable the import of multiple substrates such as UDP-Gal, UDP-GalNAc and UDP-GlcA. The majority of transporters are embedded in the Golgi membrane but can be localized at the ER membrane in some cases. Data from the last three years have shown that some nucleotide-activated sugars, e.g. GDP-Fuc, are imported in the ER and in the Golgi through different transporters. For example, the inactivation of the Golgi GDP-Fuc transporter leads to a defect of terminal fucosylation in N- and O-glycans, whereas core O-fucosylation, which takes place in the ER, is not affected.

The substrate for glycan sulfation, 3'-phosphoadenosine 5'-phosphosulfate (PAPS), is also imported into the Golgi apparatus through an antiporter, which exports back AMP into the cytosol. A phosphate transporter ensuring the export of the phosphate anion to the cytosol is likely to exist, although such a transporter has not been identified yet.



Nucleotide	ER	Golgi
CMP-Sia	-	+++
GDP-Fuc	+	++++
UDP-Gal	+	++++
PAPS	-	++++
GDP-Man	-	++++
UDP-GlcNAc	++	++++
UDP-GalNac	++	++++
UDP-Xyl	++	++++
UDP-GlcA	++++	++++
UDP-Glc	++++	+

Figure 8 – Cellular distribution of nucleotide-activated sugar transporters. Distinct transporters can mediate the import of a nucleotide-activated sugar in the ER and Golgi apparatus (e.g. GDP-Fuc).