Structure and function of vertebrate CMP-sialic acid synthetase

Author
K. Muenster-Kuehnel, Anja, Tiralongo, Joe, Krapp, Stephan, Weinhold, Birgit, Ritz-Sedlacek, Valentina, Jacob, Uwe, Gerardy-Schahn, Rita

Published
2004

Journal Title
Glycobiology

DOI
https://doi.org/10.1093/glycob/cwh113

Copyright Statement
This is a pre-copy-editing, author-produced PDF of an article accepted for publication in Glycobiology following peer review. The definitive publisher-authenticated version Glycobiology 2004 14(10):43R-51R is available online at: http://dx.doi.org/10.1093/glycob/cwh113.

Downloaded from
http://hdl.handle.net/10072/5052

Link to published version
http://glycob.oxfordjournals.org/
Structure and function of vertebrate CMP–sialic acid synthetases

Anja K. Münster-Kühnel1, Joe Tiralongo3, Stephan Krapp4, Birgit Weinhold2, Valentina Ritz-Sedlacek2, Uwe Jacob5, and Rita Gerardy-Schahn1,2

1Institut für Biochemie, Abteilung Strukturforschung, Am Klopferspitz 19, D-82152 Martinsried, Germany; 2Institute for Glycomics, Griffith University, PMB 50 Gold Coast Mail Centre, QLD 9726, Australia; 3Proteros Biostructures GmbH, Am Klopferspitz 19, D-82152 Martinsried, Germany; and 4Max-Planck-Institut für Biochemie, Abteilung Strukturforschung, Am Klopferspitz 18a, D-82152 Martinsried, Germany

Received on September 22, 2003; revised on May 18, 2004; accepted on June 10, 2004

Activation of sugars into nucleotide sugars is critical for their entry into biosynthetic pathways. In eukaryotic cells, the activation of the acidic nine-carbon sugar sialic acid to CMP–sialic acid takes place in the cell nucleus, whereas all other nucleotide sugars are made in the cytoplasm. Molecular cloning of vertebrate CMP–sialic acid synthetases confirmed the nuclear localization and introduced new molecular tools for directly exploring the functional mechanisms of the enzymes, as well as the physiological relevance of their nuclear transport. Although major advances have been made in understanding structure–function relationships and defining elements involved in the nuclear transport, the riddle surrounding the physiological relevance of nuclear localization awaits resolution.

Key words: CMP-sialic acid synthetase/nuclear localization/sialic acids/structure-function aspects

Introduction

Sialic acids make up a family of over 50 related sugars derived from neuraminic acid (systematic name: 5-amino-3,5-dideoxy-D-glycero-D-galacto-2-nonulopyranos-1-onic acid) or deaminoneuraminic acid (KDN, systematic name: 3-deoxy-D-glycero-D-galacto-2-nonulopyranos-1-onic acid). The largest structural variation of naturally occurring sialic acids is at carbon 5, which can be substituted with either an acetamido, hydroxycetamido, or hydroxyl moiety to form N-acetylenuraminic acid (Neu5Ac), N-glycolylenuraminic acid (Neu5Ge), or KDN, respectively (Schauer and Kamerling, 1997). The three major sialic acid derivatives are shown in Figure 1.

Due to their exposed terminal position on glycoconjugates, sialic acids have been implicated in a variety of biological processes, including cell–cell recognition and communication processes (for reviews see Crocker and Varki, 2001; Munday et al., 1999; Schauer and Kamerling, 1997). Several tumors overexpress sialic acid and polysialic acid (polySia) (for reviews see Inoue et al., 1998; Malykh et al., 2001a; Scanlin and Glick, 2000), a linear homopolymer of α-2,8-linked sialic acids, and the increased sialic acid levels seem to correlate with the metastatic potential of these tumours (Daniel et al., 2000; Glüer et al., 1998). In bacteria, sialic acids and polySia are found as components of lipooligosaccharides and capsules and are often important virulence factors (reviewed in Hood et al., 1999; Jann and Jann, 1990; Moran et al., 1996).

Neu5Ac and KDN are believed to be the biosynthetic precursors for other members of the sialic acid family (for reviews see Schauer, 2000; Varki, 1992). The biosynthetic pathways providing Neu5Ac are well known in bacteria and vertebrates (for review see Bliss and Silver, 1996; Roseman, 1968). Even though the biosynthetic pathway leading to KDN seems to be very similar to that leading to Neu5Ac (Angata et al., 1999a,b), the primary structures of many of the enzymes involved in the biosynthesis and catabolism of KDN remain to be resolved. However, the activation of Neu5Ac, as well as KDN to its CMP diester (CMP-Neu5Ac and CMP-KDN, respectively), which is a prerequisite for the incorporation into glycoconjugates, is conserved from bacteria through humans. CMP–sialic acid is synthesized by the following reaction:

$$\text{CTP} + \text{sialic acid} \rightarrow \text{CMP–sialic acid} + \text{PPi}$$

This reaction is catalyzed by the CMP–sialic acid synthetases (EC 2.7.7.43, CMP-Neu5Ac synthetase and CMP-KDN synthetase, respectively). The product of this reaction, CMP–sialic acid, is, in eukaryotic cells, specifically transported into the Golgi apparatus, where it acts as a substrate for sialyltransferases.

The activation of sialic acids is unique in the following ways. (1) The majority of sugars are activated by GDP or UDP. Apart from sialic acids, activation by CMP only occurs in the case of 3-deoxy-D-manno-octulosonate (KDO), generating CMP-KDO (Ghalambor and Heath, 1966). KDO is an eight-carbon sugar and an essential component of gram-negative bacterial cell wall oligo- and some capsular polysaccharides. In addition KDO is found in plants and some green algae (for review see Royo et al., 2000). (2) Free sialic acids, not the phosphorylated forms, are used as substrates (Comb et al., 1966). (3) In contrast to all other eukaryotic nucleotide sugar synthetases, which are cytoplasmic enzymes, CMP–sialic acid synthetases are localized in the nucleus. The nuclear localization has been

1To whom correspondence should be addressed; e-mail: gerardy-schahn.rita@mh-hannover.de

Glycobiology vol. 14 no. 10 © Oxford University Press 2004; all rights reserved.
documented in a number of sophisticated studies, the earliest of which probably provides the most impressive example. Studying the expression of the sialic acid activating enzyme in ocular tissue E. L. Kean (1969) realized that only the nucleated cells in the lens epithelial layer and not the enucleated fiber cells derived from these cells express CMP-Neu5Ac synthetase activity. The multitude of biochemical studies addressing the issue of nuclear localization of the sialic acid–activating enzymes has been intensively reviewed (Kean, 1991; Kean et al., 2004; Vionnet et al., 1999).

Several hypotheses have been advocated to explain the unusual subcellular localization of the CMP–sialic acid synthetase, including the following. (1) The nuclear environment may be a prerequisite for enzymatic activity. (2) The CTP concentration may be higher in the nuclear compartment. (3) Activated sialic acids may be required for sialyltransferases localized in the nucleus (Richard et al., 1975). (4) Activation of Neu5Ac in the nucleus may protect the nucleotide sugar from subsequent modifications by cytoplasmic enzymes, such as the CMP-Neu5Ac hydrolase that generates Neu5Gc (Malykh et al., 2001b; Shaw and Schauer, 1988) or (5) from degradation by the CMP-Neu5Ac hydrolase (Kean and Bighouse, 1974; van Dijk et al., 1976). (6) Because CMP-Neu5Ac has been shown to be an allosteric inhibitor of UDP-GlcNAc 2-epimerase/kinase, the enzyme that controls Neu5Ac synthesis (Hinderlich et al., 1997; Kornfeld et al., 1964; Stäsche et al., 1997), sequestration of CMP-Neu5Ac in the nucleus may be necessary to prevent early inactivation of cytoplasmic UDP-GlcNAc 2-epimerase/kinase (Kornfeld et al., 1964). (7) The CMP-Neu5Ac synthetase may have a second, as yet unidentified function within the nucleus.

Cloning and characterization of vertebrate CMP–sialic acid synthetases

In 1998, the first mammalian CMP–sialic acid synthetase was cloned by a complementation approach. In the Chinese hamster ovary cell (CHO) mutant LEC29.Lec32, the lec32 defect inactivates endogenous CMP-Neu5Ac synthetase, leading to a lack of sialylated and polysialylated glycotopes on the cell surface (Potvin et al., 1995). Reappearance of polySia, which can be easily detected with monoclonal antibody 735 (for review see Mühlennonf et al., 1998) was used as an assay of complementation of the lec32 defect and led to the isolation of a mouse CMP-Neu5Ac synthetase cDNA (Münster et al., 1998).

The isolated murine enzyme is a protein of 432 amino acids with a calculated molecular mass of 48.1 kDa. The overall sequence identity between the murine and bacterial CMP-sialic acid synthetases ranges between 41% and 47% and, as shown in Figure 2A, is concentrated to five conserved amino acid stretches (Münster et al., 1998). This finding, which provides strong evidence for a common ancestral gene, received support through the cloning of the human CMP-Neu5Ac synthetase (Lawrence et al., 2001) and the CMP-KDN synthetase from rainbow trout (Nakata et al., 2001). Both enzymes also contain the five conserved motifs (Figure 2A). Thus the CMP–sialic acid synthetases provide the first identified sialic acid–metabolizing enzymes, which conserve specific sequence motifs from bacteria through humans. Therefore it was not surprising to see that the murine CMP-Neu5Ac synthetase cDNA complements Escherichia coli EV5, a polySia-capsule negative mutant of the neuroinvasive E. coli K1 (Vimr and Troy, 1985). The capsular defect in E. coli EV5 results from a genetic defect in the endogenous CMP-Neu5Ac synthetase (Vimr et al., 1989). Expression of the murine CMP-Neu5Ac synthetase reconstituted the K1 phenotype (Münster et al., 1998, 2002).

In silico analyses of the existing gene databases carried out by Angata and Varki (2002) revealed homologous genes also in certain types of insects, archaea, and bacteria, but the enzymatic activity of these proteins remains to be elucidated. Remarkably, high sequence similarities (39–58%) were found between CMP-sialic acid and CMP-KDO synthetases (EC 2.7.7.38). As shown in Figure 2B, three out of the five conserved motifs are present in the family of CMP-KDO synthetases. Motif I has been shown to be involved in nucleotide binding in both CMP-KDO (Jelakovíc and Schulz, 2001) and CMP-sialic acid synthetases (Krapp et al., 1996; Tullius et al., 1996). The described similarities prompted the suggestion of phylogenetic relationships between the CMP-KDO and CMP-sialic acid synthetases (Angata and Varki, 2002). Both enzyme groups are, however, highly specific for their substrates (Ambrose et al., 1992; Kean and Roseman, 1966; Kohlbrenner et al., 1987) and, unlike the CMP–sialic acid synthetases, the only available eukaryotic CMP-KDO synthetase from Zea mays, has been suggested to localize in endomembranous system of the plant cell (Royo et al., 2000).

In Figure 3 a phylogenetic tree is shown, which includes the confirmed CMP–sialic acid and CMP-KDO synthetase.
sequences. The tree indicates a common ancestor for CMP–sialic acid and CMP-KDO synthetases. Within the family of CMP–sialic acid synthetases, the vertebrate genes cluster into a separate branch, but the tree clearly demonstrates phylogenetic connection with the bacterial enzymes. These data confirm an earlier extended phylogenetic study by Angata and Varki (2002), which includes two enzymes of the sialylation pathway (Neu5Ac([-9-phosphate]) synthase and CMP–sialic acid synthetase). Based on their data, the authors suggested that genes involved in the biosynthesis of sialic acids predate the split of deuterostomes (vertebrates, ascidians, and echinoderms) and protostomes (arthropods and mollusks), possibly even the split of the three domains of life (Angata and Varki, 2002).

KDN, which was first isolated from rainbow trout egg polysialoglycoproteins (Nadano et al., 1986), is now generally accepted as being an ubiquitous component of vertebrate glycoconjugates (Angata et al., 1998, 1999b; Inoue et al., 1998; Ziak et al., 1996). Different from the other sialic acid derivatives, KDN is made by an independent biosynthetic pathway. The cDNA of the CMP-KDN synthetase has been isolated from rainbow trout testis (Nakata et al., 2001). The protein displays 53.8% identity with the murine enzyme, including the five structural important motifs (Figure 2A, motifs I–V, Nakata et al., 2001), however, comparative activity tests carried out with the recombinant bacterial expressed murine and fish enzymes showed that the fish enzyme exhibits a broader substrate spectrum. The fish CMP-KDN synthetase efficiently synthesizes CMP-KDN and CMP-Neu5Ac, whereas the murine enzyme preferentially activates Neu5Ac (factor 15 over KDN) (Nakata et al., 2001). These findings are consistent with earlier results obtained with partially purified CMP-Neu5Ac synthetase (Higa and Paulson, 1985) and CMP-KDN synthetase (Terada et al., 1993) from calf brain and rainbow trout, respectively. Because a second CMP–sialic acid synthetase has not been identified in the murine or human genome, the low affinity of the mammalian CMP–sialic acid synthetases for KDN may explain the trace amount of KDN found in animal glycoconjugates. Like the mammalian enzyme, the fish enzyme has been suggested to be a nuclear resident (Nakata et al., 2001; Terada et al., 1993).

The human CMP–sialic acid synthetase (94% identity to the murine enzyme) was isolated by Lawrence and co-workers (2001) and has been shown to be transported to the nucleus if expressed in insect cells. In vitro testing of activity demonstrated that Neu5Ac and Neu5Gc are the preferred substrates also of the human enzyme (Lawrence et al., 2001).

Nuclear transport of vertebrate sialic acid synthetases

Most nuclear proteins with molecular weights of greater than 40 kDa must be actively transported through the nuclear pore complex (Yoneda, 2000). Various pathways are used simultaneously in the cell to achieve the nuclear import of different substrates (for reviews see Görlich and Kutay, 1999; Wente, 2000; Yoneda, 2000). Often nuclear transport depends on the existence of a nuclear localization signal (NLS) sequence, which, in the case of canonical NLSs, are short stretches of basic amino acids, so-called basic clusters (BCs) recognized by soluble factors of the nuclear import system (for reviews see Barry and Wente, 2000; Görlich and Mattaj, 1996). NLS motifs are not cleaved from the mature protein and do not fit a tight consensus motif (for review see Chook and Blobel, 2001; Garcia-Bustos et al., 1991).

Sequence analysis revealed one putative NLS in the human (Lawrence et al., 2001), two in the fish (Nakata et al., 2001), and three in the murine CMP–sialic acid synthetases.

### Table: Vertebrate CMP–sialic acid synthetases

<table>
<thead>
<tr>
<th>Species</th>
<th>motif I</th>
<th>motif II</th>
<th>motif III</th>
<th>motif IV</th>
<th>motif V</th>
<th>acc. nr</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homo sapiens</td>
<td>456</td>
<td>457</td>
<td>458</td>
<td>459</td>
<td>460</td>
<td>A054215</td>
</tr>
<tr>
<td>Mus musculus</td>
<td>461</td>
<td>462</td>
<td>463</td>
<td>464</td>
<td>465</td>
<td>A054215</td>
</tr>
<tr>
<td>B. musculus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homo sapiens</td>
<td>466</td>
<td>467</td>
<td>468</td>
<td>469</td>
<td>470</td>
<td>A054215</td>
</tr>
<tr>
<td>Mus musculus</td>
<td>471</td>
<td>472</td>
<td>473</td>
<td>474</td>
<td>475</td>
<td>A054215</td>
</tr>
</tbody>
</table>

**Fig. 2.** Areas representing conserved amino acid stretches as identified in all confirmed CMP–sialic acid synthetases (A) and CMP-KDO synthetases (B) are shown in a multisequence alignment blot. The sequences have been aligned to display maximum homology (MULTALIN 5.4.1; Corpet, 1988) and the consensus motif (for review see Chook and Blobel, 2001; Görlich and Mattaj, 1996). NLS motifs are not cleaved from the mature protein and do not fit a tight consensus motif (for review see Chook and Blobel, 2001; Garcia-Bustos et al., 1991).
synthetase (Münster et al., 2002). With the help of deletion mutants, the NLS responsible for nuclear import of the murine CMP-Neu5Ac synthetase could be narrowed down to BC2 (Münster et al., 2002). The sequence in BC2: K<sup>198</sup>RPRR fits well with the four-residue NLS motif suggested by Chelsky (K-R/K-X-R/K, Chelsky et al., 1989) and is strictly conserved in the human enzyme and highly conserved in the rainbow trout CMP-KDN synthetase (Figure 4A). In fact, K<sup>200</sup>RPR has been suggested to be the human NLS (Lawrence et al., 2001). A detailed study carried out by site-directed mutagenesis demonstrated that each basic amino acid residue in BC2 is of importance for the correct nuclear import of the murine enzyme (Münster et al., 2002). In contrast, no information is available to date on the factors that mediate the nuclear transport of CMP-Neu5Ac synthetases.

As mentioned before, the nuclear residence of the vertebrate CMP–sialic acid synthetases has been frequently shown by subcellular fractionation techniques (for reviews see Kean 1991; Kean et al., 2004) and could be confirmed in transfection studies with the use of the epitope-tagged CMP–sialic acid synthetase cDNAs (Lawrence et al., 2001; Münster et al., 1998, 2002; Nakata et al., 2001). The biological relevance of the nuclear transport remains unclear, particularly because it has been demonstrated that the functional activity of the enzyme is independent of the nuclear transport. Cytoplasmic variants of the epitope-tagged murine CMP-Neu5Ac synthetase (point mutants in which Lys-198 and Arg-201 in BC2 were exchanged by alanine) were made and transfected into CMP–sialic acid synthetase–deficient LEC29.Lec32 cells (Potvin et al., 1995). Full reconstitution of the wild-type sialylation pattern was observed with all mutants proven to be active in a parallel bacterial test system (complementation of E. coli EV5; Münster et al., 2002). These studies carried out at the cellular level suggest that nuclear localization of CMP–sialic acid synthetase is not essential for its function as a sugar-activating enzyme. Respecting the possibility that CMP–sialic acid synthetases may have additional functions in the nucleus of the vertebrate cell, the situation may be different if the enzyme is falsely located in the multicellular organism.

---

**Fig. 3.** Phylogenetic analysis of functionally confirmed CMP–sialic acid and CMP-KDO synthetase sequences. The analysis has been carried out with ClustalX (Thompson et al., 1997), using the BLOSUM matrix and the phylogenetic tree was calculated using the neighbor joining (NJ) method of Saitou and Nei (1987). The accession numbers of the sequences are shown in Figure 2.
Structure–function relationships in CMP–sialic acid synthetases

When the guidance functions of BC2: K198RPRR inside the cell was studied in more detail, it became obvious that the NLS (BC2) contains elements that are crucial also for the catalytic function of the murine CMP-Neu5Ac synthetase. Single point mutations introduced in BC2 (Arg-199 to Ala and Arg-202 to Ala; see Figures 4A, 5B and 5C) completely abolished the catalytic activity. The exchange Arg199Ala was additionally accompanied by cytoplasmic retention, whereas the mutant Arg202Ala predominantly (490%) localized in the nucleus (Münster et al., 2002), indicating that activity and nuclear transport can be dissociated. In Figure 4 a multialignment highlights the area surrounding BC2 (shaded in gray) in the murine enzyme. Sequences of all known CMP–sialic acid (Figure 4A) and CMP-KDO synthetases (Figure 4B) are compared. The alignment reveals that Arg-202 (shown in red), essential for both activity and nuclear transport, is highly conserved throughout the family, with only one exception, Legionella pneumophila, where a cysteine residue substitutes the arginine. Legionaminic acid, so far not identified in vertebrate tissues, is a nine-carbon keto sugar structurally similar to sialic acid, and the identified enzyme is able to complement the CMP–sialic acid synthetase deficiency in E. coli EV5 (Lüneberg et al., 2000). All other amino acid residues that make up BC2 are conserved only in the eukaryotic sequences.

The crystal structures have been resolved for the full-length NmB (Mosimann et al., 2001) and the catalytic domain of the murine CMP-Neu5Ac synthetase (Krápp et al., 2003). In both enzymes the homodimer provides the functional active unit. The structure of the murine enzyme is shown in Figure 5A. Laterally organized monomers assemble into functional dimers. The dimerization domains, which comprise the essential BC2, extend out of the central β-sheets, intertwine and form the substrate loops of the opposing monomers, giving rise to composed active sites (see Figure 5B). Different from the NmB enzyme the asymmetric unit in the case of the murine enzyme is four copies of the monomer (Figure 5A). The tetramer represents a dimer of the functional dimer, stabilized by a four-helix bundle in the center and hydrophobic interactions between the monomers A and C and B and D (Krápp et al., 2003). Tetramer formation explains the migration behavior observed while purifying the enzyme from different species and tissues (for reviews see Kean, 1991; Kean et al., 2004).

The crystal structure of the catalytic domain of the murine enzyme has been obtained in its closed conformation with the product CMP-Neu5Ac bound, making visible the contacts formed between the nucleotide sugar and the protein (Figure 5B). The active site is located at the interface of the core domain of monomer A (ochre) and the dimerization domain of monomer B (blue). The two crucial residues for enzymatic activity, Arg-199 and -202, are part of the substrate loop (monomer B) and make polar contacts to the sugar moiety. Arg-199 interacts with the carbonyl oxygen of the N-acetyl group, and the identified enzyme is able to complement the CMP–sialic acid synthetase deficiency in E. coli EV5 (Lüneberg et al., 2000). All other amino acid residues that make up BC2 are conserved only in the eukaryotic sequences.

Crystal structure data about the binding of the sugar moiety have been obtained only in the case of the murine enzyme (Figure 5B). Different from the NmB enzyme the asymmetric unit in the case of the murine enzyme contains four copies of the monomer (Figure 5A). The tetramer represents a dimer of the functional dimer, stabilized by a four-helix bundle in the center and hydrophobic interactions between the monomers A and C and B and D (Krápp et al., 2003). Tetramer formation explains the migration behavior observed while purifying the enzyme from different species and tissues (for reviews see Kean, 1991; Kean et al., 2004).

The crystal structure of the catalytic domain of the murine enzyme contains four copies of the monomer (Figure 5A). The tetramer represents a dimer of the functional dimer, stabilized by a four-helix bundle in the center and hydrophobic interactions between the monomers A and C and B and D (Krápp et al., 2003). Tetramer formation explains the migration behavior observed while purifying the enzyme from different species and tissues (for reviews see Kean, 1991; Kean et al., 2004).

The crystal structure of the catalytic domain of the murine enzyme has been obtained in its closed conformation with the product CMP-Neu5Ac bound, making visible the contacts formed between the nucleotide sugar and the protein (Figure 5B). The active site is located at the interface of the core domain of monomer A (ochre) and the dimerization domain of monomer B (blue). The two crucial residues for enzymatic activity, Arg-199 and -202, are part of the substrate loop (monomer B) and make polar contacts to the sugar moiety. Arg-199 interacts with the carbonyl oxygen of the N-acetyl group, and the side chain of the highly conserved Arg-202 coordinates the carbohydrate group at position one of the sugar. Similarly, Arg-165 in the NmB enzyme, corresponding to Arg-202 in the murine enzyme, has been demonstrated to be essential for Neu5Ac binding and stabilization of the transition state in the NmB enzyme (Mosimann et al., 2001).

Crystal structure data about the binding of the sugar moiety have been obtained only in the case of the murine enzyme (Figure 5B). Different from the NmB enzyme the asymmetric unit in the case of the murine enzyme contains four copies of the monomer (Figure 5A). The tetramer represents a dimer of the functional dimer, stabilized by a four-helix bundle in the center and hydrophobic interactions between the monomers A and C and B and D (Krápp et al., 2003). Tetramer formation explains the migration behavior observed while purifying the enzyme from different species and tissues (for reviews see Kean, 1991; Kean et al., 2004).

The crystal structure of the catalytic domain of the murine enzyme has been obtained in its closed conformation with the product CMP-Neu5Ac bound, making visible the contacts formed between the nucleotide sugar and the protein (Figure 5B). The active site is located at the interface of the core domain of monomer A (ochre) and the dimerization domain of monomer B (blue). The two crucial residues for enzymatic activity, Arg-199 and -202, are part of the substrate loop (monomer B) and make polar contacts to the sugar moiety. Arg-199 interacts with the carbonyl oxygen of the N-acetyl group, and the side chain of the highly conserved Arg-202 coordinates the carbohydrate group at position one of the sugar. Similarly, Arg-165 in the NmB enzyme, corresponding to Arg-202 in the murine enzyme, has been demonstrated to be essential for Neu5Ac binding and stabilization of the transition state in the NmB enzyme (Mosimann et al., 2001).

Crystal structure data about the binding of the sugar moiety have been obtained only in the case of the murine enzyme (Figure 5B). Different from the NmB enzyme the asymmetric unit in the case of the murine enzyme contains four copies of the monomer (Figure 5A). The tetramer represents a dimer of the functional dimer, stabilized by a four-helix bundle in the center and hydrophobic interactions between the monomers A and C and B and D (Krápp et al., 2003). Tetramer formation explains the migration behavior observed while purifying the enzyme from different species and tissues (for reviews see Kean, 1991; Kean et al., 2004).

The crystal structure of the catalytic domain of the murine enzyme has been obtained in its closed conformation with the product CMP-Neu5Ac bound, making visible the contacts formed between the nucleotide sugar and the protein (Figure 5B). The active site is located at the interface of the core domain of monomer A (ochre) and the dimerization domain of monomer B (blue). The two crucial residues for enzymatic activity, Arg-199 and -202, are part of the substrate loop (monomer B) and make polar contacts to the sugar moiety. Arg-199 interacts with the carbonyl oxygen of the N-acetyl group, and the side chain of the highly conserved Arg-202 coordinates the carbohydrate group at position one of the sugar. Similarly, Arg-165 in the NmB enzyme, corresponding to Arg-202 in the murine enzyme, has been demonstrated to be essential for Neu5Ac binding and stabilization of the transition state in the NmB enzyme (Mosimann et al., 2001).

Crystal structure data about the binding of the sugar moiety have been obtained only in the case of the murine enzyme (Figure 5B). Different from the NmB enzyme the asymmetric unit in the case of the murine enzyme contains four copies of the monomer (Figure 5A). The tetramer represents a dimer of the functional dimer, stabilized by a four-helix bundle in the center and hydrophobic interactions between the monomers A and C and B and D (Krápp et al., 2003). Tetramer formation explains the migration behavior observed while purifying the enzyme from different species and tissues (for reviews see Kean, 1991; Kean et al., 2004).

The crystal structure of the catalytic domain of the murine enzyme contains four copies of the monomer (Figure 5A). The tetramer represents a dimer of the functional dimer, stabilized by a four-helix bundle in the center and hydrophobic interactions between the monomers A and C and B and D (Krápp et al., 2003). Tetramer formation explains the migration behavior observed while purifying the enzyme from different species and tissues (for reviews see Kean, 1991; Kean et al., 2004).

The crystal structure of the catalytic domain of the murine enzyme contains four copies of the monomer (Figure 5A). The tetramer represents a dimer of the functional dimer, stabilized by a four-helix bundle in the center and hydrophobic interactions between the monomers A and C and B and D (Krápp et al., 2003). Tetramer formation explains the migration behavior observed while purifying the enzyme from different species and tissues (for reviews see Kean, 1991; Kean et al., 2004).

The crystal structure of the catalytic domain of the murine enzyme contains four copies of the monomer (Figure 5A). The tetramer represents a dimer of the functional dimer, stabilized by a four-helix bundle in the center and hydrophobic interactions between the monomers A and C and B and D (Krápp et al., 2003). Tetramer formation explains the migration behavior observed while purifying the enzyme from different species and tissues (for reviews see Kean, 1991; Kean et al., 2004).

The crystal structure of the catalytic domain of the murine enzyme contains four copies of the monomer (Figure 5A). The tetramer represents a dimer of the functional dimer, stabilized by a four-helix bundle in the center and hydrophobic interactions between the monomers A and C and B and D (Krápp et al., 2003). Tetramer formation explains the migration behavior observed while purifying the enzyme from different species and tissues (for reviews see Kean, 1991; Kean et al., 2004).
Amino acid residues involved in nucleotide binding can be superimposed in the different crystal structures and in the case of the β- and γ-phosphate coordinating residues provide highly conserved positions at primary sequence level (motif I; see also Figure 2). The same is true for the coordination of the carboxylate group at Cl of the sialic acid. The polar contact made by Arg-202 in the murine enzyme is substituted by Arg-165 in the NmB synthetase and by Arg-155 in the E. coli CMP-KDO synthetase. In contrast, trials to overlay the residues involved in the binding of the sugar moiety (shaded residues in Figure 5C) demonstrate that this part of the proteins is less conserved not only in terms of primary sequence but also with respect to the 3D structure. Although the highly conserved Gln-residue of motif III (Gln-141 in Mus musculus; Gln-104 in NmB) has been demonstrated to be involved in the coordination of the O8 and N5 atoms by either modeling (NmB) or structural (murine enzyme) data, the trial to overlay this position in the crystal structures failed (this study). Similarly, imperfect overlay was obtained for the residues involved in the formation of the hydrophobic pocket that accommodates the methyl-group of the 5N-acetyl group. The partial failure of positional identity in this area correlates with the rather low conservation at primary sequence level and may explain differences in the substrate specificities between the murine and the NmB enzyme. Molecular reasons for differences in the substrate specificities observed between the murine (mainly active with Neu5Ac) and the closely related rainbow trout enzyme (active with both Neu5Ac and KDN) are presently not available. However, also in this case the dissimilarity at the dimer interface may be responsible for the changes in the substrate recognition.

In the subfamily of CMP-KDO synthetases, the highly conserved arginine (Arg-202 in mouse) corresponds to Arg-155 in E. coli (KSU5) and is part of a conserved motif (open box in Figure 4B), known from crystal structure data to be involved in the dimerization of the enzyme and in the formation of the active site (Jelakovic et al., 1996; Jelakovic and Schulz, 2001, 2002). Significant similarities in the overall structure and in the nucleotide binding pocket (see Figure 5C) support an evolutionary relationship. However, the dimerization interfaces and therefore the sugar-binding pockets in CMP–sialic acid and CMP-KDO synthetases are different and explain the strict substrate specificities of the two enzyme classes.

**Conclusions and perspectives**

Considerable progress has been made toward understanding structure–function relationships in vertebrate CMP–sialic acid synthetases and in defining elements involved in the nuclear transport of the enzymes. The obtained data allow the revision of some of the hypotheses that have been put forward to explain the enzymes nuclear residence. Both the observation that in CHO cells CMP–sialic acid synthetase activity is independent of the nuclear environment and the fact that wild-type sialylation pattern could be restored in LEC29.Lec32 cells argue against (1) a limiting CTP concentration in the cytoplasm, (2) premature feedback inhibition of UDP-GlcNAc 2-epimerase/kinase (Hinderlich et al., 1997; Kornfeld et al., 1964), and (3) increased degradation of CMP–sialic acid by the cytoplasmically localized hydrolase (Kean and Bighouse, 1974). On the other hand, it must be kept in mind that studies available to date do not exceed the cellular level, where several defects in the sialylation pathway have been described. Like LEC29.Lec32, cells of the complementation group Lec2, lacking endogenous CMP–sialic acid transporter (Deutscher et al., 1984; Eckhardt et al., 1998), grow well in the absence of sialic acids. In contrast, the creation of an asialo phenotype in an animal model by the inactivation of the UDP-GlcNAc 2-epimerase/kinase, the key enzyme in the biosynthetic pathway of sialic acids, has proven to be irreconcilable with development (Schwarzkopf et al., 2002). Embryos die at day 9 of gestation. Similarly, changes in the destination of the CMP–sialic acid synthetase could have a more pronounced effect in the multicellular organism.

The generation of transgenic mouse models is necessary to finally reveal the systemic relevance of the nuclear localization of CMP–sialic acid synthetases in animals. Moreover, the identification of potential interaction partners of the CMP-Neu5Ac synthetase in the cell nucleus may give a hint toward a possible second function. Last but not least, the bulk of genetic information and the associated ease of obtaining additional structure–function information for these enzymes should provide an efficient path toward answering the open questions with respect to substrate specificity and phylogenetic occurrence.

**Abbreviations**

BC, basic cluster; CHO, Chinese hamster ovary; KDN, 2-keto-3-deoxy-d-glycero-d-galacto-nononic acid; KDO, 3-deoxy-d-manno-octulosonate; NCAM, neural cell adhesion molecule; NLS, nuclear localization signal.

**References**

Ambrose, M.G., Freese, S.J., Reinhold, M.S., Warner, T.G., and Vann, W.F. (1992) 13C NMR investigation of the anomeric specificity of Fig. 5. (A) Crystal structure of the catalytic domain (amino acid residues 39–267) of the murine CMP-Neu5Ac synthetase. The asymmetric unit contains four copies of the monomer (ABCD). The laterally organized subunits AB and CD provide functional dimers and the tetramer is formed via dimerization of the functional unit. (B) Structure of the active site of murine CMP-Neu5Ac synthetase in complex with the product CMP-Neu5Ac. The active site is of composed nature of the highlighted in yellow (monomer A) and blue (monomer B). While the nucleotide dips into the hydrophobic pocket formed by amino acid residues of the monomer A, the sugar part is coordinated by amino acid residues of the substrate loop (blue) provided by monomer B. In accordance with results obtained by site-directed mutagenesis, residues Arg-199 and Arg-202 were found to be essential components in the enzymes active site. (C) Amino acid residues involved in the binding of the substrates CTP and sialic acid are highlighted for the murine (blue) and the NmB (red) CMP–sialic acid synthetases and are complemented by the relevant positions coordinating the substrates in the E. coli CMP-KDO synthetase (magenta). Highly conserved positions are marked by asterisks and the crucial arginine (202 in the murine enzyme) is shown in green. Amino acid residues known to fulfill identical functions in the enzymes but are differentially located in the 3D structures are shaded.


