Murine CMP-Sialic acid Synthetase:
Structural Analysis of the C-terminal Domain
and Biochemical Characterisation of
Nuclear Localisation

Von der Naturwissenschaftlichen Fakultät

der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades einer

Doktorin der Naturwissenschaften

Dr*. rer. nat.*

genehmigte Dissertation

von

Dipl.-Biochem. Melanie Oschlies

geboren am 16.11.1978 in Eckernförde

2008

Referentin: Prof. Dr. Rita Gerardy-Schahn

Korreferent: Prof. Dr. Ralf Ficner

Tag der Promotion: Montag, 30.06.2008

Schlagworte: CMP-Sialinsäure-Synthetase, Sialinsäure, Phosphatase

Key words: CMP-Sialic acid synthetase, sialic acid, phosphatase

Erklärung zur Dissertation

Hierdurch erkläre ich, dass die Dissertation *„Murine CMP-Sialic acid Synthetase: Structural Analysis of the C-terminal Domain and Biochemical Characterisation of Nuclear Localisation*" selbstständig verfasst und alle benutzten Hilfsmittel sowie evtl. zur Hilfeleistung herangezogene Institutionen vollständig angegeben wurden.

Die Dissertation wurde nicht schon als Diplom- oder ähnliche Prüfungsarbeit verwendet.

Table of Contents

Zusammenfassung

Sialinsäuren (Sia) bilden terminale Zuckerstrukturen auf Glykokonjugaten, die an der Zelloberfläche eukaryontischer Zellen präsentiert, oder in die extrazelluläre Matrix transportiert werden. Sialinsäuren spielen eine entscheidende Rolle in Entwicklung und Funktion höherer Tiere, zumal Eingriffe in die Sia Biosynthese im Mausmodell zu Letalität am Tag 9 der Embryonalentwicklung führen.

Die metabolische Energie, die für die Integration von Sia in Glykokonjugate benötigt wird, wird durch die Aktivierung zu Cytidinmonophosphat-Sia (CMP-Sia) bereitgestellt. Die Bildung dieser Nucleotid-aktivierten Sia wird von der CMP-Sialinsäure Synthetase (CSS) katalysiert. Die CSS wurde bereits aus einer Vielzahl pro- und eukaryontischer Organismen kloniert und charakterisiert. Bemerkenswerter Weise sind alle bekannten CSS vom Bakterium bis zum Menschen hoch konserviert. Die CSS variieren jedoch bezüglich ihrer Gesamtgröße und können dementsprechend in zwei Gruppen eingeteilt werden. Während Drosophila und viele Bakterien eine CSS-Kurzform exprimieren, die nur aus der für die Zucker-Aktivierung essentiellen Domäne besteht, weisen alle Vertebraten und einige wenige Bakterien eine CSS auf, die noch eine zusätzliche C-terminale Domäne besitzt. Interessanterweise besteht keine Homologie zwischen den C-terminalen Domänen der pro- und eukaryontischen CSS auf Primärsequenzebene. Ein weiteres Charakteristikum der vertebraten CSS ist die ungewöhnliche Lokalisation im Zellkern. Die biologische Relevanz der Kernständigkeit ist noch nicht geklärt.

Während die konservierten N-terminalen Domänen der CSS (CSS-NT) umfassend charakterisiert sind, so sind sowohl Struktur als auch Funktion der C-terminalen Domänen eukaryontischer CSS bisher nicht bekannt. In Folge dessen konzentrierte sich meine Doktorarbeit vor allem auf die strukturelle und funktionelle Analyse der C-terminalen Domäne (CSS-CT) der murinen CSS. Dazu wurde CSS-CT separat als stabil gefaltetes Protein hergestellt und die Kristallstruktur bis zu einer Auflösung von 1.9 Å gelöst. CSS-CT des Mausenzyms bildet in der Kristallstruktur ein Tetramer in Form eines vierblättrigen Kleeblatts und weist nahe Verwandtschaft zu Phosphatasen der *Haloacid dehalogenase* (HAD) Superfamilie auf. Hierbei muss hervorgehoben werden, dass das aktive Zentrum in seiner Aminosäure Zusammensetzung an wenigen, aber für die Aktivität kritischen Positionen von dieser Familie abweicht. Zur Charakterisierung des murinen CSS-CT wurden neben einer detaillierten Betrachtung der Kristallstruktur, *in vitro* Aktivitätsassays und gerichtete Mutagenesen zur Funktionsaufklärung durchgeführt, wobei allerdings keine enzymatische Funktionalität gefunden werden konnte. Mittels

Größenausschlusschromatographie konnte jedoch gezeigt werden, dass der CSS-CT für die Tetramerisierung der CSS verantwortlich ist. Darüber hinaus wurde das erste 3D-Modell eines vollständigen CSS Proteins, bestehend aus N- und C-terminaler Domäne, entwickelt.

Der zweite Teil dieser Arbeit fokussiert sich auf die Analyse der intrazellulären Verteilung der murinen CSS in verschiedenen Zelllinien und murinen Geweben. Wie seit langem bekannt, handelt es sich bei der CSS aus Vertebraten um ein kernständiges Protein. Dies steht im Gegensatz zu allen anderen Zucker-aktivierenden Enzymen, die ausschließlich im Zytoplasma lokalisiert sind. Interessanterweise ist dieser ungewöhnliche Aufenthaltsort zumindest auf zellulärer Ebene keine Voraussetzung für die bekannte Sia-aktivierende enzymatische Aktivität der CSS. Um die präzise subzelluläre Verteilung endogener CSS untersuchen zu können, wurden in dieser Arbeit mono- und polyklonale Antikörper, welche spezifisch für unterschiedliche Domänen der murinen CSS sind, generiert und in Bezug auf Speziesspezifität und Anwendbarkeit untersucht. Die Kernständigkeit der murinen CSS konnte in verschiedenen Zelllinien und Geweben bestätigt werden.

Abstract

Sialic acids (Sia) are terminal monosaccharides of glycan chains on glycoconjugates presented at the cell surface or transported to the extracellular space. Sialic acids play a pivotal role for the development and function of higher animals. Interference with Sia biosynthesis causes lethality at embryonic day 9 in mice. The metabolic energy required for integration of sialic acids into glycoconjugates is provided by cytidine monophosphate-activated Sia (CMP-Sia). Formation of this sugar nucleotide is catalysed by the CMP-sialic acid synthetase (CSS), which has been biochemically characterised and cloned from a number of different pro- and eukaryotic species. Remarkably, these enzymes were found to exhibit conservation from bacteria to men. However, while most bacterial CSS are composed of only one conserved domain essential for Sia activation, some bacterial and all vertebrate enzymes show an additional domain located at the C-terminus. Interestingly, no conservation at primary sequence level is found between the C-terminal domains of pro- and eukaryotic CSS. A second remarkable feature of all vertebrate CSS is the fact that these enzymes are nuclear residents.

While the conserved N-terminal part of CSS (CSS-NT) is well characterised in terms of structure and function, no information is available on the C-terminal part of the eukaryotic proteins. My PhD study, therefore, was concentrated at deducing structural and functional information for the C-terminal domain (CSS-CT) of murine CSS. To achieve this aim the C-terminal part was produced as isolated and stably folded recombinant protein and the crystal structure was solved to 1.9 Å resolution. The CSS-CT forms a tetramer with a shamrock-like appearance closely related to phosphatases that belong to the *haloacid dehalogenase* (HAD) superfamily. However, the composition of the presumptive active site comprises slight, but in terms of activity potentially crucial modifications. Using *in vitro* enzymatic activity assays wild type and site specifically mutated proteins were analysed, but so far no active protein has been obtained. In size exclusion chromatography experiments the CSS-CT was shown to mediate tetramer formation of the full-length protein. Moreover, a well-funded model of the complete murine CSS has been established by combining the separately resolved 3D structures of N- and C-terminal domain.

In a second approach the subcellular distribution of the endogenous murine CSS was studied. It is known for years that vertebrate CSS are particular in terms of subcellular localisation. While all other sugar activating enzymes are localised in the cytoplasm, the major destination of CSS is the nuclear compartment. Interestingly, interference with nuclear localisation did not affect Sia activation in a cell based

system. In order to define the precise localisation of CSS in eukaryotic cells, mono- and polyclonal antibodies directed against different domains of murine CSS have been generated in course of this study. Obtained antibodies were characterised in terms of species specificity and applicability. Using these reagents the nuclear residence of endogenous murine CSS could be confirmed in various tissues and cell lines.

General Introduction

Sialic Acid

Sugars are abundant in nature and play a key role in a broad range of biological processes, including signal transduction and immune responses. Sialic acids are usually found as terminal monosaccharide units forming the non-reducing ends of glycan chains at glycolipids and glycoproteins of the eukaryotic cell surface.

Sialic acid (Sia) is a nine-carbon sugar with a carboxylic acid residue at the C1 position (Fig. 1). Sia are the only sugars in eukaryotic glycoconjugates bearing a net negative charge and thus largely contribute to the negative surface charge of animal cells. Moreover, the negative charge in position 1 of the sugar mediates the pronounced hydrophilic character of the sugar. The presence of sialic acids is essential for regular development of higher vertebrates. Interference with Sia synthesis in a genetic mouse model showed lethality at embryonic (E) day 9 (Schwarzkopf et al., 2002, Lehmann et al., 2006). Naturally occurring Sia show an immense structural diversity. More than 50 Sia family members have been identified so far. All derive from the parent compound 2-keto-3-deoxy-5-acetamido-D-*glycero*-D-*galacto*-nonulosonic acid (*N*-acetylneuraminic acid, Neu5Ac) (Angata and Varki 2002; Varki and Varki 2007), but vary by the type of substituent (acetyl-, sufat-, methyl-, lactyl groups) added to the C4, C5, C7, C8 and C9 positions. The addition of Sia to glycoconjugates always originates from the C2 position and can be $\alpha 2,3$ and $\alpha 2,6$ if the acceptor sugar is a hexose while the glycosidic linkage is $\alpha 2,8$ if Sia is bound to Sia in higher eukaryotes. In lower eukaryotes and in bacteria also $\alpha 2,5$ and $\alpha 2,9$ linkages have been identified between Sia-residues (Reglero et al., 1993). One of the most common modifications is the hydroxylation of the *N*-acetyl group at C5, forming *N*-glycolylneuraminic acid (Neu5Gc) (Lepers et al., 1990). Different from the other sialic acid derivatives, deaminated neuraminic acid (2-keto-3-deoxy-D-*glycero*-D-*galacto*-nonulosonic acid, KDN) is synthesised in a second, independent biosynthetic pathway. Nevertheless, only the hydroxyl group at position C5 separates KDN from Neu5Ac. Furthermore, all linkage types known for Neu5Ac; α2,3-, α2,6-, and α2,8- are also found for KDN (Inoue and Kitajima, 2006).
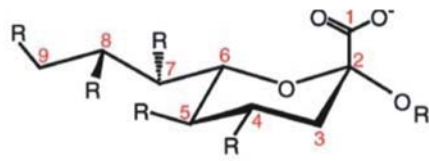
R2 = H in free Sia; alpha linkage to Gal(3/4/6), GalNAc(6),
    GlcNAc(4/6) or Sia (8/9)
R4 = H or *O*-acetyl
R5 = Amino, *N*-acetyl, *N*-glycolyl or Hydroxyl
R7 = H, *O*-acetyl
R8 = H, *O*-acetyl, *O*-methyl, *O*-sulfate or Sia
R9 = OH, *O*-acetyl, *O*-lactyl, *O*-phosphate, *O*-sulfate or Sia

Fig.1 Structural diversity (reprint from Varki, 2007)

All Sias share the common feature of having nine carbons, a carboxylic acid residue at the 1-positon, and a variety of linkages to the underlying sugar chain from the 2-positon. Various types of substitutions at the 4, 5, 7, 8, and 9 positions combine with the linkage variations to generate the diversity of Sias found in nature. Only a portion of this diversity is represented here.

The structural diversity of Sia reflects the variety of biological functions. Sia mediate formation as well as masking of cellular recognition sites and have been implicated in a variety of vital biological processes, requiring cell-cell and cell-extracellular matrix interactions. Biological ageing of circulating cells and serum glycoproteins is indicated by the loss of Sia (Morell et al., 1971; Kelm and Schauer, 1997) and the sialo-glycoconjugate array of cells changes with onco-developmental processes (Kageshita et al., 1995; Miyagi et al., 2004). In addition, Sia are crucial for a functional innate and adaptive immune system in eukaryotes. On the other hand sialic acids are often part of recognition sites to which pathogens attach (Schauer, 2000; Lehmann et al., 2006).

The most abundant Sia species in mammals is *N*-acetylneuraminic acid (Neu5Ac), followed by *N*-glycolylneuraminic acid (Neu5Gc) (Traving and Schauer, 1998; Varki and Varki 2007). A unique feature of Neu5Ac is the formation of long α2,8-linked linear homopolymers, called polysialic acid, particularly found in the developing mammalian brain as posttranslational modification of the neuronal cell adhesion molecule NCAM (Mühlenhoff et al., 1998; Rutishauser, 2008). Due to their negative net charge and hydrophilic character, Sia have a high water binding capacity. This feature is of particular relevance in the case of polysialic acid. Addition of polysialic acid to acceptor molecules like NCAM increases the hydrodynamic radius of these proteins and thus interferes with the tight apposition of neighbouring cells. This provocation of increased inter-cellular distances directly affects cellular functions like

e.g. the migration of progenitor cells in the nervous system (Varki, 2007; Rutishauser, 2008).

Neu5Gc is a common type of Sia in deuterostomes but also present in mammals. The content of Neu5Gc is higher in adults than in juvenile animals. Due to an inactivating mutation in the gene encoding CMP-Neu5Ac hydroxylase, humans in contrast to other primates are not able to generate Neu5Gc (Varki, 2001). Neu5Gc, which is, however, also found in low concentrations in some human tissues, is absorbed from dietary sources particularly from red meat (Tangvoranuntakul et al., 2003). Since degradation of incorporated Neu5Gc by sialidases and acylneuraminidate lyases is slower than of Neu5Ac, decoration with Neu5Gc extends the half live of the corresponding glycoconjugates (Schauer et al., 1995).

KDN, first discovered in rainbow trout eggs (Nadano et al., 1986), is mostly found in bacteria and lower vertebrates, but could also be identified in mammalian tissues, like in hog submaxillary gland, human red blood cells, and the surface of diverse carcinoma cells (review, Inoue and Kitajima, 2006). Megalin, a protein belonging to the low density lipoprotein (LDL) receptor family, has been identified as the sole kidney protein that carries α2,8-linked oligo- or poly-KDN as posttranslational modification (Ziak et al., 1999).

Due to their exposed position on glycoconjugates and their structural diversity, Sia confer important properties to the cell surface and provide the recognition (or part of the recognition) structures for a wide variety of specialised Sia-binding proteins. E.g. the selectins, a family of Sia binding cell adhesion molecules (CAM's), expressed on leukocytes, platelets and endothelium are involved in the regulation of cell-cell interactions during the process of inflammation. Recognition of Sia by selectins thereby mostly requires additional decoration of the glycotope with α1-3(4)-linked fucose residues (Varki, 2007). Siglecs (Sia recognising Ig-like lectins) form a second group of Sia binding molecules (Crocker, 2002). As an example, the B cell restricted Siglec-2 (CD22) is mentioned. CD22 recognises sialic acids in α2-6 linkage. Cis-interaction between CD22 and sialoglycoconjugates on the B-cell surface keeps B-cells in a silent stage. Thus sialogycoconjugates together with the Sia-recognising proteins are important components in the regulation of immune responses (Angata, 2006).

Bacterial pathogens have established mechanisms to circumvent the host immune system by incorporation of sialic acid into their cell-surface structures (Harvey et al., 2001). One prominent example for this form of 'molecular mimicry' is the encapsulated bacterium *Escherichia coli* K1, a Gram-negative neuroinvasive

bacterium that causes meningitis and sepsis in neonates. The bacterium carries a thick capsule of $\alpha2,8$-linked polysialic acid, identical with the polysialic acid bound to NCAM in the nervous system. Using this shell the pathogen is protected against the human primary immune system (Severi et al., 2007). Remarkably, sialic acids are not found in plants (Angata and Varki, 2002; Zeleny et al., 2006).

Sia biosynthesis has first been described in 1962 by Roseman *et al.* for the eukaryotic system by analysis of submaxillary glands and in the same year by Warren and Blacklow for *Neisseria meningitides.* In eukaryotic cells Sia biosynthesis starts with uridine diphospho–*N*-acetylglucosamine (UDP-GlcNAc), deriving from cytosolic glucose (Fig. 2). The first two steps are catalysed by a bifunctional enzyme, the UDP-GlcNAc 2-epimerase/*N*-acetylmannosamine (ManNAc) kinase (GNE/MNK). The GNE domain of the enzyme catalyses the epimerisation of UDP-GlcNAc to ManNAc (EC 5.1.3.14), immediately followed by ATP consuming phosphorylation of ManNAc to ManNAc-6-phosphate (ManNAc-6-P) catalysed by the MNK domain (EC 2.7.1.60). ManNAc-6-P is condensed with phosphoenolpyruvate (PEP) to *N*-acetylneuraminic acid-9-phosphate (Neu5Ac-9-P) catalysed by the Neu5Ac9P-synthase (EC 2.5.1.57). Dephosphorylation to Neu5Ac is performed by Neu5Ac9P-phosphatase (EC 3.1.3.29). The following activation of Neu5Ac, illustrated in figure 2, is an essential requirement for biosynthesis of Sia containing compounds in all biological systems. Whereas the above described Sia formation takes place in the cytosol, activation to cytidine monophosphate–sialic acid (CMP–sialic acid) takes place in the cell nucleus and is catalysed by the CMP-Sialic acid synthetase (CSS, EC 2.7.7.43). CMP activated sialic acids can subsequently be utilised in the trans Golgi network, where the sialyltransferases are assembled. Import into the Golgi apparatus is mediated by a specific CMP-Sia transporter (Eckhardt et al., 1998). Cytosolic CMP–sialic acid is a strong feedback inhibitor of GNE and binds to an allosteric site in the enzyme (Kornfeld et al., 1964). Thus, CMP-Sia contributes to the tight regulation of sialic acid biosynthesis.

Prokaryotic Neu5Ac biosynthesis differs in a few steps from the eukaryotic pathway. Here, *N*-acetylglucosamine-6 phosphate (GlcNAc-6P) serves as a substrate for GlcNAc6P-2 epimerase to form ManNAc6-phosphate (ManNAc6P). Catalysed by Neu5Ac-synthase, ManNAc6P reacts with PEP to form free Neu5Ac.

In a parallel pathway, KDN is synthesised by condensation of mannose-6 phosphate (Man6P) and PEP to KDN-9-phosphate (KDN9P), catalysed by a KDN9P synthase, though the entire pathway is not elucidated in all details yet (review, Inoue and Kitajima, 2006). However, after dephosphorylation, KDN is activated and transferred

to glycoconjugates by the same enzymes responsible for Neu5Ac activation and transfer.

Diversification of Sia occurs before as well as after Sia activation. For example *N*-Glycolylneuraminic acid (Neu5Gc) derives from CMP-Neu5Ac by the addition of an oxygen atom to the *N*-acetyl group, catalysed by cytosolic CMP-Neu5Ac hydroxylase (Schlenzka et al., 1994).
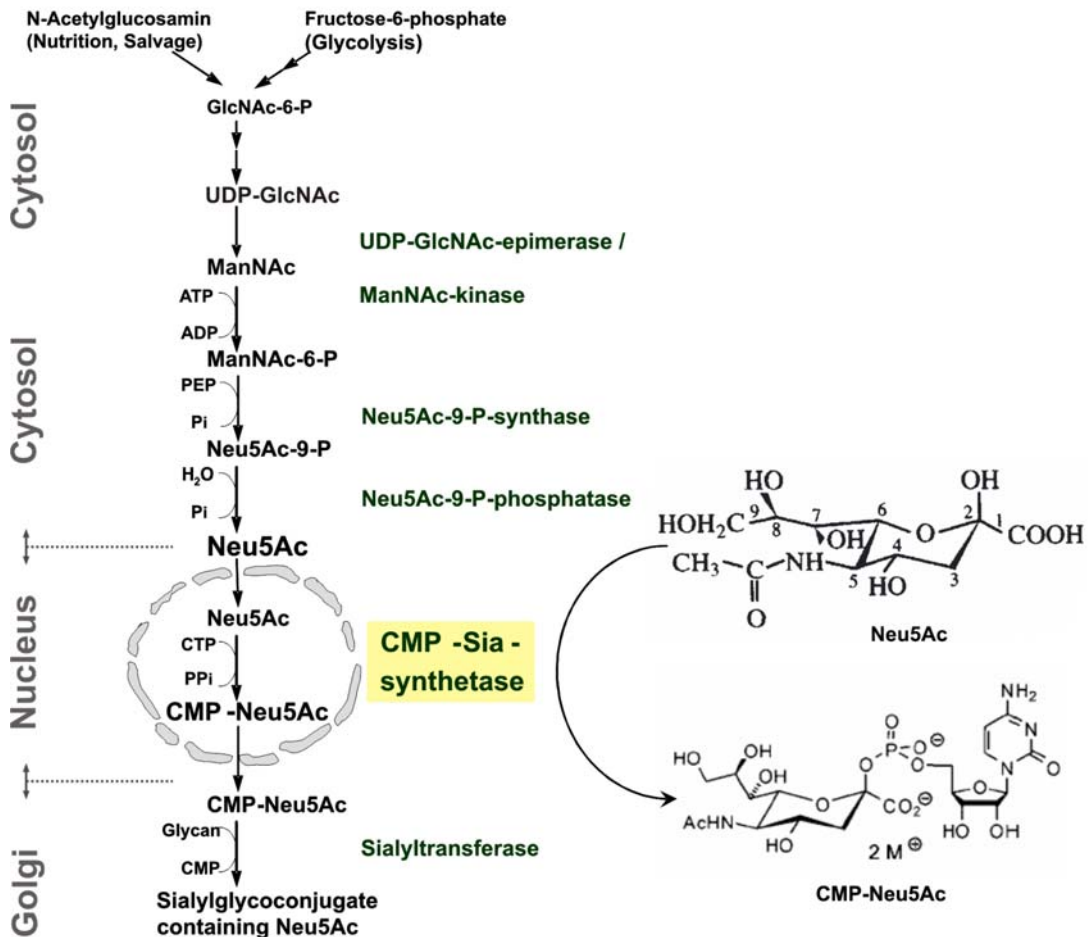
Fig. 2 Biosynthesis of sialic acid (modified from Gagiannis et al., 2007)

In many respects the activation process of sialic acid has a special character compared with that of other sugars. First, Sia is activated by a nucleotide-mono phosphate, whereas all other sugars in eukaryotes are modified by nucleotide-diphosphates. Second, free Sia serves as substrate for CSS, whereas other sugar activating enzymes require sugar 1-P donor substrates. Third, the activation itself takes place in the cell nucleus. This compartment is very uncommon for sugar activation, which usually takes place in the cytosol (Kean et al., 2004).

CMP-Sialic Acid Synthetase

In both, bacteria and vertebrates, activation of sialic acid to CMP-sialic acid is catalysed by the CMP-sialic acid synthetase (CSS, EC 2.7.7.43) using CTP and free sialic acid as substrates.

In 1962 the first isolation of eukaryotic CSS from bovine submaxillary gland was described by Roseman. In the same year Warren and Blacklow isolated the first bacterial CSS from *Neisseria meningitidis*. By now CSS encoding cDNAs have been cloned from several bacterial and animal sources: *Escherichia coli* K1 (Zapata et al., 1989), *Neisseria meningitdis* (Edwards and Frosch, 1992), *Streptococcus agalactiae* (Haft and Wessels, 1994), *Haemophilus ducreyi* (Tullius et al., 1996), *Haemophilus influenzae* (Ishige et al., 2001), *Mus musculus* (Münster et al., 1998), *Oncorhynchus mykiss* (Nakata et al., 2001), *Homo sapiens* (Lawrence et al., 2001), *Drosophila melanogaster* (Viswanathan et al., 2006) and *Clostridium thermocellum* (Mizanur and Pohl, 2007). An amino acid sequence alignment of the cloned CMP-sialic acid synthetases is shown in figure 3. The vertebrate enzymes are depicted in the upper part, the bacterial CSS in the lower part, *Drosophila melanogaster* CSS in the last lane. Five highly conserved amino acid motifs (boxed) illustrate the remarkable similarity between pro- and eukaryotic CSS, although most of the bacterial orthologs appear to be shorter and show similarity only to the N-terminal half (approximately 260 residues) of the vertebrate enzymes (review Münster et al., 2004). However, beside vertebrate CSS, two of the six bacterial enzymes, namely CSS from *Escherichia coli* K1 and *Streptococcus agalactiae* serogroup V, contain an additional C-terminal domain. If present, CSS C-terminal domains (CSS-CT) are composed of about 170 amino acids and are conserved within bacterial and vertebrate CSS, respectively, but amongst each other no similarity can be found. For example, murine and fish CSS share about 57% overall sequence identity, whereas CSS from *Streptococcus agalactiae* shows only 29% identity to the murine enzyme, because homology is restricted to the N-terminal domain. CSS from *Drosophila melanogaster* belongs to the short length CSS and shares 37% and 30% sequence identity with mouse and *E. coli* CSS, respectively. In all enzymes, the N-terminal domains, comprising the five conserved motifs, have been demonstrated to be sufficient for sialic acid activation (Krapp et al., 2003; Liu et al., 2004; Yu et al., 2006).
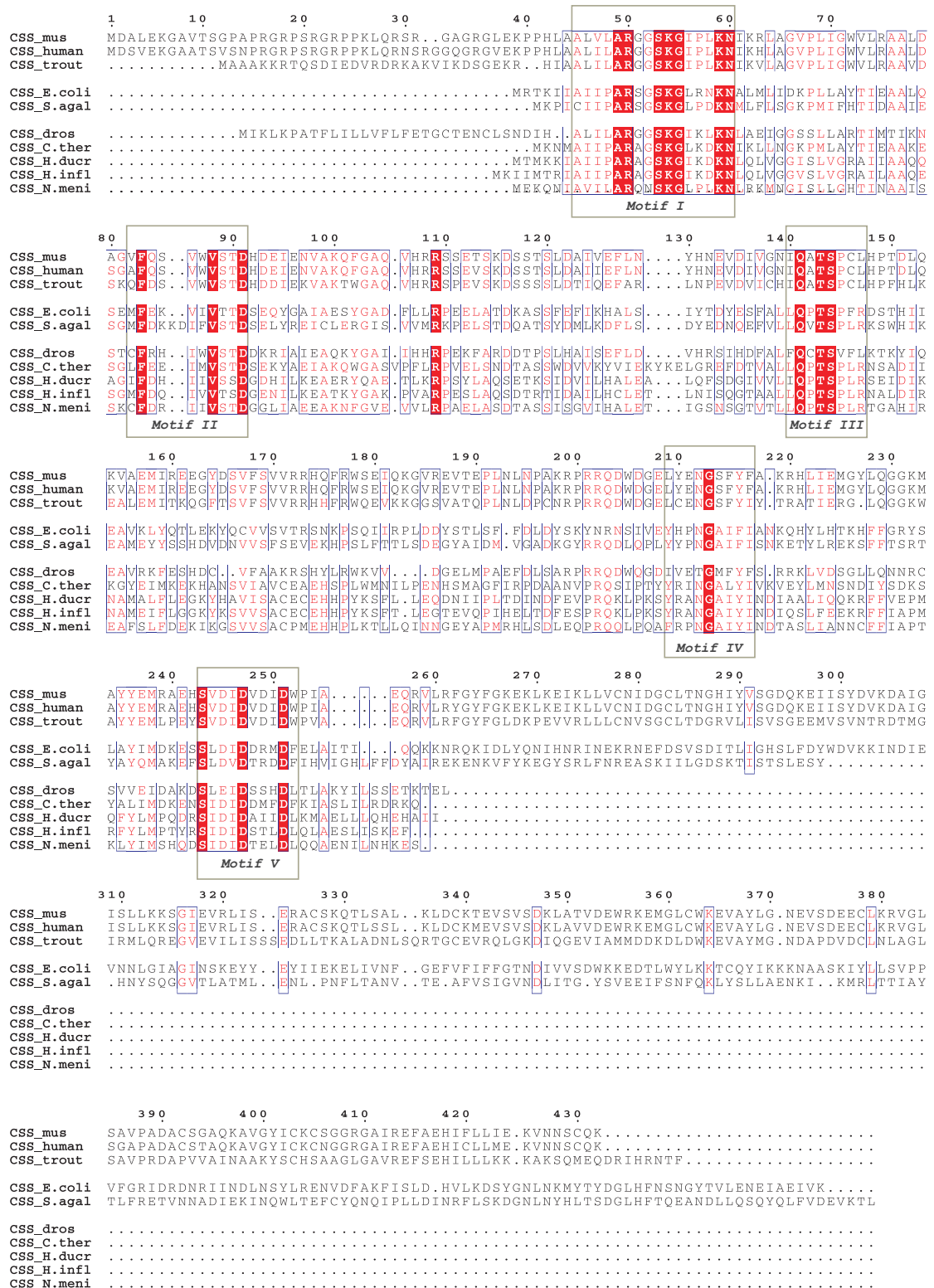
```
          1        10        20        30        40          50        60          70
CSS_mus     MDALEKGAVTSGPAPRGRPSRGRPPKLQRSR..GAGRGLEKPPHLAALVLARGGSKGIPLKNIKRLAGVPLIGWVLRAALD
CSS_human   MDSVEKGAATSVSNPRGRPSRGRPPKLQRNSRGGQGRGVEKPPHLAALILARGGSKGIPLKNIKHLAGVPLIGWVLRAALD
CSS_trout   ..........MAAAKKRTQSDIEDVRDRKAKVIKDSGEKR..HIAALILARGGSKGIPLKNIKVLAGVPLIGWVLRAAVD

CSS_E.coli  ...........................................MRTKIIAIIPARSGSKGLRNKNALMLIDKPLLAYTIEAALQ
CSS_S.agal  ...........................................MKPICIIPARSGSKGLPDKNMLFLSGKPMIFHTIDAAIE

CSS_dros    ..............MIKLKPATFLILLVFLFETGCTENCLSNDIH..ALILARGGSKGIKLKNLAEIGGSSLLARTIMTIKN
CSS_C.ther  ...........................................MKNMAIIPARAGSKGLKDKNIKLLNGKPMLAYTIEAAKE
CSS_H.ducr  ...........................................MTMKKIAIIPARAGSKGIKDKNQLVGGISLVGRAIIAAQQ
CSS_H.infl  ...........................................MKIIMTRIAIIPARAGSKGIKDKNQLVGGVSLVGRAILAAQE
CSS_N.meni  .......................................MEKQNIAVILARQNSKGLPLKNLRKMNGISLLGHTINAAIS
                                                          Motif I

          80         90       100       110       120          130        140        150
CSS_mus     AGVFQS..VWVSTDHDEIENVAKQFGAQ.VHRRSSETSKDSSTSLDAIVEFLN....YHNEVDIVGNIQATSPCLHPTDLQ
CSS_human   SGAFQS..VWVSTDHDEIENVAKQFGAQ.VHRRSSEVSKDSSTSLDAIIEFLN....YHNEVDIVGNIQATSPCLHPTDLQ
CSS_trout   SKQFDS..VWVSTDHDDIEKVAKTWGAQ.VHRRSPEVSKDSSSSLDTIQEFAR....LNPEVDVICHIQATSPCLHPFHLK

CSS_E.coli  SEMFEK..VIVTTDSEQYGAIAESYGAD.FLLRPEELATDKASSFEFIKHALS....IYTDYESFALLQPTSPFRDSTHII
CSS_S.agal  SGMFDKKDIFVSTDSELYREICLERGIS.VVMRKPELSTDQATSYDMLKDFLS....DYEDNQEFVLLQPTSPLRKSWHIK

CSS_dros    STCFRH..IWVSTDDKRIAIEAQKYGAI.IHHRPEKFARDDTPSLHAISEFLD....VHRSIHDFALFQCTSVFLKTKYIQ
CSS_C.ther  SGLFEE..IMVSTDSEKYAEIAKQWGAS.VPFLRPVELSNDTASSWDVVKYVIEKYKELGREFDTVALLQPTSPLRNSADII
CSS_H.ducr  AGIFDH..IIVSSDGDHILKEAERYGAE.TLKRPSYLAQSETKSIDVILHALEA...LQFSDGIVVLIQPTSPLRSEIDIK
CSS_H.infl  SGMFDQ..IVVTSDGENILKEATKYGAK.PVARPESLAQSDTRTIDAILHCLET...LNISQGTAALLQPTSPLRNALDIR
CSS_N.meni  SKCFDR..IIVSTDGGLIAEEAKNFGVE.VVLRPAELASDTASSISGVIHALET...IGSNSGTVTLLQPTSPLRTGAHIR
                  Motif II                                                Motif III

          160        170        180        190        200          210        220        230
CSS_mus     KVAEMIREEGYDSVFSVVRRHQFRWSEIQKGVREVTEPLNLNPAKRPRRQDWDGELYENGSFYFA.KRHLIEMGYLQGGKM
CSS_human   KVAEMIREEGYDSVFSVVRRHQFRWSEIQKGVREVTEPLNLNPAKRPRRQDWDGELYENGSFYFA.KRHLIEMGYLQGGKM
CSS_trout   EALEMITKQGFTSVFSVVRRHHFRWQEVKKGGSVATQPLNLDPCNRPRRQDWDGELCENGSFYIY.TRATIERG.LQGGKW

CSS_E.coli  EAVKLYQTLEKYQCVVSVTRSNKPSQIIRPLDDYSTLSF.FDLDYSKYNRNSIVEYHPNGAIFIANKQHYLHTKHFFGRYS
CSS_S.agal  EAMEYYSSHDVDNVVSFSEVEKHPSLFTTLSDEGYAIDM.VGADKGYRRQDLQPLYYPNGAIFISNKETYLREKSFFTSRT

CSS_dros    EAVRKFESHDCI.|.VFAAKRSHYLRWKVV|...|.DGELMPAEFDILSARPRRQDWQGDIVETGMFYFS.RRKLVDSGLLQNNRC
CSS_C.ther  KGYEIMKEKHANSVIAVCEAEHSPLWMNILPENHSMAGFIRPDAANVPRQSIPTYYRINGALYIVKVEYLMNSNDIYSDKS
CSS_H.ducr  NAMALFLEGKYHAVISACECEHHPYKSFL.LEQDNIIPLTDINDFEVPRQKLPKSYRANGAIYINDIAALIQQKRFFVEPM
CSS_H.infl  NAMEIFLGGKYKSVVSACECEHHPYKSFT.LEGTEVQPIHELTDFESPRQKLPKSYRANGAIYINDIQSLFEEKRFFIAPM
CSS_N.meni  EAFSLFDEKIKGSVVSACPMEHHPLKTLLQINNGEYAPMRHLSDLEQPRQQLPQAFRPNGAIYINDTASLIANNCFFIAPT
                                                                    Motif IV

          240        250        260        270        280        290        300
CSS_mus     AYYEMRAEHSVDIDVDIDWPIA...|...EQRVLRFGYFGKEKLKEIKLLVCNIDGCLTNGHIYVSGDQKEIISYDVKDAIG
CSS_human   AYYEMRAEHSVDIDVDIDWPIA...|...EQRVLRYGYFGKEKLKEIKLLVCNIDGCLTNGHIYVSGDQKEIISYDVKDAIG
CSS_trout   AYYEMLPEYSVDIDVDIDWPVA...|...EQRVLRFGYFGLDKPEVVRLLLCNVSGCLTDGRVLISVSGEEMVSVNTRDTMG

CSS_E.coli  LAYIMDKESSLDIDDRMDFELAITI.|..QQKKNRQKIDLYQNIHNRINEKRNEFDSVSDITLIGHSLFDYWDVKKINDIE
CSS_S.agal  YAYQMAKFFSLDVDTRDDFIHVIGHLFFDYAIREKENKVFYKEGYSRLFNREASKIILGDSKTISTSLESY..........

CSS_dros    SVVEIDAKDSLEIDSSHDLTLAKYILSSETKTEL...............................................
CSS_C.ther  YALIMDKENSIDIDDMFDFKIASLILRDRKQ.................................................
CSS_H.ducr  QFYLMPQDRSIDIDAIIDLKMAELLLQHEHAII...............................................
CSS_H.infl  RFYLMPTYRSIDIDSTLDLQLAESLISKEF..................................................
CSS_N.meni  KLYIMSHQDSIDIDTELDLQQAENILNHKES.................................................
                  Motif V

          310        320        330        340        350        360        370        380
CSS_mus     ISLLKKSGIEVRLIS..ERACSKQTLSAL..KLDCKTEVSVSDKLATVDEWRKEMGLCWKEVAYLG.NEVSDEECLKRVGL
CSS_human   ISLLKKSGIEVRLIS..ERACSKQTLSSL..KLDCKMEVSVSDKLAVVDEWRKEMGLCWKEVAYLG.NEVSDEECLKRVGL
CSS_trout   IRMLQREGVEVILISSSEDLLTKALADNLSQRTGCEVRQLGKDIQGEVIAMMDDKDLDWKEVAYMG.NDAPDVDCLNLAGL

CSS_E.coli  VNNLGIAGINSKEYY..EYIIEKELIVNF..GEFVFIFFGTNDIVVSDWKKEDTLWYLKKTCQYIKKKNAASKIYLLSVPP
CSS_S.agal  .HNYSQGGVTLATML..ENL.PNFLTANV..TE.AFVSIGVNDLITG.YSVEEIFSNFQKLYSLLAENKI..KMRLTTIAY

CSS_dros    ................................................................................
CSS_C.ther  ................................................................................
CSS_H.ducr  ................................................................................
CSS_H.infl  ................................................................................
CSS_N.meni  ................................................................................

          390        400        410        420        430
CSS_mus     SAVPADACSGAQKAVGYICKCSGGRGAIREFAEHIFLLIE.KVNNSCQK............................
CSS_human   SGAPADACSTAQKAVGYICKCNGGRGAIREFAEHICLLME.KVNNSCQK............................
CSS_trout   SAVPRDAPVVAINAAKYSCHSAAGLGAVREFSEHILLLKK.KAKSQMEQDRIHRNTF....................

CSS_E.coli  VFGRIDRDNRIINDLNSYLRENVDFAKFISLD.HVLKDSYGNLNKMYTYDGLHFNSNGYTVLENEIAEIVK.....
CSS_S.agal  TLFRETVNNADIEKINQWLTEFCYQNQIPLLDINRFLSKDGNLNYHLTSDGLHFTQEANDLLQSQYQLFVDEVKTL

CSS_dros    ................................................................................
CSS_C.ther  ................................................................................
CSS_H.ducr  ................................................................................
CSS_H.infl  ................................................................................
CSS_N.meni  ................................................................................
```

Fig. 3 Multiple sequence alignment of cloned CSS.

Alignment was constructed with Multalign version 5.4.1 (Corpet, 1998) using sequences with following accession numbers: SwissProt sequence codes: Q99KK2 (mus_CSS), Q8NFW8 (human_CSS), Q90WG6 (trout_CSS), P13266 (*E.coli*_CSS), Q9AFG9 (*S.agal*_CSS), A3DIQ9 (*C.ther*_CSS), O33868 (*H.ducr*_CSS), Q57140 (*H.infl*_CSS), P0A0Z7 (*N.meni*_CSS), Q8IQV0 (dros_CSS).

Fully conserved residues are shown as white characters with red background; similar residues are depicted as red characters. The five conserved sequence motifs are boxed.

Remarkably, prokaryotic CMP-KDO synthetases (EC 2.7.7.38) show high sequence similarities to CSS (39–58%) and share three of the five conserved motifs. These enzymes catalyse activation of a negatively charged 8-carbon sugar with physicochemical properties similar to sialic acid, named 3-deoxy-D-manno-octulosonic acid (KDO). KDO is found in Gram-negative bacteria as component of LPS. The sugar backbone functions as carrier of the hydrophobic Lipid A-anchor and the polysaccharide portion of LPS. KDO in contrast to Sia and KDN was found in plants and some green algae as a component of cell wall polysaccharides (Royo et al., 2000). Beside Sia, KDO is the only sugar which is activated by CMP.

Bifunctional CMP-Sialic Acid Synthetases

Both *Streptococcus agalactiae (S. agalactiae)* and *E. coli* K1 carry capsular polysaccharides, mimicking sialylated host structures and therefore limit the activation of an effective innate immune response. *S. agalactiae* capsules are capped by terminal α2,3-linked Neu5Ac- residues, whereas *E. coli* K1 is protected by a thick polysialic acid layer consisting of α2,8-linked Neu5Ac.

CMP-sialic acid synthetases of both strains have a C-terminal domain, composed of about 170 amino acids (CSS-CT, Fig. 3). These CSS-CTs show sequence homology to the GSDL (Gly, Ser, Asp, Leu) family of SGNH (Ser, Gly, Asn, His) hydrolases and are not related to vertebrate CSS C-terminal domains. Recent work has shown that CSS from *S. agalactiae* and *E. coli* K1 are bifunctional enzymes combining the N-terminal CMP-Neu5Ac synthetase with a C-terminal esterase activity. Initially, the esterase activity was demonstrated *in vitro* with the substrates 4-nitrophenyl-acetate and acetylated platelet-activating factor (Liu et al., 2004; Yu et al., 2006). More recently, evidence has been elaborated that O-acetylated Sia is the native substrate by demonstrating deacetylase activity for the recombinant enzymes (Lewis et al., 2007; Steenbergen et al., 2006). *In vivo* studies performed with CSS from *S. agalactiae,* furthermore revealed that overexpression decreases total Sia-O-acetylation while deletion of the CSS-gene increased the production of O-acetylated Sia. Similarly, the disruption of the *S. agalactiae* CSS esterase domain resulted in the formation of capsules in which the cap-Sia carried significantly increased levels of O-acetat (Lewis et al., 2007). Finally, CSS purified from *S. agalactiae* was demonstrated to catalyse in consecutive steps the de-O-acetylation of free Neu5,9Ac$_2$ and the activation of the resulting Neu5Ac to CMP-Neu5Ac as well as the activation of free Neu5,9Ac$_2$ followed by the de-O-acetylation of the activated sugar CMP-Neu5,9Ac$_2$ (Lewis et al., 2007). Thus, CMP activation of Sia is not a prerequisite for the Sia-O-acetyl esterase reaction. Interestingly, the degree of O-

acetylation markedly affects the interaction between *S. agalactiae* and host Siglecs (Carlin et al., 2007). Modification of the *E. coli* K1 capsule by O-acetylation in positions 7 and 9, thereby, correlated with increased virulence in patients suffering from bacteremia (Frasa et al., 1993; Deszo et al., 2005). Taken together these data illustrate that the bacterial CSS-CTs are involved in modulating the virulence and immunogenicity of the capsular polysaccharides by regulating the level of acetylation of sialic acid components in these structures.

Different to bacteria, the biological function of the C-terminal domains in vertebrate CSS are unknown. Homology has been predicted to phosphatases of the *haloacid dehalogenase* (HAD) superfamily (Parsons et al., 2002), however, enzymatic proof has not been provided yet.

Intracellular Localisation

Extraction of CSS activity from different animal sources revealed that the vertebrate enzyme mainly resides in the cell nucleus. Thus nuclear extracts from calf kidney (van Dijk et al., 1973), rat brain (Gielen et al., 1971) and hog retina (Kean, 1969) contained 77-85%, 64% and 75% of the obtained CSS activity, respectively. In cytoplasmic extracts only 0.4 to 11.3 % of the total activity was observed in tissue specimens isolated from different species (see review Kean et al., 2004). Nuclear localisation was further confirmed by fluorescent labelling of recombinant mouse, human and trout CSS, expressed in NIH 3T3, Sf9 and rainbow trout gonadal tissue-derived RTG-2 cells, respectively (Münster et al., 1998; Lawrence et al., 2001 and Tiralongo et al., 2007).

Ions, metabolites and other small molecules can enter the nucleus through the nuclear pore complex (NPS) by passive diffusion, while the translocation of cargoes larger than ~40 kDa generally requires specific transport receptors known as karyopherins (Terry et al., 2007). Karyopherins recognise and bind their cargoes via conserved sequence motifs existing for nuclear import as nuclear localisation signals (NLS) and nuclear export as nuclear export signals (NES). NLSs can be separated in so called monopartite and bipartite motifs. Monopartite NLS are generally characterised by one short stretch (4-8 residues) of positively charged amino acids with the consensus sequence K-K/R-X-K/R-R, whereas bipartite NLS motifs contain two positively charged clusters separated by a linker region of 10-12 amino acids ($(K/R)_2$-X_{10-12}-$(K/R)_{3-5}$) (Christophe et al., 2000). A single protein can contain one or more NLS consensus sequences. NESs are short leucine-rich motifs, deduced from the HIV REV and cellular protein kinase inhibitor proteins (Nigg, 1997), with the

consensus sequence $LX_{2-3}LX_{2-3}LXL$, where X stands for any amino acid and L can be some other hydrophobic amino acid (Henderson and Eleftheriou, 2000).

Vertebrate CSS are proteins with a calculated molecular mass of about 48 kDa, indicating the need of active transport. Consistently, several clusters of basic amino acids could be identified for murine and fish CSS and their functionalities were tested with the help of deletion mutants (Münster et al., 1998; Tiralongo et al., 2007). Divergent from the mouse enzyme, where the crucial NLS turned out to be the monopartite sequence $K^{198}RPRR$ (Münster et al., 2002), nuclear localisation of the fish CSS depends on a free N-terminus with the bipartite NLS $K^5KRTQSDIEDVRDRKAK$ accessible (Tiralongo et al., 2007). In Figure 4 position and sequence of the identified NLS sequences in mouse and fish are depicted. NLSs crucial for mediating nuclear import of murine and fish CSS are shown in red, whereas the corresponding basic clusters that do not form active NLSs are shown in black letters.
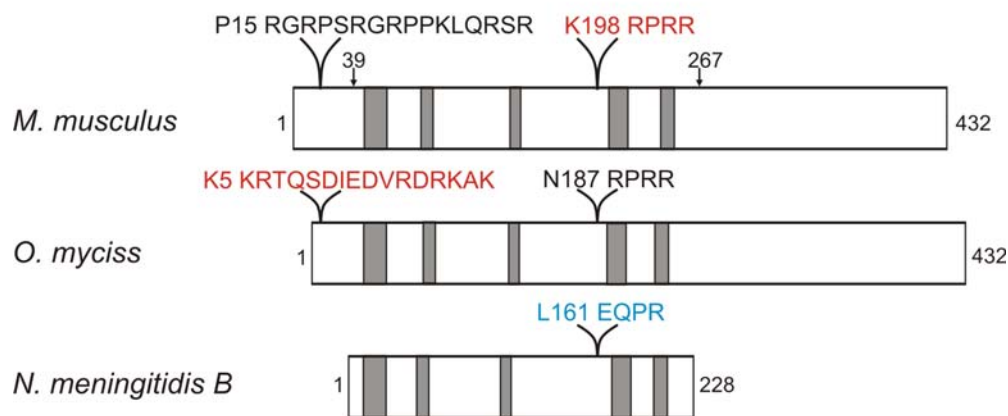
Fig. 4 Schematic representation of CMP-Sialic acid synthetases (CSS) from mouse, rainbow trout and *Neisseria meningitidis* (NmB). The highly conserved sequence motifs, common to all CSS are depicted as grey bars. Position and sequence of basic amino acid clusters for mouse and trout CSS are annotated above; equivalent sequences for NmB CSS are shown in blue. Functional nuclear localisation signals are coloured in red.

Exemplarily, *Neisseria meningitidis* CSS is included in this graph to illustrate that bacterial enzymes do not harbour basic clusters at the corresponding positions. The five highly conserved sequence motifs, important for enzymatic activity, are shown as grey bars. Interestingly, the functional NLS of the murine enzyme comprises an amino acid stretch shown to be part of the enzyme's active site (Münster et al., 2002; Krapp et al., 2003). Mutation of Arg^{202} to Ala in murine CSS reduced total activity; mutation of the corresponding residue Arg^{165} in NmB CSS completely abolished activity (Münster et al., 2002). Interestingly, mutation of the murine NLS sequence to $A^{198}RPAR$ resulted in a cytoplasmic, but active enzyme, proving that at least at the

single cell level the nuclear environment is no prerequisite for enzymatic activity (Münster et al., 2002).

Although murine CSS is predominantly found in the nucleus, a certain amount of enzyme activity can be detected in the cytoplasm (see review Kean et al., 2004) and this is particularly true, if the enzyme is overexpressed in different recipient cells. Fujita et al. (2007) recently explained this phenomenon by the existence of two nuclear export signals (NES) in the murine CSS sequence. The first motif is positioned at the very N-terminus P^{57}LKNIKRLAGVPLIGWV^{73}, the second at the very C-terminus G^{408}RGAIREFAEHIFLLIE^{424}. Deletion of the predicted NESs caused a more pronounced nuclear localisation of recombinant CSS. Because, however, fusion of the identified NES peptides to GFP was not sufficient to completely retain GFP in the cytoplasm the existing NESs seem to be subordinate to the parallel existing NLS-motif.

Different from all other eukaryotic CSS the enzyme isolated from *Drosophila melanogaster* was found to be Golgi targeted when expressed as recombinant protein in both mammalian and insect cells (Viswanathan et al., 2006).

Structural Information

So far, structural information is available of one pro- and one eukaryotic CSS. *Neisseria meningitidis* serogroup B (NmB) CSS was crystallised in the presence of CDP (Mosimann et al., 2001), whereas the 3D-structure of the catalytic active N-terminal domain (residues 39-267) of the murine enzyme was obtained in the product bound state (CMP-Neu5Ac, Krapp et al., 2003). In figure 5 the two functional dimers of mouse (A) and NmB CSS (B) and are shown as ribbon presentation. Monomers are coloured in magenta and yellow. Both enzymes share 27% sequence identity and the structures revealed remarkable similarities in the overall topology and active site architecture. Both enzymes form functional homodimers with each monomer organised in a globluar $\alpha\beta$-fold linked to a dimerisation domain.
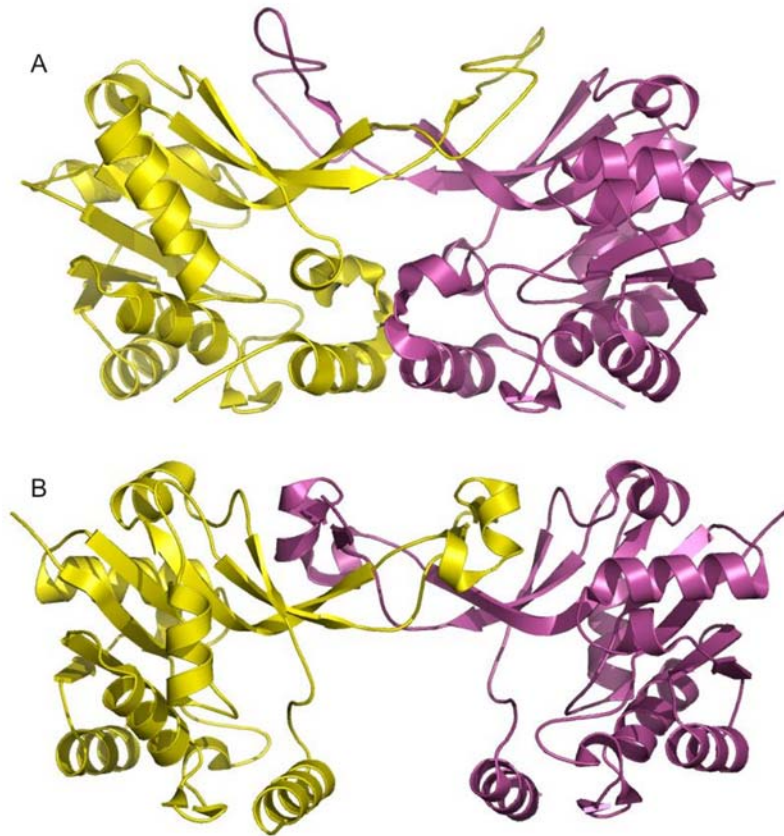
Fig. 5 Comparison of catalytic domains of murine and NmB CMP-Sialic acid synthetases
(a) Ribbon diagram of the murine Sia-activating CSS-NT dimer with the monomers coloured in magenta and yellow (PDB entry 1qwj); (b) Ribbon diagram of the CSS dimer from *Neisseria meningitidis* with the monomers shown in magenta and yellow (PDB entry 1eyr).

The active sites are located at the interfaces of the core domains of one monomer and the dimerisation domain of the other monomer. The arrangement of the dimerisation domains provides one major difference between NmB and murine CSS 3D structures. In case of the murine enzyme, the nuclear localisation signal K[198]RPRR is positioned at this dimerisation domain and is also part of the active site. Arg[202] is involved in substrate binding and mutation to Ala dramatically reduces the catalytic activity of murine CSS (Münster et al., 2002, Krapp et al., 2003). The equivalent Arg of the NmB enzyme (Arg[165]) has been shown to be essential for enzymatic activity (Münster et al., 2002). Interestingly, in the crystal structure this residue is translocated outwards the active site by more than 5 Å, in comparison with the residue of the murine CSS. Hence, the crystal structure of the product bound murine CSS represents the closed conformation, whereas NmB CSS is crystallised in the open conformation and still open for substrate access (Krapp et al., 2003).

Objectives

CMP-sialic acid synthetase (CSS) catalyses the activation of sialic acid (Sia) to CMP-sialic acid, which then serves as a substrate for Golgi resident sialyltransferases, which transfer the sugar into terminal positions of glycoconjugates. Vertebrate CSS are nuclear localised, while other sugar activating enzymes reside in the cytoplasm. Moreover, vertebrate CSS without exception are chimerical proteins containing the CSS-activity in the conserved N-terminal protein part. The C-terminal protein parts (CSS-CT) of these enzymes show also some conservation and based on primary sequence alignments were predicted to be related to bacterial phosphatases. The CSS-CTs are, however, not required for activity and not for nuclear transport of the enzymes. The aim of my study was to deduce information on the functionality of the CSS-CTs. With the murine CSS as a model structure it was my goal to solve the crystal structure of CSS-CT and to use the obtained information to reinstall phosphatase activity by site specific mutagenesis. In a parallel approach the recombinant protein should be used to generate antibodies and to use these reagents to precisely profile the expression patterns of endogenous CSS in different cell lines and murine tissues.

Chapter 1

The C-terminal domain of the murine CMP-sialic acid synthetase exhibits structural homology to phosphatases of the HAD superfamily

1.1. Introduction

Sialic acids (Sia) comprise a family of acidic 9-carbon sugars that show high structural diversity (Angata and Varki, 2002). In mammals Sia are typically found on the terminal position of glycan chains of glycolipids and glycoproteins and are required for normal embryonic development (Schwarzkopf et al., 2002). Due to their exposed position and unique physicochemical properties, Sia fulfil important functions in cell-cell recognition, cell differentiation, memory formation and immune responses (Varki, 2007; Lehmann et al., 2006; Rutishauser, 2008). Furthermore, Sia are often part of recognition sites to which pathogens attach (Angata and Varki, 2002). In bacteria, Sia are involved in formation of capsular lipooligo- and polysaccharides and represent critical pathogenicity factors (Vimr et al., 2004). An essential requirement for the *de novo* biosynthesis of sialylated structures is the activation of Sia to its cytidine monophosphate diester (CMP-Sia) catalysed by the CMP-sialic acid synthetase (CSS, E.C.2.7.7.43; CTP + Sia → CMP-Sia + PPi). Primary sequence analysis of CSS revealed homology from bacteria to man, and 5 highly conserved motifs essential for Sia activation have been identified (Nakata et al., 2001; Münster et al., 2004). X-ray analyses of the sugar activating domain of the mouse enzyme (mCSS-NT, residues 39-267) and of full-length *Neisseria meningitidis* serogroup B CSS (NmB, residues 1-228) revealed a similar overall topology and active-site architecture (Mosiman et al., 2001; Krapp et al., 2003). Both enzymes exist as functional dimers with the two monomers linked together by an intertwined dimerisation domain which is simultaneously part of the active site (Fig 10A for murine CSS-NT). In contrast to the bacterial CSS, the catalytic domain of the mammalian enzyme forms a tetramer in which the functional dimer is dimerised *via* hydrophobic interactions mediated by a C-terminal helix (Krapp et al., 2003).

In principle, CSS can be separated into two groups: (i) Short size enzymes comprising only the conserved catalytic domain. In this domain, the five highly conserved motifs shown to be essential for sialic acid activation (Münster et al., 2004) are assembled. This group contains mainly bacterial enzymes like e.g. CSS from *Neisseria meningitidis serogroup B*. (ii) Long size enzymes; these enzymes

carry a C-terminal domain of about 18 kDa in addition to the conserved N-terminal domain. All currently verified and predicted vertebrate CSS are long forms. In the bacterial kingdom CSS belonging to both groups have been identified (e.g. for long CSS, *E.coli* K1: Zapata et al., 1989; and NmB for short CSS: Edwards and Frosch, 1992)

Interestingly, if existing, the C-terminal domains are dispensable for Sia activation in both pro- and eukaryotic CSS. Moreover, no homology exists in the C-terminal parts between pro- and eukaryotic enzymes (Krapp et al., 2003; Liu et al., 2004; Yu et al., 2006). Recently, *Streptococcus agalactiae* serogroup V (*S. agalactiae* V, Yu et al., 2006) and *Escherichia coli* K1 CSS (*E. coli* K1, Liu et al., 2004) have been described as bifunctional enzymes with acetylhydrolase activity located in the C-terminal domains. According to predicted structural similarity to GDSL esterases of the SGNH hydrolase family (Yu et al., 2006) deacetylation of O-acetylated Sia monomers has been observed *in vitro* for both enzymes. For *S. agalactiae* CSS deacetylation of O-acetylated Sia was also shown by *in vivo* studies. Therefore, in addition to activation of free Sia, *S. agalactiae* and *E. coli* CSS are supposed to be involved in the regulation of the intracellular concentration of O-acetylated monomeric Sia. Consequently, these bifunctional CSS also influence the degree of acetylation of polysialic acid and thereby modulate virulence and immunogenicity of the capsular polysaccharide (Lewis et al., 2006; Steenbergen et al., 2006; Lewis et al., 2007).

In vertebrates, Sia activation represents a key reaction in the sialylation pathway, not only for providing the substrate for sialyltransferases, but also for providing a pivotal feedback inhibitor of the upstream UDP-*N*-acetylglucosamine epimerase / *N*-acetylmannosamine kinase (GNE/MNK) (Keppler et al., 1999; Blume et al., 2004). Vertebrate CSS are unique as far as they have been predominantly found in the cell nucleus. All other sugar activating enzymes are restricted to the cytoplasm (Kean et al., review 1991 and 2004). Nuclear localisation signals responsible for nuclear targeting have been identified and evaluated in both mouse and rainbow trout CSS (Münster et al., 2002; Tiralongo et al., 2007). In contrast, the biological meaning of nuclear residency is not understood. Moreover, several pieces of information promoted the speculation that nuclear residence is no more than accidental. Arguments to support this view have been (i) the observation that nuclear import is not prerequisite for CSS activity (Münster et al., 2002) and (ii) the demonstration that the relevant NLS in the murine enzyme is also part of the active site and exposed at the protein surface (Krapp et al., 2003). However, while highly conserved at primary sequence level, the active NLS-motifs are different in mouse and rainbow trout (Tiralongo et al., 2007), thus providing strong evidence for the relevance of nuclear

targeting. Furthermore, two weak nuclear export signals (NES) were recently identified for the mouse enzyme, eventually suggesting a dynamic relocation of the enzyme (Fujita et al., 2007). Beyond the enigma of nuclear sequestration, the biological function of the C-terminal domain in vertebrate CSS is also unknown. PSI-Blast sequence analysis performed with a bacterial sugar phosphatase belonging to the haloacid dehalogenase (HAD) superfamily revealed significant similarity to the murine CSS-CT (Parsons et al., 2002). Members of the HAD superfamily are structurally typified by L-2-haloacid dehalogenases, corresponding to the first protein of this family which has been crystallised (Hisano et al., 1996; Riddler et al., 1997). The family embraces proteins with different enzymatic activities; the largest subgroup, the phosphatases are followed by ATPases and phosphomutases. All family members are characterised by an alpha/beta core domain classified as a hydrolase fold, three conserved sequence motifs and -as far as analysed- a common mechanism of action (Aravind and Koonin, 1998; Allen and Dunaway-Mariano, 2004). Despite the remarkably low sequence identity of the different HAD superfamily proteins (less than 14%), the spatial arrangement of the three sequence motifs, which define the active site, including substrate as well as cofactor-binding residues, is highly conserved. Diversification of chemistry within the family is achieved by the presence or absence of a cap domain, which reorients upon substrate binding and is able to close the active site (Lahiri et al., 2006).

With the objective to shed light on the biological function of the CSS C-terminal domain in vertebrates, we determined the x-ray crystal structure of the mouse CSS C-terminal domain (CSS-CT, amino acids 267-432) at 1.9 Å resolution. Structure superposition revealed high overall similarity to phosphatases belonging to the HAD superfamily and thus confirmed primary sequence analyses. However, few but crucial differences in the active site architecture were identified. To evaluate a potential phosphatase activity, *in vitro* assays were performed with a wide range of substrates using native as well as engineered CSS-CT, in which the active site of functional HAD enzymes was approximated by site directed mutagenesis. While activity could not be installed under the conditions used *in vitro,* the combination of the structural information obtained for N- and C-terminal part of the murine CSS, enabled for the first time construction of a 3D-model of the full length murine CSS.

1.2. Experimental Procedures

1.2.1. Materials

pET expression vectors and *Escherichia coli* BL21(DE3) and *E. coli* B834(DE3) were purchased from Novagen. Glucose-6P was bought from Serva, phosphoenolpyruvat and pNPP from Fluka BioChemika and ATP from Merck. All other substrates were purchased from Sigma.

1.2.2. Cloning of CSS-CT

CSS-CT cDNA was amplified from the plasmid pBSCMPrev using the primer pair MO 34 (5′-G<u>GGATCC</u>AAAGAGAAGCTGAAGGAGATAAAG -3′) / MO 35 (5′- AA<u>CTGCAG</u>CTATTTTTGGCATGAGTTATTAAC -3). PCR products have been digested with *Bam*HI/*Pst*I (underlined), and, after gel purification (Qiagen), ligated into the corresponding restriction sites of pET43-Strep, a modified pET43 vector containing the sequence encoding an N-terminal Strep-tag II followed by a thrombin cleavage site (WSHPQFEKGALVPRGS). The resulting plasmid pET43-Strep-CMP267 drives the expression of N-terminally NusA-StrepII-tagged proteins. The insert was further subcloned into *Bam*HI/ *Xho*I sites of the expression vector pET22b-Strep, a modified pET22b vector also containing the sequence encoding an N-terminal Strep-tag II followed by a thrombin cleavage site (WSHPQFEKGALVPRGS). The sequence identity of all PCR products was confirmed by sequencing.

1.2.3. Site-directed mutagenesis

Active site mutants of CSS-CT were obtained by QuikChange site-directed mutagenesis (Stratagene) following the manufacturers guidelines. The identity of all constructs was confirmed by sequencing. Mutagenic primers were as follows (mutated base triplets underlined): N279D: MO43, 5′- CTTTTGGTTTGT<u>GAT</u>ATTGATGGATGTCTCACCAATGGCCACATTTATG -3′; MO44, 5′- CATCCATCAAT<u>ATC</u>ACAAACCAAAAGCTTTATCTCCTTCAGCTTC -3′; N371D: MO45, 5′- GGCCTATCTCGGC<u>GAT</u>GAAGTGTCTGATGAAGAATG -3′; MO46, 5′- CAGACACTTC<u>ATC</u>GCCGAGATAGGCCACTTCTTTC -3′. Plasmids containing the wild-type sequence were used as a template. PCR products were subcloned into *Kpn*I and *Xba*I sites of pBSCMPrev, resulting in pBSCMPrevN279D and pBSCMPrevN371D. Double mutants were generated by combination of the two genes using *Bsa*I restriction sites. Mutated gene fragments were subcloned into the

*Hind*III und *Not*I sites of pET43-Strep-CMP267 for expression of N-terminal NusAStrepII-tagged proteins.

1.2.4. Expression and purification of CMP-sialic acid synthetase constructs

Recombinant CSS-CT (residues 267–432) carrying a StrepII–tag (WSHPQFEKGALVPRGS, IBA) at the N-terminus was overexpressed in *E. coli* BL21 (DE3) and cultivated in PowerBroth (Athena Enzyme Systems) at 15 °C. At an optical density (OD_{600}) of 1.2, expression of CSS-CT was induced by adding 0.1 mM isopropyl-1-thio-β-D-galactopyranoside, and bacteria were harvested 24 h after induction. For protein purification, bacteria were lysed by sonication (Branson Sonifier) in appropriate loading buffer (10 mM Tris-HCl, pH 8, 100 mM NaCl) containing protease inhibitors (40 µg/ml bestatin (Sigma), 4 µg/ml pepstatin (Sigma) and 1µg/ml Leupeptin (Boehringer Mannheim)), and the soluble fraction was incubated with 4 µg/ml avidin for 30 min at 4 °C and aggregates were removed by centrifugation at 12 000 x g. The filtered supernatant (0.2 µm) were loaded on a 5 ml Strep-Tactin Superflow (IBA) column. The protein was purified by affinity chromatography utilising the N-terminal StrepII–tag according to manufacturers' instruction. Buffer exchange was performed on a HiPrep 26/10 desalting column (Amersham Biosciences). The N-terminal StrepII–tag was removed by proteolytic cleavage with biotinylated thrombin (Novagen). The StrepII–tag and biotinylated-thrombin were precipitated by a mix of Streptavidin-agarose (Fluka BioChemika) and Strep-Tactin sepharose beads (IBA). Size exclusion chromatography was performed for further purification using Superdex 200 HR 10/30 column (Amersham Biosciences) equilibrated with 10 mM Tris-HCl, pH 8, 100 mM NaCl, 4 mM DTT. The flowthrough was concentrated to 10 mg ml^{-1} in buffer containing 10 mM Tris-HCl, pH 8, and 100 mM NaCl. The selenomethionine (SeMet) derivative of CSS-CT was overexpressed in *E. coli* B834(DE3) grown in M9-minimal medium (Sigma) supplemented with 50 mg l^{-1} D- and L-SeMet (Sigma) at 30 °C and purified using the same protocol as for the native protein.

Full-length CSS and CSS-NT were expressed and purified like native CSS-CT via N-terminal StrepII-tags, which were removed afterwards by thrombin digestion.

1.2.5. Expression of *N*-acetylneuraminic acid 9-phosphate synthase

The vector pET-32a-NPSm, allowing the expression of 6xHis-tagged *N*-acetylneuraminic acid 9-phosphate (Neu5Ac9P) synthase (NPS) was kindly provided by Ken Kitajima from Nagoya University, Nagoya 464-8601, Japan (Nakata et al., 2000). Expression was performed in *E. coli* BL21(DE3) cultivated in PowerBroth (Athena Enzyme Systems) at 15 °C, as described above. For protein purification, bacteria were lysed by sonication in appropriate loading buffer (50 mM Tris-HCl, pH 8, 300 mM NaCl) containing protease inhibitors cocktail (Sigma). Soluble fractions were loaded on 2-ml HisTrap HP column (Amersham Biosciences) according to the manufacturer's guidelines. Bound protein was eluted with 150 mM imidazole in buffer A (50 mM Tris-HCl, pH 8, 300 mM NaCl, 10 % Glycerol) using a step gradient. Buffer exchange was performed twice by dialysis against the 200 fold volume of buffer A containing 2 mM DTT (buffer B). Purified protein samples were stored at -80 °C in buffer B.

1.2.6. Analysis of phosphatase activity

1H-NMR based Enzyme Assays

All NMR experiments were performed at the Institute for Glycomics at Griffith University, Gold Coast, Australia (from September to November 2005) on a Bruker Avance DRX 600 MHz spectrometer at 298 K in deuterated 100 mM $Tris_{d5}$, 150 mM NaCl, 20mM $MgCl_2$, pH 8 in a total volume of 600 µl. Data acquisition and processing were performed with XWINNMR software (Bruker) on a Silicon Graphics O2 work station.

Coupled assay for Neu5Ac9P generation and dephosphorylation

Dephosphorylation of Neu5Ac9P was analysed with the following coupled assay: ManNAc6P+ PEP → Neu5Ac9P; Neu5Ac9P→ Neu5Ac + Pi; Neu5Ac + CTP → CMP-Neu5Ac. *N*-Acetylmannosamine5-phosphate (ManNAc6P) was synthesized as described by Liu and Lee (2001). In brief, mannosamine was enzymatically phosphorylated at the C6 position to generate mannosamine-6P, followed by chemical acetylation and purification of the desired product ManNAc6P. Product formation was controlled continuously by thin-layer chromatography (TLC) and finally by MS analysis and 1H-NMR spectroscopy. A reaction mixture of ManNAc6P (10 mM), 2 mM PEP and 42 µg purified NPS was incubated for 1 h at 37°C for generation of Neu5Ac9P. After addition of CSS-CT, or calf intestine alkaline

phosphatase (CIP, NEB) as positive control, and further incubation at 37°C, CTP (f. c. 5mM) and purified CSS were added. Successful Neu5Ac9P dephosphorylation lead to the formation of CMP-Neu5Ac that could be monitored in 1H-NMR by appearance of characteristic peaks at 2.48 ppm for H3 eq. of CMP-Neu5Ac.

EnzChek

In vitro phosphatase activity assay of CSS and CSS-CT was performed using EnzChek[TM] phosphatase assay kit (Molecular Probes, Eugene, OR) following manufacturers instructions. Each 90 µl reaction mixture contained 50 mM Tris pH 7.5, 8 mM $MgCl_2$ and 1-10 mM of the various phosphorylated substrates (Table 3). After baseline development reactions were started by the addition of 2 µg enzyme (10 µl, f.c. 1-2 µM) or CIP as a positive control.

para-nitrophenyl phosphate assay

Dephosphorylation of para-nitrophenyl phosphate (pNPP) was analysed in assay solution (50 mM PBS, pH 8, 1 mM $MgCl_2$, 300 mM NaCl and 100 mM pNPP) at 37°C for 30 min. Reactions were carried out in 96-well microtiter plates in a total volume of 100 µl and quenched by the addition of 30 µl 1 M NaOH. PNP anion was detected at 405 nm (em=1.85×10^4). In addition the following buffer systems were tested: NaAc pH5, MES pH6, Hepes pH 7, Hepes pH 8, Glycine pH9.

1.2.7. Crystallisation and data collection

All Crystallisation experiments were carried out in Cryschem plates (Hampton) using the sitting drop vapor-diffusion technique. Initial Crystallisation conditions were obtained using the Crystal screen kit from Hampton-Research. CSS-CT crystals were grown at 277 K by mixing equal volumes of protein solution (10 mg ml^{-1}) and precipitant (10% (w/v) EtOH, 0.1 M Tris-HCl, pH 8, 4 mM Neu5Ac, 4 mM DTT). The crystals appeared after two days and reached a maximum size of 200 × 200 × 1,000 μm^3 within two weeks. Crystals were cryoprotected in precipitant solution containing 25% (v/v) 2-Methyl-2,4-pentanediol (Hampton) and flash-cooled in nylon-fiber loops at 100 K for data collection. The diffraction experiments were carried out at beamline BL14.1 and BL14.2 of BESSY and Free University Berlin at BESSY. X-ray data of a native crystal were collected at 100 K to a resolution of 1.89 Å, a three-wavelength MAD data set of a SeMet derivative was collected to 1.9 Å. The diffraction images of the native dataset were scaled and processed using DENZO and SCALEPACK (HKL Research) the MAD dataset was indexed, integrated and scaled with X-ray detector software (XDS, Kabsch, 1993). Crystals belonged to space group C2 containing two molecules in the asymmetric unit. Data processing statistics are summarised in Table 2.

1.2.8. Structure determination and refinement

Molecular replacement was performed using 3D coordinated from *Haemophilus influenzae* KDO8P phosphatase (PDB entry: 1J8D) and the CaspR web-server (Claude et al., 2004). The coordinates of the selenium sites were found using SHELX-D and initial phases were calculated in SHELXE (Sheldrick et al., 2001). Density modification by solvent flattening, histogram matching and phase extension was done in DM (Cowtan and Zhang, 1999). The resulting map was further improved with the program RESOLVE (Terwilliger, 2000 and 2003) whose chain-tracing feature was able to built 163 of the 332 residues. The resolve-built partial model and the obtained model from molecular replacement approach were used as a guide to manually rebuild the model with Coot. Refinement was performed with Refmac5 (Murshudov et al., 1999) with the same set of 5% of reflections reserved for Rfree crossvalidation (Brünger, 1993). The final model was obtained after iterative cycles of model building and refinement using Refmac, ARP/wARP (Perrakis et al., 1999), and Coot (Emsley and Cowtan, 2004) resulting in values for R and Rfree of 0.199 and 0.257, respectively. Refinement was carried out with non-crystallographic symmetry (NCS) restraints and TLS parameters. The Ramachandran plot gave 100% of the residues in the most favourable or additionally allowed regions, done with PROCHECK (Morris et al., 1992; Laskowski et al., 1993). Further assessment of the model quality (rotamer analysis and peptide orientation analysis) as well as automatic water oxygen assignment was performed with Coot. Final refinement statistics are summarised in Table 2. Possible hydrogen bonds and salt bridges were detected with Protein-Protein interaction server (Jones and Thornton, 1996) using default parameters. The contact area of the two NCS related monomers as well as the solvent accessible areas were calculated using the program areaimol of the CCP4 program suite (Lee and Richards, 1971). Docking of the two protein domains was performed using ClusPro (http://nrc.bu.edu/cluster/; Comeau et al. 2004). Protein fold similarity searches were done using the Dali server (Holm and Sander, 1998); molecular images were created with PyMOL (http://www.pymol.org).

The coordinates and structure factors will be deposited in the Protein Data Bank soon.

1.2.9. Size exclusion chromatography

Size exclusion chromatography was performed on a Superdex 200 HR 10/30 column (Amersham Biosciences) equilibrated with 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2

mM MgCl$_2$ and 2 mM DTT. The column was calibrated using the gel filtration molecular weight markers (MW-GF-200) from Sigma.

1.2.10. SDS-PAGE

SDS-PAGE was performed according to Laemmli (1970) under reducing conditions in the presence of 2.5% (v/v) β-mercaptoethanol and 1.5% (w/v) SDS. Gels were stained with colloidal Coomassie blue staining (Roti-Blue, Roth) to visualize protein bands.

1.2.11. Multiple sequence alignment

Multiple sequence alignment was performed with MultAlign using default parameter (http://bioinfo.genopole-toulouse.prd.fr/multalin/multalin.html; Corpet, 1988).

1.3. Results

1.3.1. Amino acid sequence analysis

The murine CSS is a nuclear resident protein composed of 432 amino acids with a calculated molecular weight of 48.01 kDa (Münster et al., 1998). While the Sia activating domain (CSS-NT) is located within the first 267 residues (Krapp et al., 2003), the function of the C-terminal domain (CSS-CT, residues 267 to 432) is unknown. The alignment of verified ($^+$) and putative vertebrate CSS-CT sequences as shown in figure 6 (upper part), illustrates that conservation exists between the C-terminal domains from zebrafish to human. Moreover, when the amino acid sequence of the murine CSS-CT was used as a template in PSI-Blast search and hits with already identified C-terminal domains of vertebrate CMP-sialic acid synthetases are excluded, the protein has highest homology to bacterial 3-deoxy-D-*manno*-octulosonic acid-8-phosphate (KDO8P) phosphatases (Fig. 6, lower part). The overall similarity in the amino acid composition between murine CSS-CT and KDO8P phosphatases was about 51 to 52% (27% to 30 % sequence identity). KDO8P phosphatases belong to the *haloacid dehalogenase* (HAD) superfamily which is characterised by a common α,β-hydrolase fold and three conserved sequence motifs (Motif I to Motif III shown in Fig. 6) defining the active site and harbouring substrate as well as cofactor-binding residues. Conserved residues, important for enzymatic activity, are marked by black circles in figure 6. The initial aspartic acid of the amino-terminal motif I (DXDXT/V) is invariant across the HAD family (Collet et al., 1998). This aspartic acid residue coordinates a divalent cation, in

most cases Mg^{2+}, and even more important, acts as nucleophile to form the phosphoaspartate intermediate with the substrate (Collet et al., 1998). The aspartylphosphate is intrinsically high in energy, which is then used to drive phosphoryl transfer to water, the predominant phosphoryl-group acceptor. In motif II (S/T) the conserved serine or threonine residue is involved in hydrogen bonding to the phosphoryl oxygen (Morais et al., 2000). The residues of motif III (KX_{18-30}(G/S)(D/S)XXX(D/N)) act as general acid/base, interact with an active site divalent metal cation, and are also involved in phosphoryl oxygen hydrogen bonding. The divalent metal ion Mg^{2+}, is essential for catalysis since it coordinates both, important Asp residues, (**D**XDXT/V and (KX_{18-30}(G/S)(**D**/S)XXX(**D**/N)) and the phosphorylated substrate, thereby providing orientation and charge shielding for nucleophilic attack (Wang et al., 2002; Allen and Dunaway-Mariano, 2004).

Alignment of vertebrate CSS-CTs demonstrate the presence of all three HAD sequence motifs, however, some residues shown to be involved in dephosphorylation are not conserved (Fig 6). In case of mammalian CSS-CTs motif I (DXDXT/V) is represented by the sequence NXDXC, motif II (S/T) by serine and motif III (KX_{18-30}(G/S)(D/S)XXX(D/N)) by KX_{21}GNXXXD. Other CSS-CT sequences basically show the same pattern with few additional modifications, especially in the first residues of motif I and motif III, respectively. Most conspicuously, all CSS-CT sequences lack the first aspartate residue of motif I, which, because of its essential role in catalysis, is the only non-exchangeable residue of HAD proteins. Furthermore the first aspartate residue of motif III is exchanged to Asn. For the last residue of motif I, which in all vertebrate CSS-CTs is a cysteine residue instead of Thr or Val, no catalytic function is described. Summarised, two strictly conserved aspartic acid residues important for catalysis and coordination of Mg^{2+} are exchanged to asparagines in CSS-CT sequences (Asn^{279} and Asn^{371} in mCSS-CT). Beside sequence motifs common to all HAD proteins, KDO8P phosphatases are unified by an additional sequence motif at the C-terminus (Wu and Woodard, 2003). Interestingly, this motif, defined by the amino acid sequence GGXGAXRE, is highly conserved in all vertebrate CSS-CTs (Fig. 6, marked by diamonds). The exact function of this sequence motif remains to be determined.
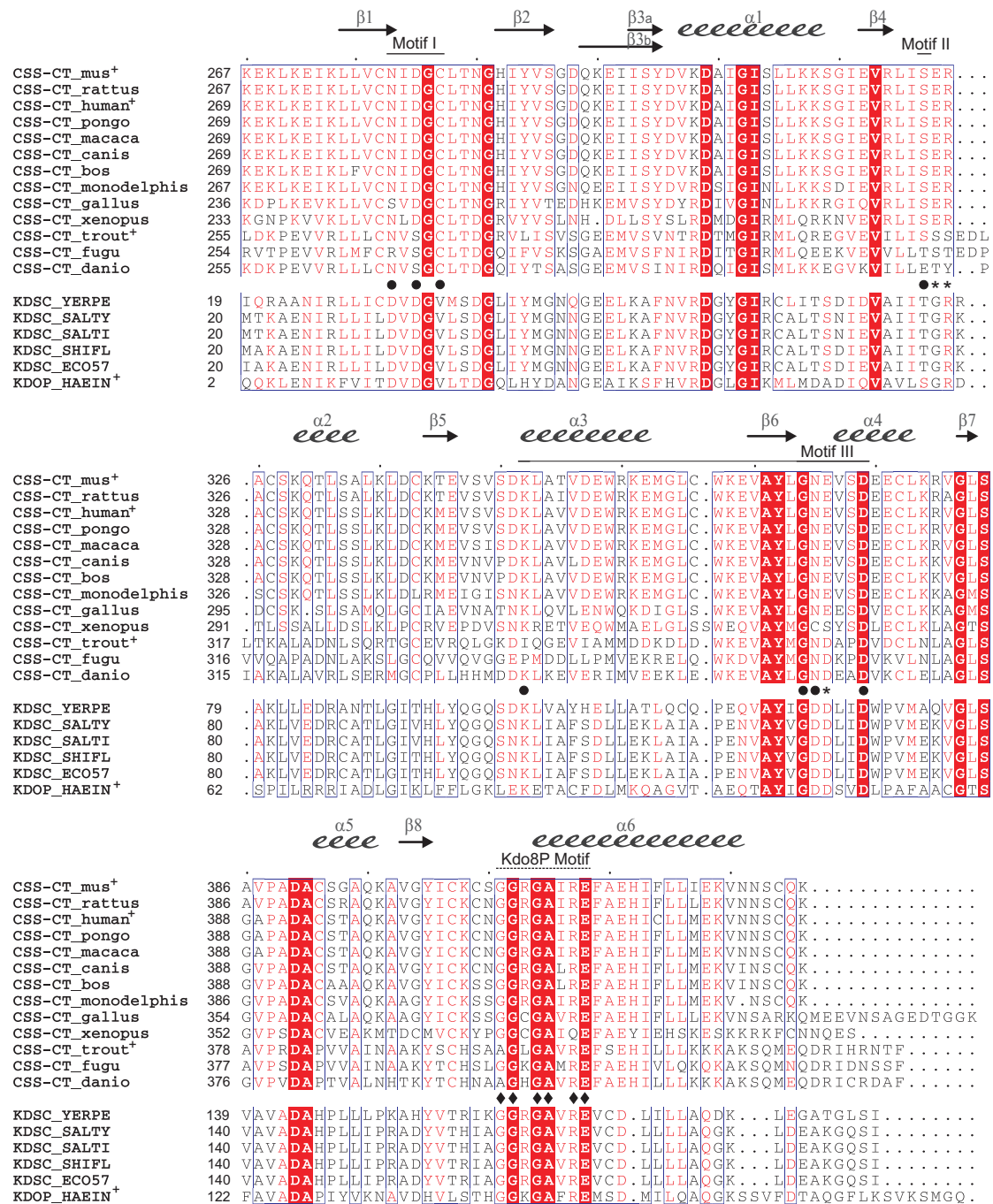
Fig. 6 Multiple sequence alignment of murine CSS-CT with other vertebrate CSS-CT sequences (upper part) and bacterial Kdo8P phosphatases (lower part) obtained by PSI-BLAST search. Alignment was constructed with Multalign version 5.4.1 (Corpet, 1998). The secondary structural elements of mCSS-CT are shown on top as arrows (β-strands) and corkscrews (α-helices). Fully conserved residues are shown as white characters with red background; similar residues are depicted as red characters. Residues that are homologous to conserved residues in motif I, II and III of the HAD superfamily are indicated by circles; residues conserved in KDO8-P phosphatases are marked by diamonds. Asterisks indicate residues partially occluding the active site in case of mCSS-CT crystal structure. (⁺) marks amino acid sequences with verified protein function.

Sequences with following accession numbers were used (see next page).

Fig. 6 SwissProt sequence codes: Q99KK2 (CSS-CT_mus), P69060 (CSS-CT_rattus), Q8NFW8 (CSS-CT_human), Q5R6R5 (CSS-CT_pongo), Q3SZM5 (CSS-CT_bos), Q90WG6 (CSS-CT_trout), Q8ZB47 (KDSC_YERPE), Q8ZLS0 (KDSC_SALTY), Q8Z3G5 (KDSC_SALTI), P67654 (KDSC_SHIFL), P0ABZ5 (KDSC_ECO57), P453134 (KDOP_HAEIN); accesion numbers of putative CSS: XP_001099303 (CSS-CT_macaca), XP_534866 (CSS-CT_canis), XP_001362472 (CSS-CT_monodel), XP_416429 (CSS-CT_gallus), NP_001090750 (CSS-CT_xenopus), CAG28328 (CSS-CT_fugu), CAK18993 (CSS-CT_danio).

1.3.2. Enzymatic activity

To investigate the physiological function of vertebrate CSS-CT domains, *in vitro* activity assays were performed with the recombinantly expressed and purified murine enzyme. Thereby, both full-length CSS as well as CSS-CT were tested with various substrates using EnzChek TM phosphatase assay Kit (Molecular Probes), which is based on measurement of free inorganic phosphate. Due to the above described amino acid exchanges in critical positions of HAD motif I and III, mutants were generated reconstituting the classical active site of HAD phosphatases. Therefore the two asparagine residues of motif I and III were exchanged to aspartic acid (N279D and N371D), and single as well as double mutants of full-length CSS and CSS-CT were investigated. Activity assays were supplemented with $MgCl_2$, since all phosphotransferases of the HAD superfamily require Mg^{2+} as cofactor.

Due to high sequence homology to KDO8P phosphatases, murine CSS-CT was tested for phosphatase activity with several commercial available sugar-phosphates, listed in table 1. Additionally dephosphorylation of the generic substrate 4-nitrophenyl phosphate, nucleotides, pyrophosphate and phosphorylated amino acids was investigated (Table 1). However, no phosphatase activity could be obtained with native or mutated murine CSS-CT and full-length CSS, respectively. Since the eight-carbon sugar KDO is not present in eukaryotes and its phosphorylated form KDO8P not purchasable, *N*-Acetylneuraminic acid 9-phospate (Neu5Ac9P) was investigated as substrate, a closely related sugar-phosphate to KDO8P in eukaryotes. In the eukaryotic sialic acid metabolic pathway Neu5Ac9P dephosphorylation precedes CSS catalysed sialic acid activation and at the beginning of this work, no specific Neu5Ac9P phosphatase was identified. Hence Neu5Ac9P is not commercially available, it was generated by condensation of phosphoenolpyruvate and 2-acetamido-2-deoxy-D-mannose-6 phosphate (ManNAc6P), which itself was not purchasable and thus synthesized as described by Liu and Lee (Liu and Lee, 2001). For visualisation of Neu5Ac9P dephosphorylation a coupled assay was established in which CMP-Neu5Ac biosynthesis was simulated in the NMR tube. Starting with ManNAc6P and PEP, consecutive addition of recombinantly expressed and purified

murine Neu5Ac9P synthase, CSS-CT (or calf intestinal alkaline phosphatase, CIP, as positive control) and full length CSS was performed. Product formation was monitored by 1H-NMR spectroscopy. Generation of the intermediate product Neu5Ac9P was constantly observed, but characteristic chemical shifts for CMP-Neu5Ac were only obtained when using the non specific phosphatase CIP, proving that neither native, nor mutated murine CSS-CTs are able to dephosphorylate Neu5Ac9P *in vitro*.

Table 1. Compounds used for CSS-CT activity assays

Phosphatase activity assay
tested compounds
p-Nitrophenylphosphate
Glucose–1-phosphate Glucose–6-phosphate Glucosamine-6-phosphate GlcNAc-1-phosphate GlcNAc-6-phosphate Fructose-6-phosphate Neu5Ac-9-phosphate Mannose-1-phosphate Ribose-5-phosphate
Glycerol-1-phosphate Phosphorylethanolamine Phosphocholine Phosphocreatine 6-Phosphogluconic acid Phosphoenol pyrovate
Phosphoserine Phosphothreonine Phosphotyrosine
ATP, CTP, CDP, UMP Pyrophosphate

1.3.3. Crystallisation of CSS-CT

Significant similarity of vertebrate CSS-CTs to bacterial KDO8P phosphatases was obtained on primary sequence level, but no activity could be observed by investigation of native as well as engineered murine protein in the above described *in vitro* assays. To verify if homology to functional HAD phosphatases is also reflected in 3D structure, the crystal structure of CSS-CT (aa 267-432) was determined. A cDNA-fragment encoding the C-terminal domain of mouse CSS-CT was expressed in *E.coli* as a StrepII-fusion protein and the resulting gene product was affinity purified to homogeneity (verified by SDS-PAGE). To improve diffraction quality, the StrepII-tag had to be removed by proteolytic cleavage and precipitation with Streptavidin-Agarose. Finally, crystallisation trials resulted in well diffracting crystals using EtOH as precipitant. Since the initial molecular replacement approach using a native dataset and the 3D structure of KDO8P phosphatase from *Haemophilus influenzae* (PDB entry: 1J8D) did not result in a reliable model, further phase information was included, provided by an additional multiple-wavelength anomalous dispersion (MAD) dataset. The resulting CSS-CT crystal structure was refined to a resolution of 1,9 Å. Native and selenomethionine (SeMet)-substituted crystals contained two monomers in the asymmetric unit. The final model was refined to a free R value of 25.7%, ranging from lysine 271 to serine 429. Crystal data collection and refinement statistics are shown in Table 2. The quality of the model is well-defined with 90.5% of the residues adopting ϕ, ψ angles in the "most favoured" and 9.5% in the "additionally allowed" regions of the Ramachandran plot (obtained with PROCHECK).

Table 2. Data collection, phasing and refinement statistics

	Native	SeMet		
Data collection				
Space group	C2	C2		
Cell dimensions (Å)				
a	121.4	121.2		
b	37.5	37.6		
c	78.7	79.0		
β	116.3	116.6		
		edge	*peak*	*remote*
Wavelength	0.9184	0.9797	0.9799	0.9117
Resolution (Å)	50 – 1.89	70 – 1.90	70 – 1.90	70 – 1.90
	$(1.97 - 1.90)^a$	(2.0 – 1.90)	(2.0 – 1.90)	(2.0 – 1.90)
R_{meas}^{b} (%)		4.9 (37.9)	4.7 (32.5)	5.4 (42.7)
R_{sym}^{c} (%)	4.2 (25.2)	3.7 (28.5)	3.6 (24.5)	4.0 (32.0)
I/σ	27.60 (3.40)	14.78 (3.20)	15.36 (3.70)	14.03 (2.86)
Completeness (%)	95.6 (69.5)	98.1 (98.9)	98.3 (98.8)	97.8 (99.0)
Redundancy	4.2 (2.7)	2.04 (1.96)	2.04 (1.96)	2.06 (1.99)
Refinement				
Resolution (Å)	50 – 1.89			
No. of unique reflections	23 173			
R_{work} / R_{free}	0.199 / 0.257			
No. residues				
Protein	2408			
Water	219			
B-factors				
Protein	38.9			
Water	45.9			
RMSD				
Bond length (Å)	0.022			
Bond angles (°)	1.69			
Ramachandran plot statistics^d (%)				
Most favoured	90.5			
Additional allowed	9.5			
Disallowed	0			

^{a} Values in parentheses are for the highest-resolution shell

^{b} R_{meas}: redundancy independent R-factor

^{c} R_{sym}: $\Sigma_{hkl}\Sigma_i$ |$I_{hkl, i}$ - <I>$_{hkl}$| / Σ |<I>$_{hkl}$|

^{d} Ramachandran plot statistics were obtained with PROCHECK

1.3.4. Overall structure

CSS-CT is a complex of four subunits as shown in figure 7A, with two pairs of slightly different polypeptides, called A (green and magenta) and B (yellow and blue). The twofold rotational symmetry axis of the A2B2 tetramer with the approximate dimensions of 70 Å x 70 Å x 35 Å corresponds to a unit-cell symmetry axis. The

asymmetric unit is a single AB dimer with the two molecules related by a non-crystallographic 2-fold axis of symmetry. Due to tetramer formation each monomer stays in contact to two neighbouring monomers; therefore roughly 19% (1480 $Å^2$) of the surface area of each monomer is buried. Both polar and hydrophobic interactions are present within this buried surface to each of the two adjacent monomers. Eleven hydrogen bonds and one salt linkage gives a total of 24 (2x12) contacts per monomer. Fig. 7B shows a ribbon diagram of tertiary structure of monomer B with the N-terminus coloured in blue and the C-terminus coloured in red. Each monomer constitutes a single domain arranged in an α/β-hydrolase fold. The centre of six parallel β-strands, with strands ordered as β5, β4, β1, β6, β7, β8, is surrounded by six α-helices in right-handed β-a-β connectivity. An extended loop (HP1, residues His^{288} to Tyr^{301}), consisting of the two antiparallel β-sheets β2 and β3, is inserted between strand β1 and helix α1 as depicted in schematic representation of the secondary structure topology in figure 7C. The HP1 hairpins of the four monomers form a central eight stranded beta-barrel, stabilising the whole tetramer (Fig. 7A). The inside of the barrel has a width of maximal 13 Å. An overlay of monomer A (blue) and monomer B (yellow) is shown in figure 7D, with major mainchain deviations indicated by arrows. The most prominent difference between the two monomers can be found in the beta-hairpin orientation. Another conformational variance between monomer A and B was found in helix α2 (Lys^{329} to Ser^{333}) that stays in contact with HP1 of the neighbouring monomer. Compared with other regions of the protein, residues in this helix have increased temperature factors, indicating a higher degree of flexibility.
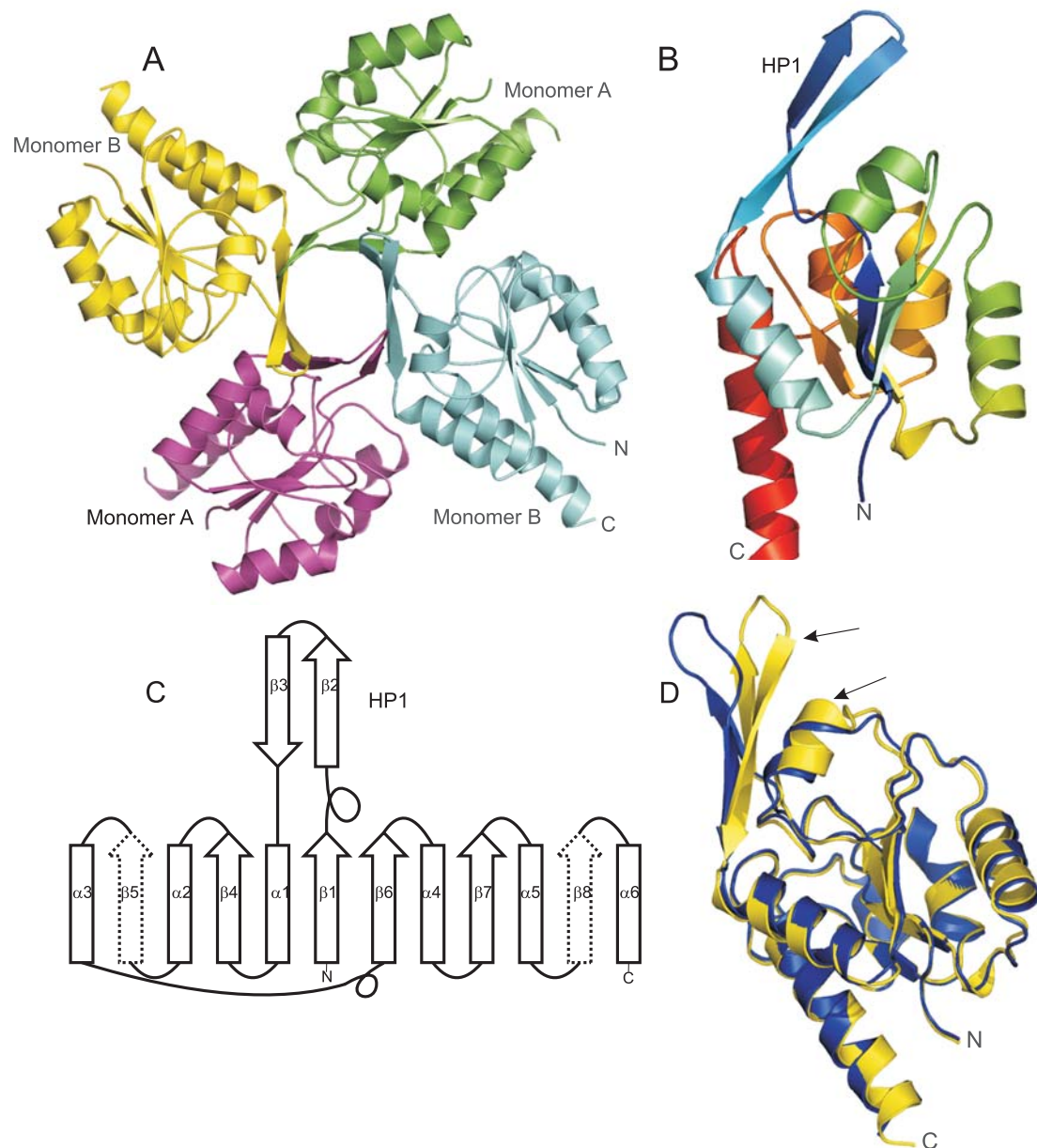
Fig. 7: Crystal structure of mouse CSS-CT

(A) The quaternary structure is represented as a ribbon diagram, monomers A are green and magenta, monomers B yellow and blue; (B) ribbon diagram of tertiary structure of monomer B. The chain is coloured from blue at the N terminus to red at the C terminus; (C) representation of the secondary structure topology; α-helices and β-sheets are depicted as bars and arrows, respectively. The inserted hairpin (HP1) is comprised of β-sheets β2 and β3. Dotted lines indicate imperfect secondary structures due to incomplete refinement. (D) Overlay of monomer A (blue) and B (yellow), main structural deviations between the two monomers are indicated by arrows.

1.3.5. Related structures

The automated structure superposition of CSS-CT against the Protein Data Bank (PDB) with the Dali algorithm (Holm 1996), resulted in a list of structural neighbours, all belonging to the HAD superfamily (Table 3). Proteins with highest structural agreements were the hypothetical protein AQ2171 from *aquifex aeolicus* (PDB entry 2P9J), the cobalt bound form of the above mentioned KDO8P phosphatase from *haemophilus influenzae* (PDB entry: 1K1E), and the HAD-like phosphatase Yida from *Escherichia coli* (PDB entry 1rkq) with a RMSD of 1.8 Å, 2.0 Å and 1.9 Å, respectively. Reflecting the strong structural relation, all functional characterised proteins listed in table 3 comprise phosphatase activity, with phosphoglycolate phosphatases in the main (PDB entries: 1wr8 and 1l6r). Remarkably, all proteins are of bacterial origin with half of them belonging to thermostable species.

Table 3. **Results of a DALI search for protein structures similar to CSS-CT**

PDB code	Protein	Source	Z score	RMSD (Å)	Sequence identity (%)
2p9j	hypothetical protein AQ2171	*Aquifex aeolicus*	23.7	1.8	34
1k1e	KDO8-P phosphatase	*Haemophilus influenzae*	23.2	2.0	28
1rkq	HAD-like phosphatase Yida	*Escherichia coli*	20.1	1.9	16
1nrw	HAD-like hydrolase	*Bacillus subtilis*	19.9	1.7	17
1wr8	phosphoglycolate phosphatase	*Pyrococcus horikoshii*	19.3	2.2	22
1l6r	phosphoglycolate phosphatase	*Thermoplasma acidophilum*	19.1	2.1	22

The CSS-CT monomer exhibit the α,β –hydrolase fold characterising the HAD superfamily (Li et al., 1998) (Fig. 7B, C). The beta-hairpin insertion HP1 is topologically equivalent to the insertion site of cap domains in the HAD superfamily and lead CSS-CT to subclass III of the HAD superfamily. Similarity of mouse CSS-CT to KDO8P phosphatases at primary sequence level is reflected by impressive similarities in 3D structure. Like CSS-CT, KDO8P phosphatases belong to subclass III of the HAD superfamily. Superposition of the 3D coordinates of CSS-CT and *H. influenzae* KDO8P phosphatase resulted in a RMSD of 2.0 Å, underlining their outstanding structural similarity, which is already apparent by visual inspection of the two structures, because of only minor variations in Cα-trace (Fig. 8A).

1.3.6. Active site

By homology to HAD phosphatases the CSS-CT active site could be identified at the subunit interface; thus the whole tetramer comprises four putative active sites. Residues of motif I to III cluster in the putative active site of CSS-CT, however, no divalent cation was found within the active site pocket. Despite the amino acid exchanges (Asn^{279} and Asn^{371}), the CSS-CT active site superimposes extremely well with other HAD superfamily members, illustrated by overlay of the active sites of CSS-CT and *H. influenzae* KDO8P phosphatase in figure 8B.
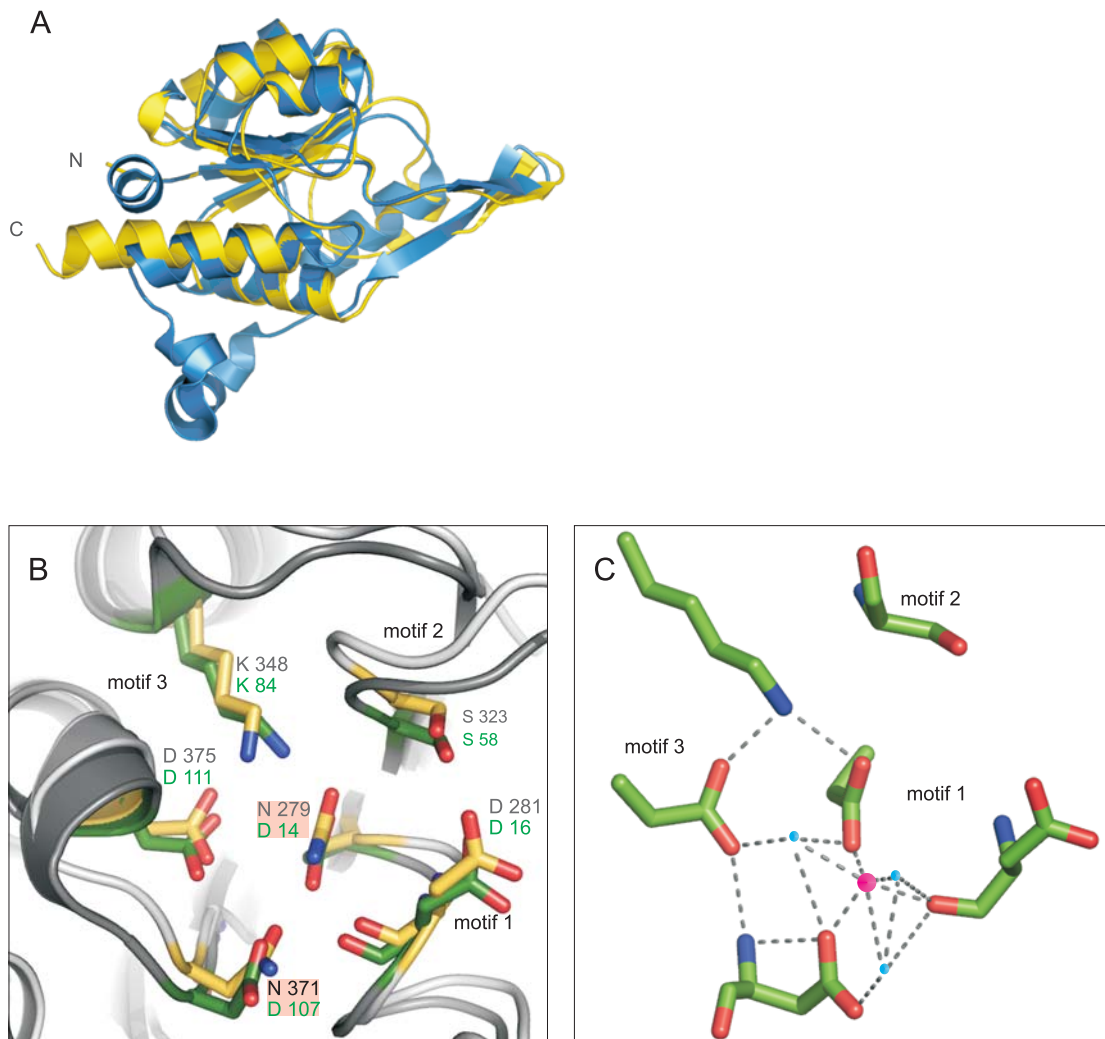
Fig. 8 Superposition of CSS-CT and KDO8-P phosphatase from *H. influenzae* (PDB entry 1J8D). Overlay of the Cα-backbones (A) and active sites (B). In Figure (A) CSS-CT is coloured in yellow, KDO8P phosphatase in blue; in (B) the active site residues are beige and green, the Cα-traces light and dark grey for CSS-CT and KDO8P phosphatase, respectively. (C) Typical arrangement of HAD phosphatase active site with coordinated Mg^{2+} and water molecules, prior to substrate binding. Mg^{2+} and water are shown as magenta and blue circles, respectively. Atomic colours are as follows: red for oxygen atoms; blue for nitrogen atoms.

Residues important for KDO8P phosphatase activity are shown in green; the corresponding residues of CSS-CT are shown in beige; mutated residues are highlighted in red. However, substitution of the two conserved asparagines entails dramatic consequences on the active site of CSS-CT. First, the surface of the active site pocket is markedly less negatively charged than in other HAD phosphatases, and second, no residue is available for nucleophilic attack and formation of the phospho-protein intermediate (Fig. 8B). Furthermore Asp[14] and Asp[107] of CMP-KDO8P phosphatase that are exchanged to Asn in CSS-CT are involved in coordination of the critical cofactor Mg^{2+}, as illustrated in Fig. 8C by arrangement of a typical HAD-phosphatase active site prior to substrate binding. Mg^{2+} is hexacoordinated and directly involved in phosphate group orientation. In case of CSS-CT, addition of $MgCl_2$ to crystallisation conditions did not impair crystal growth apparently, but diffraction ability was dramatically worsened.

Strikingly, the CSS-CT 3D-structure revealed a significantly narrowed active site entry in comparison to other HAD superfamily members. Surface presentation in Figure 9 illustrates the differences between CSS-CT monomer A and B (Fig. 9A and B) and KDO8P phosphatase from *haemophilus influenzae* (Fig. 9C; PDB entry 1J8D). Catalytic residues of the active site are coloured in red, whereas protruding amino acids are coloured in yellow. The active side entry is seen at the centre, the red spot top left represents the origin of the catalytic lysine of motif III. The two CSS-CT monomers are identical in orientation of the active site residues, but show tremendous differences in residues encircling the active site entry. Monomer A reveals a closed conformation, due to positioning of Glu[324] and Arg[325] in front of the active site entry. The displacement of Arg[325] is stabilised by a salt bridge to a glutamate residue of motif III (Glu[372]). This occlusion is less pronounced in monomer B, but in comparison with the functional KDO8P phosphatase, the active site entry is narrowed of about 0.2 Å in width. These changes are based on a more spacious turn of the motif II loop, due to sequence variation from S/TGR to S323ER in case of CSS-CT. Residues, responsible for active site occlusion are marked by asterisks in the sequence alignment in figure 6. Interestingly, residues forming the salt bridge (Glu[372] and Arg[325] in mCSS) are conserved within the mammalian enzymes.
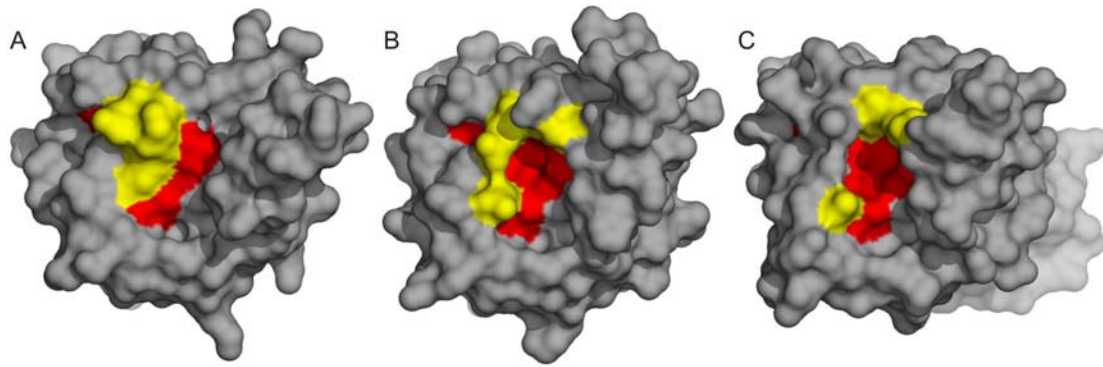
Fig. 9 Surface presentations looking down on the active-site pocket. (A) CSS-CT monomer A; (B) CSS-CT monomer B, (C) Kdo8-P phosphatase from *H. influenzae* (PDB entry 1J8D). Catalytic residues of motif I – III are coloured in red, amino acids occluding the active site are coloured in yellow:

CSS-CT: red: Asn^{279}, Asp^{281}, Ser^{323}, Lys^{348}, Asn^{371}, Asp^{375}; yellow: Glu^{324}, Arg^{325}, Glu^{372}

KDO8P-P: red: Asp^{14}, Asp^{16}, Ser^{58}, Lys^{84}, Asp^{107}, Asp^{111}; yellow: Gly^{59}, Arg^{60}, Asp^{108}

1.3.7. 3-D Model of the full-length CMP-Sialic acid Synthetase

Prior to the present study, providing structural information about the C-terminal domain of murine CMP-sialic acid synthetase, Krapp et al. (2003) solved the crystal structure of the N-terminal domain (CSS-NT). Native CMP-sialic acid synthetase comprises 432 amino acid residues, and the structures of the two separate domains complement one another by defining 3D coordinates of residues 40 to 267 (CSS-NT, Fig. 10A) and 271 to 429 (CSS-CT, Fig.10B), respectively. By combination of both structures, a model of the entire protein was developed, which is shown in Fig. 10C. An initial model was generated manually, by inspection of both structures and electrostatic surfaces. The model was later confirmed by ClusPro, a web server for protein-protein docking (Comeau et al., 2004). In this model the CSS-CT tetramer acts as a platform for two independently acting CSS-NT dimers, placed site by site on top of one CSS-CT tetramer. A ribbon diagram of one functional dimer of CSS-NT is shown in Fig. 10A, and a side-view of the CSS-CT platform is depicted in Fig. 10B. In the final model, the C-terminal ends of the CSS-NT chains reside in close proximity to the N-termini of CSS-CT chains, so that both CSS domains could be linked without impairing the overall folds. A surface representation of the quaternary structure of the modelled full-length CSS tetramer is depicted in Fig. 10C. CSS monomers are shown in yellow, blue, green and red, respectively. Two intertwined dimerisation domains which build the active sites of the CSS-NT for Sia activation are visible on top and are freely accessible. The putative active sites of the CSS-CT, identified by homology to the active sites of other HAD proteins, are located at the bottom of the tetramer. Exemplarily, illustrating their exposed positions, active sites

located at the interface of the red and green monomer, are marked by white circles in figure 10C.
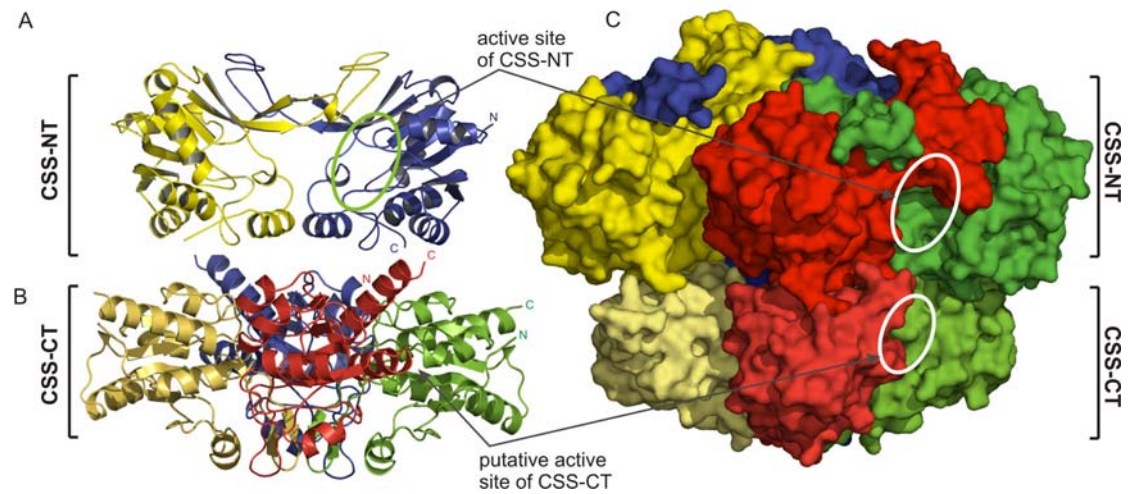
Fig. 10 3D-Model of mature murine CMP-Sialic Acid Synthetase

(A) Ribbon diagram of the Sia-activating CSS-NT dimer with the monomers coloured in blue and yellow (PDB entry 1qwj); (B) Ribbon diagram of CSS-CT tetramer (side-view) with the monomers shown in green, blue, yellow and red. (C) Surface representation of the modelled mature murine CSS with the full-length monomers given in blue, yellow, green and red at with the N-terminal domains coloured in a slightly darker shade, than the corresponding C-terminal domains. Two CSS-NT dimers are positioned side by side on top of one CSS-CT tetramer. White circles indicate the position of active sites at the red and green monomer interface.

In agreement with the CSS 3D model, the recombinantly expressed and purified full-length CSS exclusively forms tetramers in solution, as obtained from size exclusion chromatography experiments (Fig. 11B). In figure 11G, the calculated as well as the observed molecular weights (Mr) of CSS-NT, CSS-CT and full-length CSS are depicted. On basis of the observed Mr the oligomerisation status was calculated. Summarised, results obtained from size exclusion chromatography experiments are consistent with the 3D model of full-length CSS. In contrast, CSS-NT exclusively forms homodimers in solution (Fig. 11C), supporting two independent CSS-NT dimers on top of CSS-CT platform as assumed for CSS modelling. Hydrophobic areas which mediate dimerisation of two functional CSS-NT dimers in the protein crystal are positioned at the CSS-NT - CSS-CT interface in the CSS full-length model. CSS-CT forms a homotetramer in solution (Fig. 11D) as well as in 3D structure and is responsible for full-length CSS tetramerisation. We further investigated whether the separately expressed domains spontaneously assemble to oligomeric complexes in solution.
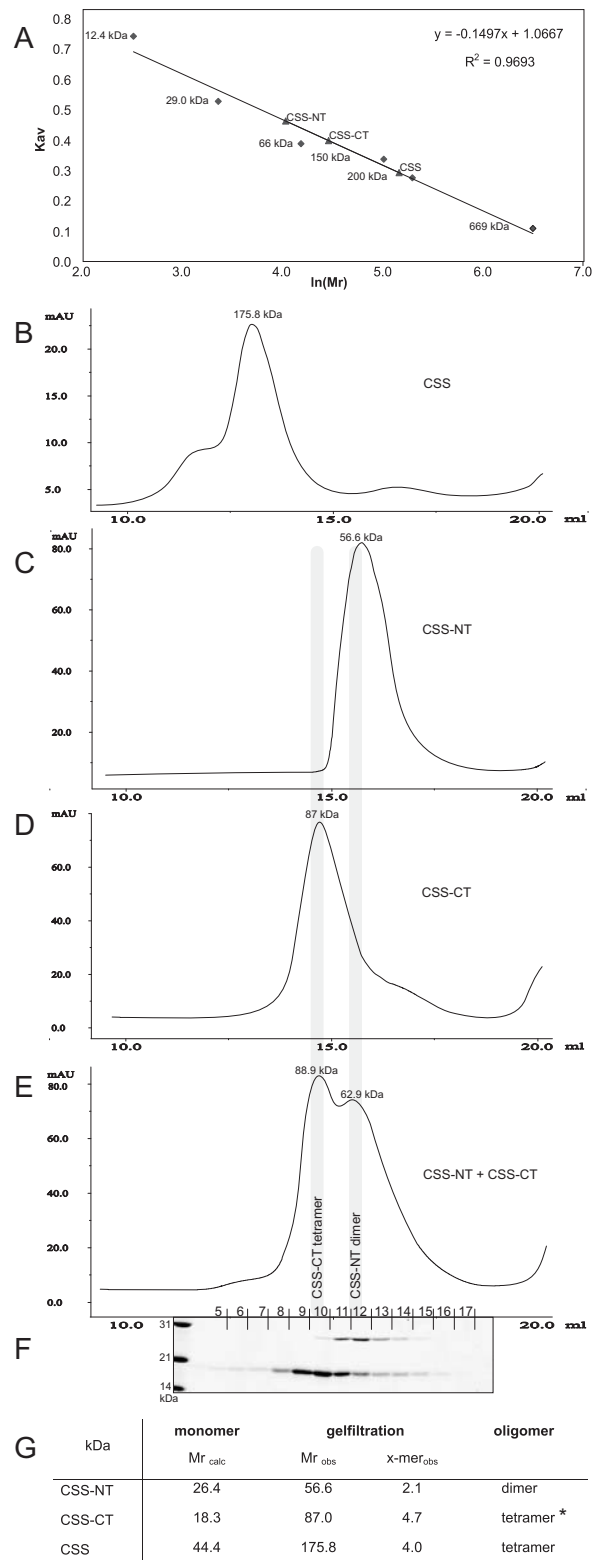
Fig. 11 Size exclusion chromatography experiments with murine CSS, (A) calibration curve, standard proteins represented by squares, samples by triangles; (B) full-length protein (CSS, residues 39-432); (C) N-terminal domain (CSS-NT, 39-267); (D) C-terminal domain (CSS-CT, residues 267-432); (E) mixture of CSS-NT and CSS-CT showing that both proteins behave independently, (f) fractions were analysed by SDS-PAGE and coomassie staining; (G) evaluation table.

*further gelfiltration experiments and the 3D structure confirmed the tetrameric organisation of CSS-CT

However, size exclusion chromatography of a protein mixture containing CSS-NT and CSS-CT revealed an independent behaviour of both CSS domains, without forming any mixed multimers (see Fig. 11E). SDS-PAGE analysis of this chromatogram is depicted in figure 11F, proving that the smaller CSS-CT domain (residues 267-432, Mr 18 kDa) elutes earlier than the bigger CSS-NT domain (residues 36-267, Mr 26 kDa) from size exclusion column, indicating a higher oligomeric state.

1.4. Discussion

The C-terminal domain of murine CSS (CSS-CT), with a calculated size of about 18 kDa, is not required for the sugar activating function of the mature enzyme (Krapp et al., 2003). Interestingly, CSS-CT is conserved within all known vertebrate CSS, and shows homology to sugar phosphatases of the HAD family (Fig. 6); especially to 3-deoxy-D-manno-octulosonate 8-phosphate (KDO8P) phosphatases (E.C. 3.1.3.45) that catalyze the hydrolysis of KDO8P to KDO and inorganic phosphate. KDO is an 8-carbon sugar that links the polysaccharide moieties of the lipopolysaccharide to lipid A in Gram-negative bacteria but is absent in Gram-positive bacteria and eukaryotes. KDO is synthesized in a similar way to sialic acids. After dephosphorylation of KDO8P, KDO is activated to its Cytidine-monophosphodiester (CMP-KDO) prior to transfer to glycolconjugates. Activation is performed by the CMP-KDO synthetase, which shows homology to murine CSS-NT in primary sequence as well as in 3D structure (Jelakovic and Schulz, 2002). Thus, the two domains of vertebrate CSS, CSS-NT and CSS-CT, are homologous to separated bacterial enzymes, responsible for KDO activation and KDO8P dephosphorylation, respectively. Moreover, Sia and KDO are the only sugars that are activated by linkage to CMP. Synthesis and activation of sialic acid and KDO thus appear to be mechanistically similar processes, as suggested by the phylogenetic relationships between some of the biosynthetic enzymes (Angata and Varki, 2002).

Primary sequence analysis of murine CSS-CT revealed homology to HAD proteins, but also disclosed major aberrations in conservation of the three HAD sequence motifs. By structural analysis, murine CSS-CT was finally classified into subgroup III of HAD phosphatases, because of (i) the overall α/β-hydrolase fold with the beta-hairpin insertion between motif I and II, (ii) the incomplete but apparent presence of the three sequence motifs conserved throughout HAD-family proteins (iii) the perfect agreement in steric arrangement of active site residues and (iv) presence of GGXGAXRE sequence motif which is proposed to be a common feature of KDO8P phosphatases, also belonging to subgroup III of HAD superfamily. The function of the

residues in the GGXGAXRE motif is so far unknown. However, residues might be involved in substrate specificity or closure of the active site, since it is positioned above the active site cavity of the adjacent monomer.

HAD proteins are classified to the same family due to their common overall fold and active site architecture, but the homology at primary sequence level is remarkably low. However, amino acids involved in catalysis are strictly conserved (motif I to III, Fig. 6). The motif I sequence was denoted above as DXDXT/V, but only phosphotransferases display the DXD sequence in motif I. The corresponding sequence in P-type ATPases and dehalogenases is DXT and DXY, respectively (Kim et al., 2004; Collet et al., 1998). Though the HAD superfamily unifies a variety of proteins with different enzymatic activities, CSS-CT, possessing a NXD sequence, is closer related to phosphotransferases than to any other known HAD protein functionality. Notably, all known vertebrate CSS-CT sequences display equivalent mutations of catalytic aspartic acid residues to asparagines (mCSS: N279 in motif I, and N371 in motif III) known to be critical for substrate as well as cofactor binding (see figure 6 for sequence alignment). In case of already characterised HAD phosphatases, each of these mutations can lead to a complete loss of activity (Selengut, 2001). Especially the first aspartic acid residue, conserved in motif I (DXDXT/V), is indispensable for activity, because of its role as nucleophile (Allen and Dunaway-Mariano, 2004; Li et al., 1998).

These critical mutations in motif I and III indicate that if vertebrate CSS-CTs harbour enzymatic activities, they probably use an alternative mechanism compared to other HAD enzymes.

Based on sequence homology, KDO8P seems to be the most likely substrate for CSS-CT. However, this phosphorylated eight-carbon sugar is not present in eukaryotes. Thus the closely related sugar-phosphate *N*-acetylneuraminic acid 9-phosphate (Neu5Ac9P) was provided as substrate for wild type and mutant murine CSS-CT in the *in vitro* activity assay. In the sialic acid metabolic pathway, Neu5Ac9P dephosphorylation occurs directly upstream of the CSS catalysed sialic acid activation, and at the beginning of this study no specific Neu5Ac-9-Pase was known at all. Consequently Neu5Ac9P dephosphorylation could be a possible second function of a bifunctional CSS. However, no activity for CSS-CT with Neu5Ac9P as substrate could be obtained. These results are consistent with the recent isolation of a highly specific Neu5Ac9P phosphatase from rat liver (Maliekal et al., 2006), also belonging to Mg^{2+} dependent phosphatases of the haloacid dehalogenase family. Interestingly, the homologue protein from mouse is structurally highly related to CSS-CT with a Z-score of 13.8 and a RMSD of 2.8 Å (PDB-entry 2GFH), whereas on

primary sequence level no significant similarity can be found (10% sequence identity). Furthermore CSS-CT was tested for esterase activity (data not shown). However, no deacetylation ability could be obtained *in vitro* with the generic substrate 4-nitrophenyl acetate.

HAD phosphatases are Mg^{2+} dependent enzymes, with the hexacoordinated Mg^{2+} positioned in the active site (Fig. 8C). Reflecting the high affinity, the Mg^{2+} was observed in the 3D structure of *E. coli* YbiV, a HAD phosphatase, although not added to crystallisation conditions (Roberts et al., 2005). In the case of CSS-CT crystal structure, no Mg^{2+} is present in the active site. This is consistent with observations done by Wang et al (Wang et al., 2002), who did not find Mg^{2+} in the active site of phosphoserine phosphatase (PSP) from *Methanococcus jannashii* after mutation of the catalytic aspartic acid residue of motif I to asparagine (D11N), although Mg^{2+} was present at 5 mM during crystal formation. In addition, the mutated enzyme was 25-fold less active than the wild type protein. Possibly, this amino acid composition as present in native CSS-CT and mutated PSP, preclude proper Mg^{2+} coordination.

Finally, the lacking phosphatase activity of native CSS-CT can be explained by the exchange of two key residues of the active site and the inability of Mg^{2+} binding. However, reconstitution of HAD phosphatase active site by mutation of the two asparagines to aspartic acid, did not rescue phosphatase activity *in vitro* neither in the full-length enzyme nor in the separately expressed CSS-CT. However, CSS-CT 3D structure revealed further drawbacks hindering substrate entry into the putative active site. Monomer A displays a closed conformation with a lid-forming saltbridge (Arg^{325} – Glu^{372}) above the active site pocket. Although no salt bridge is formed in monomer B, the width of the active site entrance is about 0.2 Å smaller than observed for KDO8P phosphatese from *haemophilus influenzae* (Fig. 9).

Apart from structural reasons, the lack of activity might be explained by the fact that both, native and engineered CSS-CT, are probably dependent on selection of the correct substrate. HAD phosphatases can be highly specific and barely catalyse dephosphorylation of other compounds. For example KDO8P phosphatases from *E.coli* and *Haemophilus influenzae* process their natural substrate KDO8P with high turnover rates, but hardly catalyse dephosphorylation of the generic substrate pNPP (Parsons et al., 2002; Wu and Woodard, 2003). In addition it might be that a so far unknown cofactor is needed for enzymatic activity or posttranslational modifications like glycosylation that are missing in proteins expressed in bacteria but are essential for enzymatic activity.

Beside the possibility of not having used the right substrate or cofactor, CSS-CT might have evolved a non-enzymatic function. Homology to HAD phosphatases

points towards a common evolutionary origin, but the marked differences in active site architecture advert to a physiological function different from phosphatases, and possibly might be linked with the oligomeric state of mature CSS. As proven by the presented size exclusion chromatography experiments, CSS-CT mediates tetramer formation of mature CSS (Fig. 11). Though a CSS-NT dimer is sufficient for sialic acid activation, it is not able to arrange as tetramer in solution. Hence, tetramer formation of the CSS-NT 3D structure as reported by Krapp et al. (2003) apparently seems to be a crystallisation artefact. Advantages of tetramer formation might be an increased stability towards environmental factors, like changes in temperature, pH or proteases supported by striking structural similarities of CSS-CT to proteins of thermo stable bacteria (PDB entries 2p9j, 1wr8, 1l6r in Table 3). Further advantage of tetramerisation might be the increased overall size of CSS from 30 to 48 kDa for better regulation of its subcelluar distribution. Proteins smaller than 40 kDa do not require active transport and are able to pass the nuclear pore complex (NPC) by diffusion for entering or leaving the cell nucleus.

In most cases, bifunctional proteins catalyse two steps of the same metabolic pathway (Hinderlich et al., 1997; Brilli and Fani, 2004); less often, catalytic and regulatory properties are combined (Suck and Ficner, 1996). But there are also examples in nature, where significance of bifunctionality is not clear (for review, Nagradova, 2003). A possible advantage of bifunctional enzymes is that specific interactions between domains may provide the structural basis for conformational transitions modulating the catalytic properties of an enzyme. The generated model of mature CSS suggests that a direct interaction of the two active sites does not take place. But CSS-CT might play a regulatory role for CSS activity and localisation by influencing posttranslational modifications, like phosphorylations, glycosylations or SUMOylations.

This study presented the first 3D structure of a C-terminal domain of a CMP-sialic acid synthetase. Though the association of CSS-CT with a second function is still likely, its physiological role remains unclear. Nonetheless, detailed insights into enzymatic organisation and plausible explanations for the lack of activity were provided. To further investigate the impact of CSS-CT, it would be reasonable to create a mouse model lacking the C-terminal domain of the CMP-sialic acid synthetase.

Chapter 2

Production and characterisation of polyclonal and monoclonal antibodies specific for mouse CMP-Sialic acid Synthetase

2.1. Introduction

In mammalian cells glycan structures of glycoproteins and glycolipids are often terminated with sialic acid (Sia) residues. Due to its negative charge and hydrophilic character, the nine-carbon sugar Sia forms highly hydrated structures. Thus, in normal development cell surface Sia have regulatory effects on cell-cell interactions; whereas altered Sia patterns are associated with tumorigenesis and oncogenic transformations in vertebrates (Varki and Varki 2007; Seidenfaden et al., 2003; Hildebrandt et al., 1998). In addition, Sia is essential for embryonic development in mammals (Schwarzkopf et al., 2002). Furthermore, Sia serve as ligand for endogenous lectins in course of immune responses and are important factors in regulation of protein half-lives (Morell et al., 1971; Traving and Schauer, 1998).

In eukaryotic cells Sia biosynthesis starts with uridine diphospho–*N*-acetylglucosamine (UDP-GlcNAc), deriving from cytosolic glucose. The first two steps are catalysed by a bifunctional enzyme, the UDP-GlcNAc 2-epimerase/*N*-acetylmannosamine (ManNAc) kinase (GNE/MNK). The GNE domain of the enzyme catalyses the epimerisation of UDP-GlcNAc to ManNAc (EC 5.1.3.14), immediately followed by ATP consuming phosphorylation giving ManNAc-6-phosphate (MaNAc-6-P) catalysed by the MNK domain (EC 2.7.1.60). ManNAc6P is further condensed with phosphoenolpyruvate, followed by dephosphorylation of the generated *N*-acetylneuraminic acid-9-phosphate (Neu5Ac9P or Sia-9P). Both reactions, catalysed by Neu5Ac9P-synthase (EC 2.5.1.57) and Neu5Ac9P-phosphatase (EC 3.1.3.29), respectively, take place in the cytoplasm of the cell. Prior to transfer onto nascent glycoconjugates in a α(2,3), α(2,6) or α(2,8) linkage in the Golgi apparatus, Sia has to be activated as a cytidine monophosphate (CMP) diester. This reaction takes place in the nucleus and is catalysed by the CMP-sialic acid synthetase (CSS, EC 2.7.7.43), transferring the CMP moiety of cytidine triphospate (CTP) to the anomeric hydroxyl group of Sia in the presence of Mg^{2+}. This transport intense process is unique to Sia, since biosynthesis as well as activation of other sugars takes place in the cytoplasm.

CSS cDNAs have been cloned and characterised from *Escherichia coli* K1 (Zapata et al., 1989), *Neisseria meningitidis* serogroup B (*Nm*B; Edwards and Frosch, 1992), *Streptococcus agalactiae* (Haft and Wessels, 1994), *Haemophilus ducreyi* (Tullius et al., 1996), *Haemophilus influenzae* (Ishige et al., 2001), *Mus musculus* (Münster et al., 1998), *Oncorhynchus mykiss* (Nakata et al., 2001), *Homo sapiens* (Lawrence et al., 2001), *Drosophila melanogaster* (Viswanathan et al., 2006) and *Clostridium thermocellum* (Mizanur and Pohl, 2007). Interestingly, CSS are conserved from bacteria to man and five highly conserved sequence motifs have been identified (Münster et al., 2004). Figure 12 shows a schematic alignment of selected CSS with the conserved motifs highlighted as grey bars. Crystal structures, obtained from *Nm*B CSS and the catalytic active N-terminal domain of the murine enzyme (residues 39-267) revealed formation of functional dimers, with the active site formed by clustering of the five sequence motifs (Mosimann et al., 2001; Krapp et al., 2003).
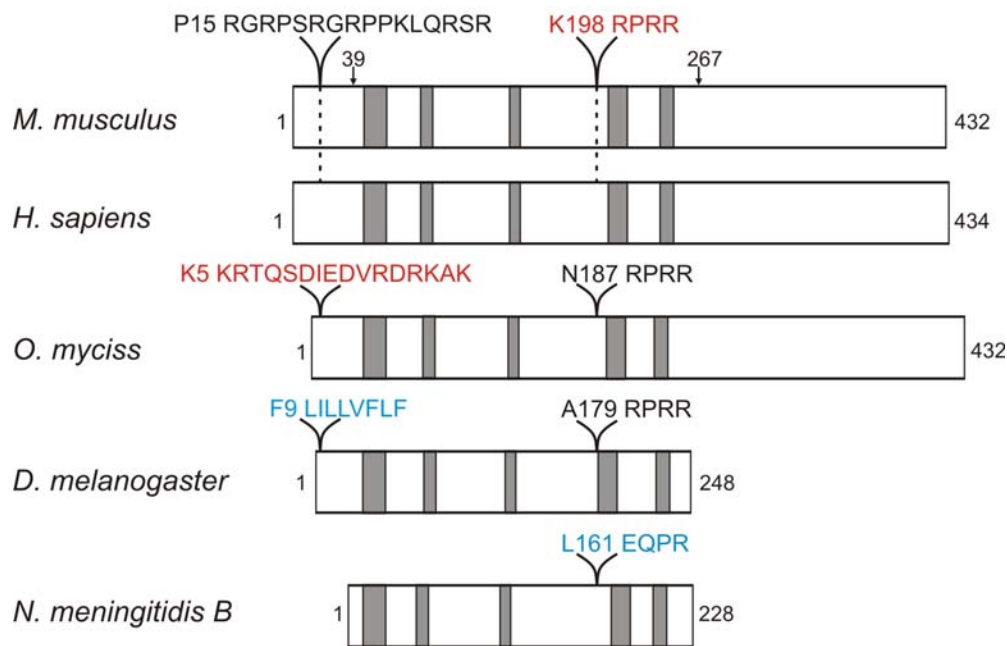
Fig. 12 Schematic representation of CMP-Sialic acid synthetases from *Mus musculus*, *Homo sapiens*, *Onkorhynchus mykiss*, *Drosophila melanogaster* and *Neisseria menigitidis* serogroup B (NmB). Highly conserved sequence motifs common to all CMP-sialic acid synthetases (CSS) are depicted as grey bars. The N-terminal domain of murine (CSS-NT), comprised of amino acids 39 to 267, is sufficient for Sia activation. The C-terminal domain (CSS-CT) of the murine enzyme is defined by amino acids 267 to 432 and of unknown function. C-terminal domains of human and trout are homologous to the murine enzyme. Position and sequence of basic amino acids stretches (black) of the murine and trout CSS are annotated above with functional nuclear localisation signals depicted in red. Equivalent sequences in *D.melanogaster* and NmB CSS are coloured in blue. Basic clusters of the human enzyme are identical to murine CSS.

Nuclear localisation of mammalian CSS was first described by Kean, 1969 and was further proven by series of studies carried out by preparation of subcellular fractions of various tissues and species in combination with detection of CSS activity (for reviews see Kean et al., 1991 and 2004). Immunofluorescent labelling of the recombinant murine, human and trout enzymes expressed in NIH 3T3 cells, Sf9 cells and rainbow trout gonadal tissue-derived RTG-2 cells, respectively confirmed nuclear sequestration (Münster et al., 1998; Lawrence et al., 2001; Tiralongo et al., 2007). Only proteins with a molecular weight of less than 40 kDa are able to enter the cell nucleus by passive diffusion. Translocation of larger proteins often depends on the existence of a nuclear localisation signal (NLS) sequence mediating nuclear import of the mature protein. Classical NLS are mono- or bipartite clusters of basic amino acids with the consensus sequences K-K/R-X-K/R-R and $(K/R)_2-X_{10-12}(K/R)_{3-5}$, respectively (Christophe et al., 2000). In the murine CSS (mCSS) two basic clusters were identified that were able to transport the enhanced green fluorescent protein (eGFP) to the nuclear compartment. However, only one (K^{198}RPRR) is necessary and sufficient to mediate nuclear import of murine CSS (Münster et al., 2002). Interestingly, this NLS was also found to be crucial for CSS activity (Münster et al., 2002). Data obtained from crystal structure analysis showed that the NLS (K^{198}RPRR) forms part of the active site (Krapp et al., 2003). In contrast, the basic cluster at the very N-terminus (P^{15}RGRPSRGRPPKLQRSR) of murine CSS is not sufficient to target the parent protein to the cell nucleus, when the NLS is deleted. Both basic clusters are annotated in figure 12, with the crucial NLS coloured in red. The overall sequence of human CSS is very similar to the murine one (93% identity at primary sequence level) and harbours basic clusters which are identical with the murine ones (Fig. 12). Consequently, recombinant human CSS is reported to reside inside the cell nucleus, though so far only investigated by overexpression in insect cells (Lawrence et al., 2001; Viswanathan et al., 2006). Rainbow trout (*Onkorhynchus mykiss*) CSS is also a nuclear protein, but its amino acid sequence at the equivalent position of murine NLS is slightly different (NRPRR instead of KRPRR, Fig.12). Studies carried out by Tiralongo et al. (2007) revealed that rainbow trout CSS uses a NLS different from the murine enzyme, defined by a stretch of basic amino acids at the very N-terminus (Fig 12, NLS (K^5KRTQSDIEDVRDRKAK, coloured in red). In contrast to the above described CSS from mouse, human and trout, the *Drosophila melanogaster* CSS sequence does not contain clusters of basic amino acids, which could serve as NLS (Fig 12). Concordantly the drosophila enzyme is reported to be a Golgi resident protein (Viswanathan et al., 2007).

Recently, two nuclear export signals (NES) were identified for the mouse enzyme, causing a small, but significant cytoplasmic population of overexpressed murine CSS in several cell lines. After NES deletion in murine CSS a reduced amount of cytoplasmic mCSS was found by indirect immunofluorescent analysis. However, these signal sequences seem to have a relatively weak effect on CSS localisation. Although the biological significance of NES, which are conserved in mammalian and fish CSS, but not in bacterial or insect CSS, is still unclear, their presence suggest a more complex regulation of CSS localisation than thought so far.

Interestingly, most bacterial orthologous, e.g. CSS from *Neisseria meningitidis*, as well as CSS from *Drosophila melanogaster* appear to be shorter and show similarity to only the N-terminal half (CSS-NT, about 260 residues) of the vertebrate enzymes (Fig 12). C-terminal domains of vertebrate CSS (CSS-CT) are highly conserved and comprise about 170 residues. It has been proposed (Krapp et al., 2003; Münster et al., 2004; Tiralongo et al., 2007), that vertebrate CSS-CTs fulfil a second yet unknown function, since they show homology to bacterial phosphatases of the *haloacid dehalogenase* (HAD)-superfamily. Notably, CSS from *Escherichia coli* and *Streptococcus agalactiae* have already been characterised as bifunctional enzymes, with the Sia activating function in the N-terminal domain and a C-terminal domain, which has been shown to deacetylate O-acetylated Sia monomers *in vitro* (Steenbergen et al., 2006; Lewis et al., 2007). However, no homology has been found between pro- and eukaryotic CSS-CT at primary sequence level.

Investigation of the intracellular localisation of endogenous CSS, so far was only possible by cell fractionation studies in combination with enzymatic activity assays (Kean et al., 1991 and 2004). Intracellular localisation studies at single cell level were performed by the use of recombinant epitope-tagged proteins. To further investigate the localisation of the vertebrate CSS it was essential to generate appropriate tools for direct illustration of the endogenous enzyme.

Aiming at generation of polyclonal antibodies directed against different domains of murine CSS three female New Zealand rabbits each were immunised with recombinantly expressed and purified CSS-NT (residues 39-267), CSS-CT (residues 267-432) or CSS (residues 39-267), respectively. In addition, for generation of monoclonal antibodies, immunisation of two New Zealand Black (NZB) mice was performed with full-length murine CSS. Antibodies were further characterized in terms of species specificity and applicability. Finally, nuclear sequestration of endogenous murine CSS was confirmed by different techniques.

2.2. Experimental Procedures

2.2.1. Materials

Strep-Tactin® was purchased from IBA GmbH (Göttingen, Germany). pET expression vectors and *Escherichia coli* BL21(DE3) were purchased from Novagen. α-Hsp70 and a-myc antibodies were purchased from Stressgen and Cell Signalling Technology, respectively. All other substrates for activity assays were purchased from Sigma.

For immunisation purpose, New Zealand rabbits (Charles River) New Zealand Black mice (NZB, Harlan Winkelmann, Paderborn) were used. For preparation of brain extracts and histological analyses wild type C57 Bl/6J mice were used.

Following plasmids were already available in the laboratory:

Mouse CSS-NT (residues 39-267) was encoded in pGEX-6P-1 (Pharmacia) driving the prokaryotic expression of N-terminal GST-tagged CSS-NT. Murine CSS-NT (residues 39-267) and full-length CSS (residues 39-432) and CSS from *drosophila melanogaster* (residues 1-248) were encoded in pET43-Strep, a modified pET43 (Novagen), driving the prokaryotic expression of N-terminal NusAStrepII-tagged proteins. Furthermore murine CSS-NT (residues 39-267) and full-length CSS (residues 39-432) as well as CSS from human (residues 1-434), and *neisseria meningitidis* (residues 1-228) were encoded in the vector pET22b-Strep, a modified pET22b vector (Novagen), for prokaryotic protein expression of N-terminal StrepII-tagged proteins. Rainbow trout CSS (residues 1-432) was encoded in pASK-IBA7 (IBA) driving the prokaryotic expression of N-terminal StrepII-tagged proteins.

For inclusion body generation of the full-length murine CSS (residues 1-432), the appropriate gene was present in vector pET3a (Novagen), driving prokaryotic expression of N-terminal T7/His-tagged proteins.

All prokaryotic expressed proteins carrying a N-terminal FLAG- and Myc-tag (EQKLISEEDL) were encoded in the vector pTrAM22, a modified pTrc99A (Amersham Biosciences), for prokaryotic protein expression; including all murine CSS (residues 1-432) with amino acid exchanges K198A/R201A, R202A, P200A, Q203A, respectively, the deletion construct lacking P196AKRPRR and truncated murine CSS comprised of residues 63-267 and 83-267.

2.2.2. Cloning of CSS-CT

CSS-CT cDNA encoding residues 267-432 was amplified from the plasmid pBSCMPrev using the primer pair MO 34 (5′-G<u>GGATCC</u>AAAGAGAAGC TGAAGGAGATAAAG -3′) / MO 35 (5′- AA<u>CTGCAG</u>CTATTTTTGGCATGAG TTATTAAC -3′). Because the primers contained *Bam*HI and *Pst*I restriction sites (underlined), products could be digested with *Bam*HI/*Pst*I and, after gel purification (Qiagen), ligated into the corresponding restriction sites of pET43-Strep. The resulting plasmid pET43-Strep-CMP267 drives the expression of N-terminally NusA-StrepII-tagged proteins in bacteria. Because of *Xho*I restriction site of the pET43-Strep multiple cloning site, the insert could be subcloned into *Bam*HI/ *Xho*I sites of the expression vector pET22b-Strep, a modified pET22b vector (Novagen) containing the sequence encoding an N-terminal Strep-tag II. The insert war further subcloned from pET43-Strep-CMP267 into *Not*I / *Bam*HI sites of pGEX-6P-1 (Pharmacia) for prokaryotic expression of CSS-CT with N-terminal GST-tag. cDNA for N-terminal extended CSS-CT (residues 234-432) was amplified from the plasmid pBSCMPrev using the primer pair MO 33 (5′- CG<u>GGATCC</u>GCATATTATGAAATGCGAGCTG-3′) / MO 35 (5′- AA<u>CTGCAG</u>CTATTTTTGGCATGAGTTATTAAC -3′). Primers contained *Bam*HI and *Pst*I restriction sites (underlined) and were cloned in pET43-Strep (Novagen) as described above.

The sequence identity of all PCR products was confirmed by sequencing.

2.2.3. Protein expression and purification via StrepII-tag

All proteins encoded in pET expression vectors, carrying a StrepII-tag at the N-terminus were expressed and purified according to the following protocol. This concerned all Strep-tagged proteins, with the exception of rtCSS encoded in pIBA7-Kdn, which is described below.

Recombinant CSS-constructs carrying a StrepII–tag (IBA) at the N-terminus were overexpressed in *E. coli* BL21 (DE3) and cultivated in PowerBroth (Athena Enzyme Systems) at 15°C. At an optical density (A600) of 1.2, expression of the recombinant protein was induced by adding 0.1 mM isopropyl-1-thio-β-D-galactopyranoside, and bacteria were harvested 24 h after induction. Thereafter, cells were harvested by centrifugation (4000 x g, 15 min, 4°C), and pellets were washed once with phosphate-buffered saline (PBS) and stored at -20°C until needed. The thawed pellet (3 mg wet weight) was resuspended in 25 mL buffer W (100 mM Tris-HCl, pH 8, 150 mM NaCl) and incubated for 1 hour on ice in the presence of a protease inhibitor cocktail (#P8465, Sigma-Aldrich), 2.5 mg lysozyme, 1 mM DTT and 0,1 % TritonX-

100. The suspension was then sonicated (Branson Sonifier) for 4 min at 0°C. Non-lysed cells and membrane constituents were removed by centrifugation at 12 000 x g. The supernatant was incubated with 1 mg avidin for 30 min at 4°C and aggregates were removed by centrifugation at 12 000 x g. The proteins were purified immediately by affinity chromatography utilising the N-terminal StrepII–tag (IBA) according to manufacturers' instructions. Buffer exchange was performed on a HiPrep 26/10 desalting column (Amersham Biosciences). Fractions containing the fusion protein were pooled and finally concentrated to 1-2 mg/ml using Amicon Ultra centrifugal devices (Millipore).

Expression of rainbow trout CSS

Recombinant rtCSS (residues 1-432) carrying a StrepII–tag (IBA) at the N-terminus was overexpressed in *E. coli* BL21, and cells were grown in Luria-Bertani (LB) medium with 200 µg/ml carbenicillin at 15°C to an A600 nm of 0.6. Then protein expression was induced by the addition of 20 mg anhydrotetracyclin per litre culture medium, and the culture continued for 50 h. Harvest, lysis and protein purification was done as described above for other StrepII-tagged proteins.

2.2.4. Purification via GST tag

Protein expression and cell lysis was performed as described for Strep tagged rainbow trout CSS, but expression of the protein was induced by adding 0.1 mM isopropyl-1-thio-β-D-galactopyranoside, and bacteria were harvested 24 h after induction. Additionally no incubation with avidin was performed.

The resulting supernatant was applied to a GST-Sepharose column (1 ml), equilibrated with buffer A (PBS, 1 mM DTT). Unbound material was washed off with ten column volumes of buffer A and the GST fusion protein was eluted with buffer A containing 10 mM reduced glutathione. Fractions containing the GST fusion protein were pooled and finally concentrated to 1 mg*ml−1 using Amicon Ultra centrifugal devices (Millipore).

2.2.5. Separation of soluble and insoluble fractions

Lysis of bacterial pellets was performed with BugBuster™ Protein Extraction Kit (Novagen) according to manufacturer recommendation. Soluble and insoluble fractions were obtained by centrifugation (16 000 x g, 20 min, 4°C), with the resulting pellet representing the insoluble faction.

2.2.6. Inclusion body (IB) preparation and solubilisation

Bacteria from 500 ml LB culture cells were harvested (6000 x g, 15 min), washed with PBS and the pellet was resuspended in 25 ml buffer I (50 mM Tris pH 8.0, 2 mM EDTA) containing 250 µl protease inhibitor cocktail (#P8465, Sigma-Aldrich) and 5 mg lysozyme. After incubation at 30°C for 15 min bacteria were lysed by sonification with a microtip (Branson Sonifer, 50% duty cycle, output control 5) using five pulses of 1 min intermitted by breaks of 30 s. Non-lysed cells and membrane constituents were removed by centrifugation at 15 000 x g. The pellet was resuspended in 30 ml buffer I containing 0.5% LDAO and 250 µl protease inhibitor cocktail followed by homogenisation in a dounce homogeniser and centrifugation (15 000 x g) in four subsequent cycles (4°C). Thereafter three cycles of redispersion in 30 ml buffer J (20 mM Tris pH 8.0, 300 mM NaCl) containing 250 µl protease inhibitor cocktail using a dounce homogeniser followed by centrifugation were performed. Finally the pellet was washed for three times in buffer J without protease inhibitor cocktail using dounce homogeniser and subsequent centrifugation.

The resulting pellet was stored in aliquots corresponding to 100 ml LB culture at -20°C until needed.

For immunisation IB were further purified by resuspendion of 1 mg IB in 70 µl PBS and 80 µl 2x Laemmli using an Eppendorf dounce homogeniser and boiling for 10 min, followed by separation on preparative 10% SDS-PAGE. The gel band carrying the recombinant protein was excised and chopped into small pieces and the protein was eluted with 1 ml PBS by incubation at RT for 4h, followed by subsequent buffer exchanges with reducing volumes for at least 4h. The excised region was controlled by colloidal Coomassie blue staining (Roti-Blue, Roth) of a small gel stripe to visualize position of protein bands.

2.2.7. Immunisation of rabbits

Antisera against murine CSS were prepared by subcutaneous injection of nine female New Zealand rabbits (Charles River) with 500 µg of purified recombinant protein, corresponding to the full length, N-terminus or C-terminus of murine CSS.

Solubilised inclusion bodies of CSS-NT fused to a GST epitope tag were used for immunisation of three rabbits #62, #69 and #82. Recombinantly expressed and affinity purified CSS-CT, carrying a N-terminal NusA-StrepII-tag was applied to rabbits #58, #59 and #86. Rabbits #51, #56 and #61 were immunised with solubilised inclusion bodies of T7/His-tagged full length CSS (see also Table 5). After preparation of pre-immune serum, the priming was done by subcutaneous injection

of 500 µg antigen mixed with complete Freund's adjuvant (Difco) in the back of the rabbits. After three weeks, the first boost with 500 µg antigen and incomplete Freund's adjuvant (Difco) was performed. In the following, every sixth week a further booster injection was done. 37 weeks after the primary immunisation, which is equivalent to 7 boosts, the final bleedings were performed. Blood was collected 10 days after the last injection. The prepared antisera specifically recognize both recombinant and native forms murine CSS.

2.2.8. Serum preparation

The blood was incubated at RT for 4 hours followed by incubation over night at 4°C to allow agglutination. The serum was obtained by centrifugation of the clocked blood at 3 600 x g for 15 min at 10°C. The antibody titer of the serum was tested in ELISA and the serum stored at -20°C.

2.2.9. Serum purification

Total lysates of mouse embryo fibroblasts were separated by 10% SDS-PAGE and blotted onto nitrocellulose. After saturation of Western blottings with blocking buffer (5% milk, 0.2%Tween in PBS) polyclonal sera were applied for 1 h at RT. Thus unspecific elements were removed from polyclonal sera. The resulting purified sera were stored in 100 µl aliquots at -20°C until needed.

2.2.10.   Immunisation of mice

Two nine-week-old female New Zealand Black mice (NZB, Harlan Winkelmann, Paderborn) were intraperitoneally immunised with solubilised inclusion bodies of recombinant T7-His-tagged murine CSS (residues 1-432). 100 µg antigen with complete Freund's adjuvant (cFA, Difco) were initially injected. For booster injections 50 µg antigen with incomplete Freund's adjuvant (iFA) were applied. Three consecutive days directly before spleen preparation, 40 µg antigen was applied without adjuvant. The mice were splenectomised on day 60 and 67, respectively. All experiments were carried out with an approval according to the animal protection law. Table 4 shows the immunization pattern.

Table 4. – Immunisation scheme of NZB mice

Day		Injection	Adjuvant
mouse 1	mouse 2		
0	0	Priming	cFA[a]
19	19	Boost	iFA
35	35	Boost	iFA[b]
40	40	Boost	iFA
47	47	Boost	iFA
54	54	Boost	iFA
	61	Boost	iFA
57	64	Boost	PBS
58	65	Boost	PBS
59	66	Boost	PBS
60	67	Fusion	

[a] complete Freund's Adjuvant; [b] incomplete Freund's Adjuvant

2.2.11. Cell culture

Mammalian cells were maintained in a humidified 7.5% CO_2 atmosphere at 37°C.

CHO C6 cells (derived from chinese hamster ovary cells) were cultured in α-MEM (Invitrogen) supplemented with 10% fetal calf serum and 2 mM L-glutamine (Seromed). Hybridomas were cultured in Dulbecco's modified Eagle's medium (DMEM) (Biochrom) containing 20% foetal bovine serum (Invitrogen), 2 mM L-glutamine (Seromed), penicillin-streptomycin (Sigma), HAT-supplement (Gibco) and 50 U/ml IL-6 (Boehringer). Myeloma cells X63-Ag8.653 were kept in DMEM containing 5% FCS, Pen/Strep and sodium pyruvat (1mM).

Stably transfected CHO C6 cells, expressing the murine CSS with N-terminal Flag/Myc-tag, and mouse embryo fibroblasts (MEFs) were already available in the laboratory.

2.2.12. Fusion of myeloma and B-cells cells

Cell fusion and hybridoma cloning were performed using Galfre's method (1981) with some modifications. Splenocytes from the immunised mice were subjected to cell fusion with myeloma cells X63-Ag8.653 (Kearney et al., 1979) using 50% (w/v) polyethylene glycol (PEG) 1500 in 75 mM Hepes buffer (Boehringer 783641) in a ratio of 800 µl PEG solution for 50 million spleen cells. Hybridomas were cultured for two weeks in Dulbecco's modified Eagle's medium containing 20% foetal bovine serum (Invitrogen), 2 mM L-glutamine (Seromed), penicillin-streptomycin (Sigma), HAT-supplement (Gibco) and 50 U/ml IL-6 (Boehringer). Antibody-secreting

hybridomas were cloned by the limited dilution method. Three weeks after fusion, HAT-supplement was replaced by HT (Gibco).

2.2.13. Determination of antibody titre in rabbit sera and hybridoma supernatants

96-well microtiter plates (U-form, Greiner) were coated with 100 ng per well recombinantly expressed and purified murine CSS at RT for 1 hour. As negative control for hybridoma supernatant analysis, plates were coated with bovine serum albumin (BSA, Applichem). The plates were washed twice with PBS and blocked with 150 ml 1% BSA/PBS over night. After two washing steps, residual buffer was removed and the plates used directly. Coated microtiterplates were incubated with serial serum dilutions (1:200 to 1: 64 000 in 1% BSA/PBS) or pure hybridoma cell culture supernatant. As a positive control anti-epitope monoclonal antibody (mAB) and as a negative control serial dilutions of the preimmune sera were used. After 1 hour at RT, the plates were washed 3 times and incubated with 10 µl/well secondary antibody dilution (α-rabbit-pox 1:2000, α-mouse-pox 1:2000) in 1% BSA/PBS for 1 hour at RT. The plates were washed three times with PBS and signals were detected with the substrate 3,3',5,5'-Tetramethylbenzidin (TMB) according to manufacturer instruction (Sigma). The samples were analysed at 450 nm using the ELISA-Reader 'DigiScan' (Asys Hitech)

To exclude false positive signals obtained by antibodies directed against epitope-tags differently tagged proteins were used for immunisation and titer monitoring. Rabbits immunised with the GST-tagged N-terminus of CSS were analysed with NusA-StrepII-CSS-NT. Anti StrepII-CSS-CT sera were characterised with GST-CSS-CT, and sera derived from immunisation with the full length T7/His-tagged CSS and hybridoma cell supernatants were analysed by the use of NusA-StrepII tagged full-length CSS, as well as the separate domains.

2.2.14. Subtype determination of mAB LF6

Subtype determination of mAB LF6, specific for murine CSS, was performed by Mouse IgG ELISA Kit (Roche) according to manufacturer's instruction. In brief, microtiter plates were coated with CSS antigen, as described above for titre monitoring. As primary antibody hybridoma cell culture supernatant was used. Biotinylated secondary antibodies specific for IgG (all subtypes), IgM, IgG1, IgG2a and IgG2b, respectively were used and signals were detected with an avidin-alkaline phosphatase detection system.

2.2.15.   Purification of mAB LF6 via protein-A sepharose

mAB LF6 from cell culture supernatants were adjusted to pH 7-8 and filtered (Sartorius Minisart 0.45 µm). The resulting supernatant was applied over night to a 1 ml protein A-sepharose column (HiTrap rProtein A Fast Flow, Amersham-Pharmacia) equilibrated with buffer E1 (20 mM sodium phosphate pH7). Unbound material was washed off with 20 column volumes of buffer E1, and LF6 was eluted with buffer E2 (100 mM sodium citrate pH 4) and immediately supplemented with sodium chloride (f.c.200 mM) and neutralised by stepwise addition of 1M Tris buffer, pH9. Fractions containing mAB LF6 were pooled and IgG concentration was determined using BCA protein assay reagent (Pierce). Purification of one litre cell culture supernatant yielded in about 17 mg mAB LF6.

2.2.16.   Preparation of nuclear and cytoplasmic extracts

One frozen mouse brain (post natal day 1) or cell pellet (100 µl cell volume) was lysed in 200 µl hypotonic buffer (10 mM hepes pH7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT, 0.5 mM PMSF, 0.5 µg/ml aprotinin) and homogenised using Eppendorf dounce homogeniser. After 15 min on ice, suspension was homogenised by passing 10 to 15 times through a cannula of 0.4 mm in diameter. After centrifugation at 3300 x g for 5 min at 4°C, supernatant represents the cytoplasmic extract. Pellet was resuspended in 0.5 pellet volumes (40 - 60 µl) low salt buffer (20 mM hepes pH7.9, 20 mM KCl, 1.5 mM MgCl2, 0.2 mM EDTA, 1 mM DTT, 0.5 mM PMSF, 0.5 µg/ml Aprotinin, 20% Glycerol) followed by slow addition of the same volume of high salt buffer (identical with low salt buffer, but 1.2 M KCl). After incubation at 4°C for 30 min and centrifugation (3300 x g for 15 min at 4°C), the resulting supernatants represent nuclear extracts.

2.2.17.   SDS-PAGE analysis and immunoblotting

SDS-PAGE was performed according to Laemmli (1970). Protein samples were separated on SDS-polyacrylamide gels composed of a 3% stacking gel and a 10% separating gel (if not differently denoted). For Western blot analysis, proteins were transferred to nitrocellulose membranes (Schleicher & Schüll). Western blots were developed using 5 µg/ml of primary antibody followed by the anti-mouse or anti-rabbit alkaline phosphatase conjugate (Dianova). Nitro blue tetrazolium (NBT) and 5-bromo-4-chloro-3-indoyl phosphate (BCIP) were used as substrates for alkaline phosphatase.

When peroxidase (pox) -coupled secondary antibodies were used, the ECL system was used. Western blottings were incubated for about 5 min in ECL detection solution composed of Luminol enhancer solution (250 mM Luminol, 90 mM p-coumaric acid, 1 M Tris pH 8.5) and peroxide solution (1M Tris pH 8.5, 30% H_2O_2) in a 1:1 ratio. Signals were displayed by exposed for 1-10 min to a film (Hyperfilm MP).

2.2.18. Histology

Postnatal mice (P0.5-2.5) were anesthetised by hypothermia and killed by decapitation. Brain, liver and kidney were removed and fixed by immersion in 4% paraformaldehyde/PBS over night. After dehydration and embedding in paraffin tissues were sectioned at 3 µm. For standard histological examination, sections were air dried at 80°C for 1h, deparaffinised and stained with polyclonal serum #59 (1:4000) followed by anti-rabbit IgG-biotin conjugate and incubation with ABC-complex (Vectastain Elite ABC reagent Vector Laboratories, Inc.). Sections were dehydrated and mounted by the entellan (Merck). Micrographs were archived digitally using AxioVision version 3.1 software (Zeiss).

2.2.19. Identification of 70 kDa protein by 2D PAGE and MALDI-TOF MS

Sample Preparation

Cytoplasmic extracts were prepared of two wild type post natal day 0.5 mouse brains (C57 Bl/6J). Proteins were precipitated with TCA (10 to 20 %) over night and pelleted by centrifugation (15,000 x g, 30 min). The protein pellet was washed with acetone and solubilised in lysis buffer (30 mM Tris-HCl pH 8.0, 7 M urea [Roth, Karlsruhe,Germany], 2 M thiourea [Sigma], and 4% [w/v] CHAPS [Roth]) to a protein concentration of 5 mg/ml. Insoluble material was removed by centrifugation at 16,000 x g for 5 min in a benchtop centrifuge.

2D dimensional SDS-PAGE and MS analysis were done by Dr. Büttner in the department of microbiology at TiHo Hannover in the group of Prof. Gerlach.

First Dimension

After rehydration of Immobiline DryStrips (24 cm, pI 3-11, non linear; Amersham) with 450 µl of rehydration buffer per strip containing 2 M thiourea (Sigma), 7 M urea (Roth), 4% (w/v) CHAPS (Roth), 1% (v/v) IPG buffer (Amersham), and 0.2% (w/v) dithiothreitol (Roth) for 12-18 h samples were loaded and subsequently focused using an EttanTM IPGphorTM (Amersham) for 21 h in the following series of time

blocks with increasing voltage: 3 h at 150 V, 3 h at 300 V, 6 h at a 1000 V gradient, 3 h at an 8000 V gradient, and 6 h at 8000 V. For second dimension, the strips were equilibrated twice for 15 min by rocking in a solution of 100 mM Tris-HCl (pH 8.0), 6 M urea (Roth), 30% (v/v) glycerol (Roth), and 2% (w/v) sodium dodecyl sulfate (SDS; Serva, Heidelberg, Germany) supplemented with 0.5% (w/v) DTT (Roth) for the first equilibration step and 4.5% (w/v) iodoacetamide (Sigma) for the second equilibration step.

Second Dimension

The Immobiline DryStrips (Amersham) were placed on the top of 12.5% polyacrylamide gels containing 12.5% (v/v) acrylamide (Serva) and 0.33% bisacrylamide (Serva). The second-dimension SDS-PAGE was carried out at 12°C and 50 V for 3 h followed by 100 V for 15 h using the EttanTM DALTsix Electrophoresis System (Amersham).

In order to pick spots that were identified by Western blotting, the 2D-PAGE was directly stained with colloidal Coomassie blue G-250 (Candiano et al., 2004). Spots of interest were excised; the protein was trypsinated and removed from the gel according to the method of Wilm and Mann (Shevchenko et al., 1996). The solution containing the extracted peptides was completely evaporated in a vacuum centrifuge. Peptides were redissolved in 3–5 µl of 50% (v/v) ACN (Merck), 0.1% (v/v) formic acid (Merck).

Protein Identification by MALDI-TOF MS

Matrix Assisted Laser Desorption / Ionization Time-of-Flight (MALDI-TOF) mass spectrometry (MS) was carried out on a Voyager-DETM Pro (Applied Biosystems, Forster City, USA); 1 µl of the peptide solution was mixed with 1 µl of matrix (α-cyano-4-hydroxy-cinnamic acid [ACHC, Bruker Daltonics, Billerica, USA], 5 mg/ml in 50% ACN with 0.1% trifluoroacetic acid [TFA]) and then spotted on the target plate. Peptide spectra were acquired in positive reflection mode, averaging about 1000 laser shots per MALDI-TOF spectrum. The mass spectra were analysed using the Applied Biosystems Data Explorer® V 4.8 software and peptide mass lists were generated. Peptide masses were applied to the peptide mass fingerprint algorithm on the Mascot web site (http://www.matrixscience.com). The search algorithm was set to allow carbamidomethylation on cysteine residues, oxidation on methionine residues and a maximum of 1 missed cleavage. The peptide mass tolerance was set 0.2 Da. proteins were considered as identified when the probability-based score was above the significance threshold ($p < 0.05$). Identification was confirmed by comparing the calculated molecular mass and isoelectric point values from the identified proteins with the observed values on the 2D gel.

2.3. Results

The CMP-sialic acid synthetase (CSS) of all vertebrate species analysed so far has predominantly been found in the cell nucleus. However, nuclear localisation is not a prerequisite for enzymatic activity as demonstrated at cellular level for the recombinant murine enzyme (Münster et al., 2002). In the past, several unsuccessful attempts to obtain mCSS specific antibodies had been carried out in cooperation with various companies. The following work focuses on generation and characterisation of poly- as well as monoclonal antibodies specific for murine CSS. In this context, subcellular distribution of the endogenous mCSS has been analysed in different experimental setups.

2.3.1. Generation of mono- and polyclonal antibodies directed against murine CMP-sialic acid synthetase

With the aim to generate antibodies specific for the single domains or mCSS, three different constructs were used for immunisation of three female New Zealand rabbits each: the N-terminal catalytic active domain (CSS-NT, residues 39-267), the C-terminal domain (CSS-CT, residues 267-432) and the full length CSS (residues 1-432), respectively (Fig. 12 and Table 5). Table 5 gives an overview about the used antigens, their epitope-tags and calculated molecular weights. Furthermore the antigen application forms as well as the 'names' of the immunised rabbits (rabbit identifier) are summarised. Titre development was observed over the whole immunisation period in an ELISA with serum samples taken 10 days after each boost. 37 weeks after the primary immunisation, which is equivalent to 7 boosts, the final bleedings were performed. Using ELISA plates coated with purified recombinant CSS, a specific signal could be detected in sera of all nine rabbits with dilutions ranging from 1: 8 000 (anti-CSS-CT sera of rabbits #58, #59, #86) to 1: 16 000 (anti-CSS-NT sera #62, #69, #82 and anti-CSS sera #51, #56, #61). To exclude false positive signals obtained by antibodies directed against epitope-tags differently tagged proteins were used for immunisation and titer monitoring. Details are described in 'experimental procedures' (2.2.13).

Table 5. Antigens used for generation of polyclonal antibodies

Antigen	Residues	Epitope-tag	kDa all (tag+CSS)	Antigen Status	Rabbit Identifier
CSS-NT	39-267	GST	53 (27+26)	IB, solubilised	#62, #69, #82
CSS-CT	267-432	NusA-Strep	76 (56+20)	affinity purified	#58, #59, #86
CSS	1-432	T7/His	48 (0+48)	IB, solubilised	#51, #56, #61

IB: inclusion bodies

The aptitude of all sera for application in Western Blot analysis was proven by incubation of nitrocellulose stripes carrying recombinantly expressed and purified CSS with serial dilutions of each serum. Exemplarily, titrations with the CSS-NT specific serum #62, the CSS-CT specific serum #86 and the full-length serum #61 are shown in Figure 13A. Anti-CSS-NT sera could be diluted up to 1:1000, anti-CSS-CT sera to 1: 4000 and sera directed against full length CSS 1:1500 (Fig 13A).
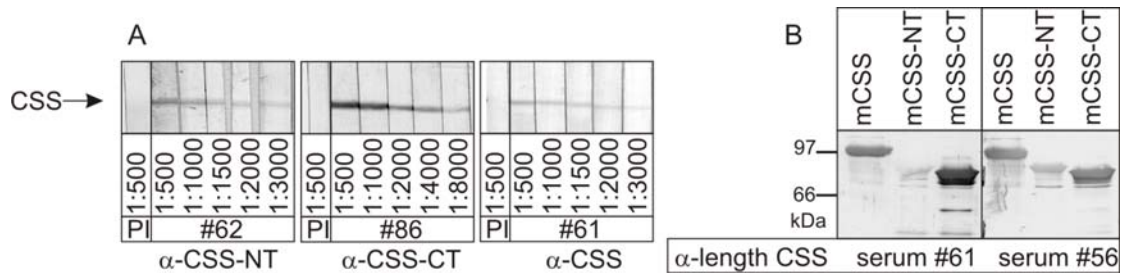
Fig. 13 (A) Titration of polyclonal antibodies for recombinant mCSS in Western Blot analysis. For titration of anti-CSS-NT and anti-CSS sera 5 µg/lane purified NusA-StrepII-tagged CSS (aa 39-432), and for titration of anti-CSS-CT sera 5 µg/ml purified GST-tagged CSS-CT was separated by SDS-PAGE and blotted onto nitrocellulose. Western blots were developed with serial dilutions of the generated polyclonal sera (exemplarily shown for α-CSS-NT #62 (left), α-CSS-CT #86 (middle) and α-CSS #61 (right)), and AP-coupled secondary antibody using BCIP/NBT as substrate. Preimmunsera (PI) were used as negative control. The situation after the fifth of seven boots is presented.

(B) Serum #56 and #61 generated against full length mCSS recognize the entire protein as well as separate domains (CSS-NT (residues 39-267) and CSS-CT (residues, 267-432)). 1.5 µg recombinantly expressed and purified mCSS was separated by 10% SDS-PAGE and blotted onto nitrocellulose. All proteins, full-length (CSS, aa 39-432, Mcalc: 44 kDa), N-terminus (CSS-NT, aa 39-267, Mcalc: 26 kDa) and C-terminus (CSS-CT, aa267-432, Mcalc: 18 kDa) carried a NusA-StrepII epitope tag with the calculated molecular weight of 56 kDa, resulting in expected molecular weights of 100, 82, and 74 kDa for CSS, CSS-NT and CSS-CT, respectively. The Western blots were developed with polyclonal sera #61 and #56, generated against full length mCSS, respectively (1:1000) and AP-coupled secondary antibody using BCIP/NBT as substrate. In parallel approaches staining of StrepII-epitope tags was performed (data not shown).

Staining with preimmuneserum was used as negative control. All nine sera recognize recombinantly expressed and purified mCSS in Western Blot analysis. Furthermore,

sera #56 and #61, both immunised with T7/His-tagged full length CSS, were able to detect the separate N-terminal (CSS-NT) and C-terminal (CSS-CT) domains of the murine protein (Fig 13B).

For the intricate purpose of generating an immune response in mice against a mouse protein, the New Zealand Black (NZB) strain was used. These autoimmune animals exhibit hyperactivity of the B- and T-cell system and have been shown to give positive responses to weak antigens (Frosch et al., 1985). Two female NZB mice were immunised by intraperitoneal injection of solubilised inclusion bodies of recombinantly expressed T7/His-tagged full length mCSS. For priming 100 µg antigen was injected with complete Freund's adjuvant (cFA). Up to seven booster injections were applied every second week containing 50 µg antigen in combination with incomplete Freund's adjuvant (iFA). At three consecutive days directly before spleen preparation, antigen was applied in PBS. Respectively, 60 and 67 days after priming the two mice were splenectomised and spleen cells were fused to X63-Ag8.653 myeloma cells (Kearney et al., 1979). Thereafter, hybridoma cells were cloned by limiting dilution. From a total of 654 cell hybridoma supernatants screened by indirect ELISA with murine CSS (details in 'experimental procedures', 2.2.14.), 17 hybridoma clones gave positive signals. Signal intensity of positive clones ranged from two to ten fold above background signal, defined by binding to BSA. However, only the clone with the best signal to noise ratio, recognized purified mCSS in Western blot analysis (Fig 14). In the following, this work will focus on characterisation of this hybridoma clone, named LF6. The isotype was determined by the use of Mouse IgG ELISA Kit (Roche), revealing predominantly the IgG2b epitope, but also significant amounts of IgG2a epitope. The presence of both epitopes remained also after extensive subcloning. Due to good affinity to protein A, LF6 was purified via protein A-Sepharose columns (HiTrap rProtein A Fast Flow, Amersham-Pharmacia) with a yield of 17 mg LF6 per litre cell culture supernatant (details in 'experimental procedures', 2.2.16.).
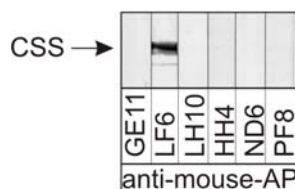
Fig. 14 Western blot analysis of hybridoma supernatants. 5 µg/lane purified NusA-Strep tagged full length mCSS (aa 39-432), was analysed by Western blotting, developed with cell culture supernatants of hybridoma clones GE11, LF6, LH10, HH4, ND6 and PF8, respectively, and AP-coupled secondary antibody using BCIP/NBT as substrate.

2.3.2. Species specificity of polyclonal antibodies and mAB LF6 directed against murine CMP-sialic acid synthetase

All known CSS, from bacteria to man, contain highly conserved sequence motifs, which are in case of the murine enzyme restricted to the N-terminal domain. Thus, the sera directed against CSS-NT (#62, #69, #82) and the full-length protein (#51, #56, #61), respectively, were tested for species specificity to possibly extend the application to CSS in other organisms. For this purpose recognition of affinity purified human, rainbow trout, *Neisseria meningitidis* and *Drosophila melanogaster* CSS, carrying N-terminal StrepII-tags, were analysed by Western blotting. Primary sequence of the human enzyme is to 93% identical with the murine one. Reflecting this close relation all sera detected the StrepII-tagged human enzyme (hCSS, M_{calc} 50 kDa) in Western Blot analysis as shown in figure 15A. NusA-StrepII-tagged full-length mCSS (M_{calc} 104 kDa) served as positive control. Sera directed against mCSS-NT, #62, #69 and #82, were all able to detect hCSS. (1:1000 dilutions). Because of high background signals, serum #51 is just of limited use. Furthermore, anti-mCSS sera #56 and #61 were tested for their specificity to the trout enzyme. Primary sequence comparison revealed 57% identity and 71% similarity between mouse and trout CSS. Notably, only serum #56 was able to detect the CSS derived from trout (Fig 15B). In addition, although the enzyme from *Drosophila melanogaster* is just to 36% identical with the mCSS, all anti-CSS-NT sera (#62, # 69 and # 82) were able to specifically detect the purified drosCSS (Fig 15C), whereas anti-mCSS sera #56 and #61 were not able to detect drosCSS (data not shown). Futhermore, anti-CSS-NT sera (#62, # 69 and # 82) were tested for their ability to detect *Neisseria meningitidis* serogroup B (NmB) CSS, which shows just 28% identity to the murine enzyme at primary sequence level. However, the conserved motifs are present in the bacterial CSS and serum #82 gave a strong signal for the NmB enzyme. A faint band for NmB CSS was detected with serum #69; however, #62 did not recognize the bacterial enzyme (Fig 15D).

In addition, species specificity of monoclonal antibody (mAB) LF6 directed against murine CSS was analysed with purified recombinant StrepII-tagged CSS from human, rainbow trout and *Nm*B. Interestingly, as shown in figure 15E LF6 recognizes the almost identical human enzyme (hCSS), whereas figure 15F illustrates that CSS deriving from fish (rtCSS) or bacteria (NmB-CSS), respectively, are not detected. Purified recombinant StrepII-tagged murine CSS served as positive control. For control reasons an equivalent Western blot was developed with Strep-Tactin-AP conjugate (iba; Fig 15F, left).
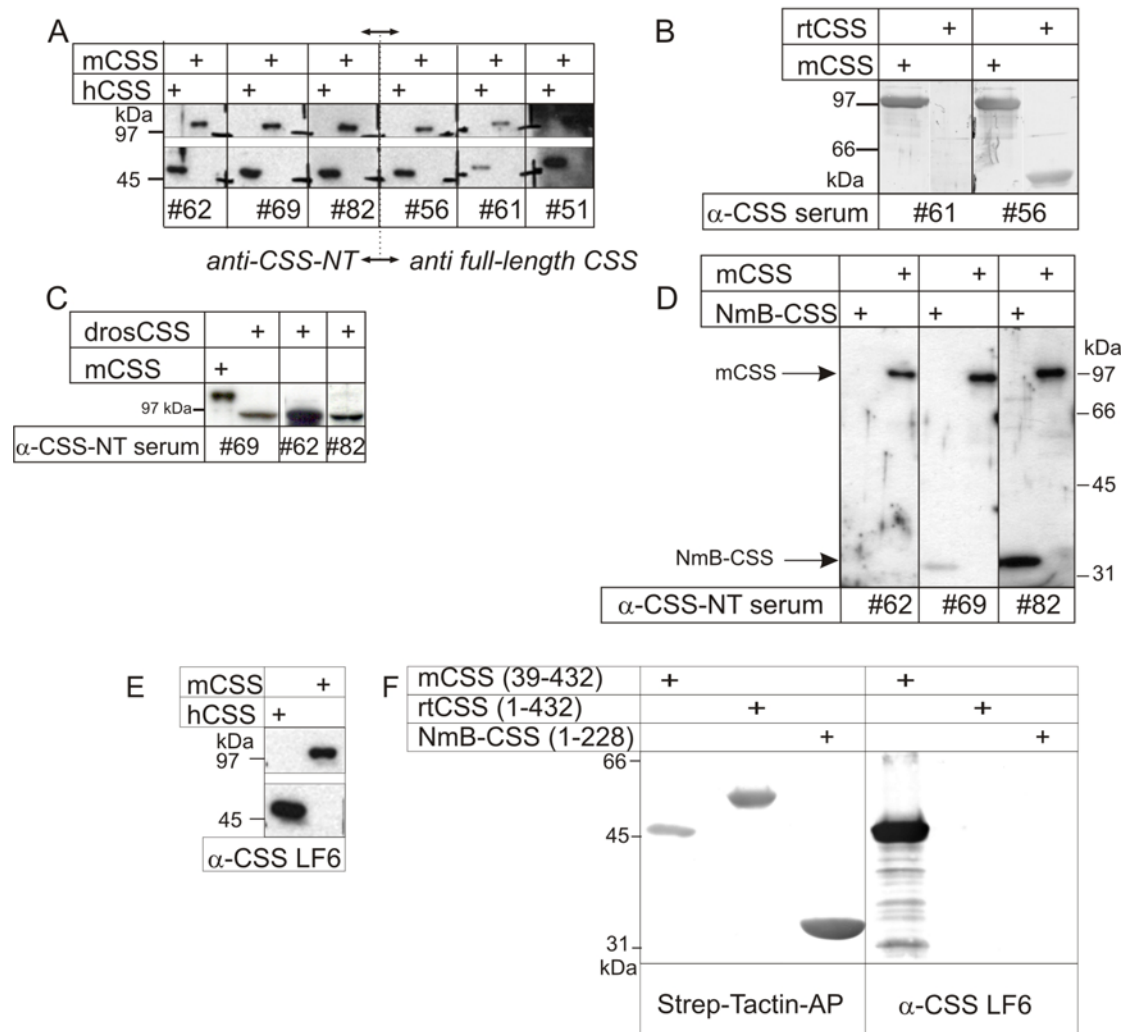
Fig 15. Species specificity of mono- and polyclonal antibodies generated against mCSS analysed by Western Blot analysis. Purified (A) human (100 ng/lane), (C) drosophila (100 ng/lane), (D) NmB (100 ng/lane) and mouse CSS (50 ng per lane) were separated by 10% SDS-Page and blotted onto nitrocellulose. (A) Western Blots development was performed with polyclonal antibodies (1:1000) raised against CSS-NT (#62, #69, #82) or full-length CSS (#56, #61, #51), peroxidase-conjugated secondary antibody and ECL detection. (B) 1.5 µg/lane of recombinantly expressed and purified murine (mCSS) and rainbow trout (rtCSS) CSS were separated by 10% SDS-Page and blotted onto nitrocellulose. Western Blot development was performed with polyclonal antibodiy #61 and #56 (anti-CSS, 1:1000), AP-conjugated secondary antibody and BCIP/NBT detection. Full-length mCSS (aa 39-432, M_{calc}: 44 kDa) carried a NusA-StrepII epitope tag of the calculated molecular weight of 56 kDa. Full length rtCSS (aa1-432, M_{calc}: 48 kDa) carried a StrepII-Tag at the N-terminus. In parallel approaches staining of StrepII-epitope tags was performed (data not shown). (E, F) mAB LF6 specifically binds to murine and human CSS; CSS from rainbow trout (rt) and *Neisseria menigitidis* B (NmB) are not detected. (E) Western blot analysis of purified murine and human CSS was done with mAB LF6 and ECL detection. (F) 2 µg of affinity purified StrepII tagged mCSS, rtCSS, and NmB-CSS, respectively, were separated by 12% SDS-Page and blotted onto nitrocellulose. The Western blotting was developed with mAB LF6 and AP-coupled secondary antibody using BCIP/NBT as substrate (right side). StrepTactin-AP conjugate was used as positive control for development of an identical approach (left side).

2.3.3. Nuclear sequestration of endogenous murine CSS

Specificity of the generated antibodies was analysed with whole bacterial and whole CHO C6 cell lysates overexpressing full-length mCSS by Western blotting. However, all polyclonal sera gave strong background signals and had to be purified by retroblotting. Therefore, whole cell lysates from mouse embryo fibroblasts (MEF), in which CSS expression ranges below the detection limit (see Fig 16D), were separated by SDS-PAGE and blotted onto nitrocellulose membranes. Sera were purified by incubation with the MEF-blot membranes and could further be used in Western blot analysis with dilutions ranging from 1:500 to 1:1000. However, as expected, the background on bacterial lysates remained.

CSS expression was analysed in cytoplasmic (CE) as well as nuclear extracts (NE) of (i) native *Chinese hamster ovary* cells (CHO C6, wt) and (ii) CHO C6 cells stably transfected with N-terminally Flag-tagged murine CSS (+CSS) by Western blotting (Fig. 16A). Sera directed against CSS-NT (#62, #69, #82) detect endogenous hamster CSS (Mr probably about 48 kDa) in nuclear extracts (NE) of CHO C6 cells (wt). However, so far no sequence information is available for hamster CSS, but due to high homology within mammalian CSS, the observed Mr of about 48 kDa is in perfect agreement with the calculated Mr of e.g. mCSS (Mr 48 kDa). In stably transfected CHO cells (+CSS), Flag-tagged mCSS (Mr 50 kDa) is detected in addition to the endogenous enzyme. The corresponding region is marked by a dotted frame in figure 16A. Expression of Flag-tagged mCSS was confirmed with anti-Flag mAB M5 (data not shown). Sera directed against mCSS-CT (#58, #59, #86) show relatively high background staining, but also detect endogenous as well as Flag-tagged mCSS in nuclear extracts. However, sera #59 and #86 additionally detect low amounts of CSS in the cytoplasmic extracts. Additional bands at about 35 kDa and 40 kDa are detected with sera #58 and #59, respectively. In case of sera directed against full-length mCSS (#51, #56, #61) only #56 and #61 showed clear bands of the expected sizes in nuclear extracts of the transfected cells. A good signal to noise ratios was observed for both sera, whereupon signal of #61 was less intense than that of serum #56. However, Flag-tagged mCSS was also detected in cytoplasmic extract by serum #56. Although purified by retroblotting, serum #51 gave dramatic background signals, which limit the usability of this serum for Western blot analyses. In figure 16B endogenous CSS as well as Flag-tagged murine CSS is detected in nuclear extracts (NE) of CHO C6 (wt) and stably transfected cells (CHO CSS) with anti-CSS mAB LF6. Due to overexpression low amounts of Flag-tagged CSS are detected in the cytoplasmic extract of stably transfected CHO cells (+CSS).

According to these experiments endogenous CMP-sialic acid synthetase of CHO C6 cells is conjugated to the cell nucleus.
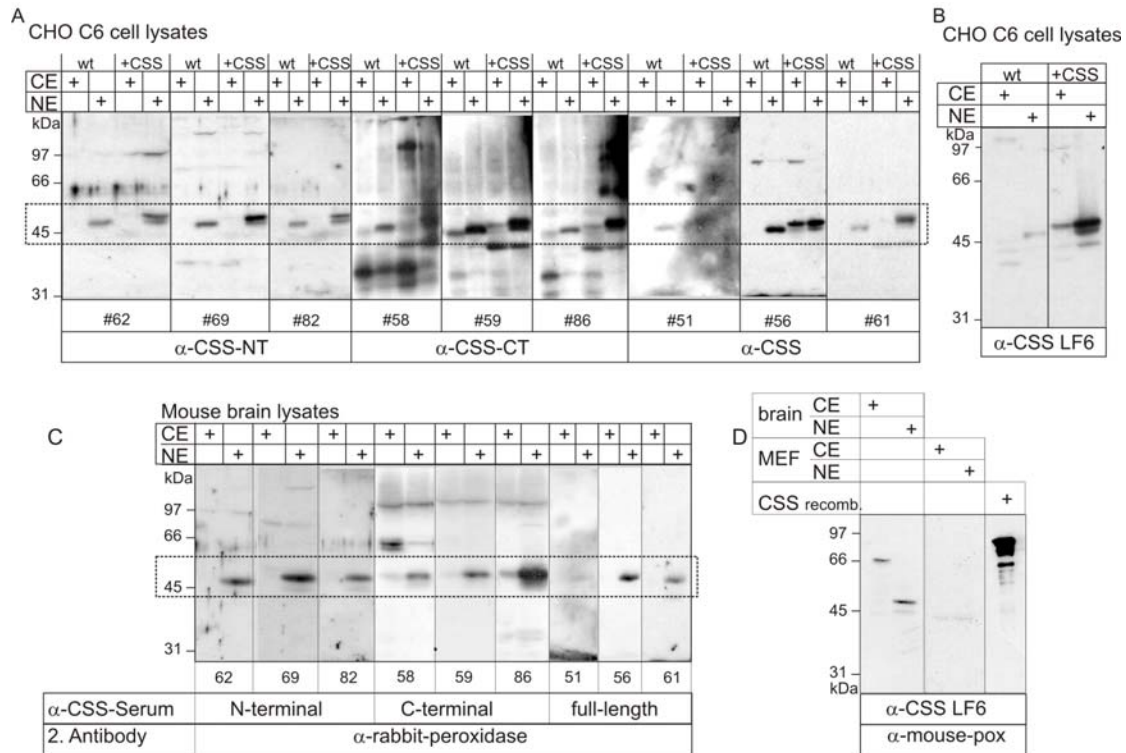

Fig. 16 Detection of endogenous and recombinant mCSS in nuclear and cytoplasmic extracts of CHO C6 cells and mouse brain by Western blot analysis. Western Blot analysis of nuclear and cytoplasmic extracts from (A, B) CHO C6 (wt) cells as well as CHO cells stably transfected with Flag/Myc-tagged murine CSS (+CSS) and (C, D) mouse brain (P0.5) as well as mouse embryo fibroblasts (MEF). 20 µg cytoplasmic extract (CE) or nuclear extract (NE) were separated by 10% SDS-PAGE and blotted onto nitrocellulose. Development was performed with (A, C) polyclonal antisera (1:1000), (B, D) mAB LF6 or anti-Flag antibody as a positive control (not shown), peroxidase-conjugated secondary antibody and ECL detection. The box indicates the position of endogenous and epitope-tagged CSS. Furthermore, NusAStrepII-tagged mCSS-NT (NusA-StrepII tag, aa 39-267, M_{calc}: 82 kDa) was loaded as positive control (D, 100 ng/lane).

Endogenous CSS in mouse brain extracts (C57 Bl/6J, postnatal day 0.5) was detected by Western blot analysis with the generated polyclonal antibodies (Fig 16C). With the exception of serum #51, all sera showed a clear band at the expected size of 48 kDa for endogenous CSS (boxed) in nuclear extracts (NE). Interestingly, additional bands with a molecular weight of about 110 kDa and 66 kDa were detected with anti-CSS-CT sera, and serum #58, respectively. However, so far, these bands could not be assigned to CSS. Nuclear sequestration was confirmed by Western blot analysis of mouse brain extract developed with mAB LF6 (Fig 16D). A 50 kDa band was detected with mAB LF6 in nuclear extract of mouse brain, however, no band is visible in mouse embryonic fibroblast (MEF) cells (Fig 16D). Purified

recombinant NusA-StrepII-tagged mCSS served as a positive control. Remarkably, an additional band of about 70 kDa was detected in cytoplasmic extracts of mouse brain by mAB LF6 (Fig. 16D). The nature of this band is further analysed below.

All polyclonal sera, with the exception of serum #51, together with mAB LF6 are of good quality and usable for detection of recombinant, as well as endogenous CSS by Western blot analysis. Polyclonal sera directed against the mCSS N-terminus are all of reasonable quality and are sensitive to recombinant as well as endogenous mouse CSS. C-terminal sera are of limited usability for Western blot analysis of endogenous and recombinant CSS, because of some additional protein bands, which most probable do not correspond to CSS. Quality of sera deriving from immunisation with full length mCSS is very inhomogeneous. In case of serum #51 specific bands are hardly detectable, due to high background signals. For further use this serum needs to be additionally purified. In contrast to this serum, especially serum #56 but also serum #61 show excellent signal to noise ratios and detected recombinant as well as endogenous CSS in CHO C6 and brain lysates. Anti-mCSS serum #56, anti-mCSS-NT serum #69 and anti-mCSS-CT serum #86 gave the best signal to noise ratios. In table 6 all obtained specificities are summarised.

One of the major goals of this study was the direct visualisation of endogenous CSS in murine cells and tissues. Therefore sera were tested for application in histological analysis of mouse brain, kidney and liver. After fixation in 4% paraformaldehyde (PFA), dehydration and embedding in paraffin, tissues were sectioned at 3 μm. Staining was performed with various CSS specific polyclonal sera, and anti-rabbit IgG-biotin conjugate as secondary antibody followed by ABC-complex (Vector Laboratories, Inc.) based staining. Preimmunesera were used as negative control and initial titrations for determination of the appropriate serum dilutions were done on murine brain sections. Preliminary tests revealed that the anti-CSS-CT serum #59 was especially applicable for immunohistochemical approaches. Histological analysis of mouse brain (P 0.5), kidney (P0.5) and liver (adult) with anti-CSS-CT serum #59 are depicted in figure 17. Staining observed with preimmune serum (not shown) was mirrored by staining only with the secondary antibody, leaving out serum #59. Thus, in figure 17 (left) only the negative control with peroxidase-coupled secondary antibody is shown. In all three organs nuclear localisation of endogenous CSS was observed (Fig. 17 right), further approved by haematoxylin / eosin (HE) staining of analogous samples (data not shown). Beside the additional analysis of murine pancreas and spleen sections, also different ages of the above mentioned organs

were analysed by staining with serum #59 (data not shown). In all organs investigated so far, CSS exclusively resides in the cell nucleus. Notably, these are the first in-situ experiments done, proving the nuclear localisation of endogenous mCSS, and supporting the previously performed studies with recombinant protein (Münster et al., 2002).
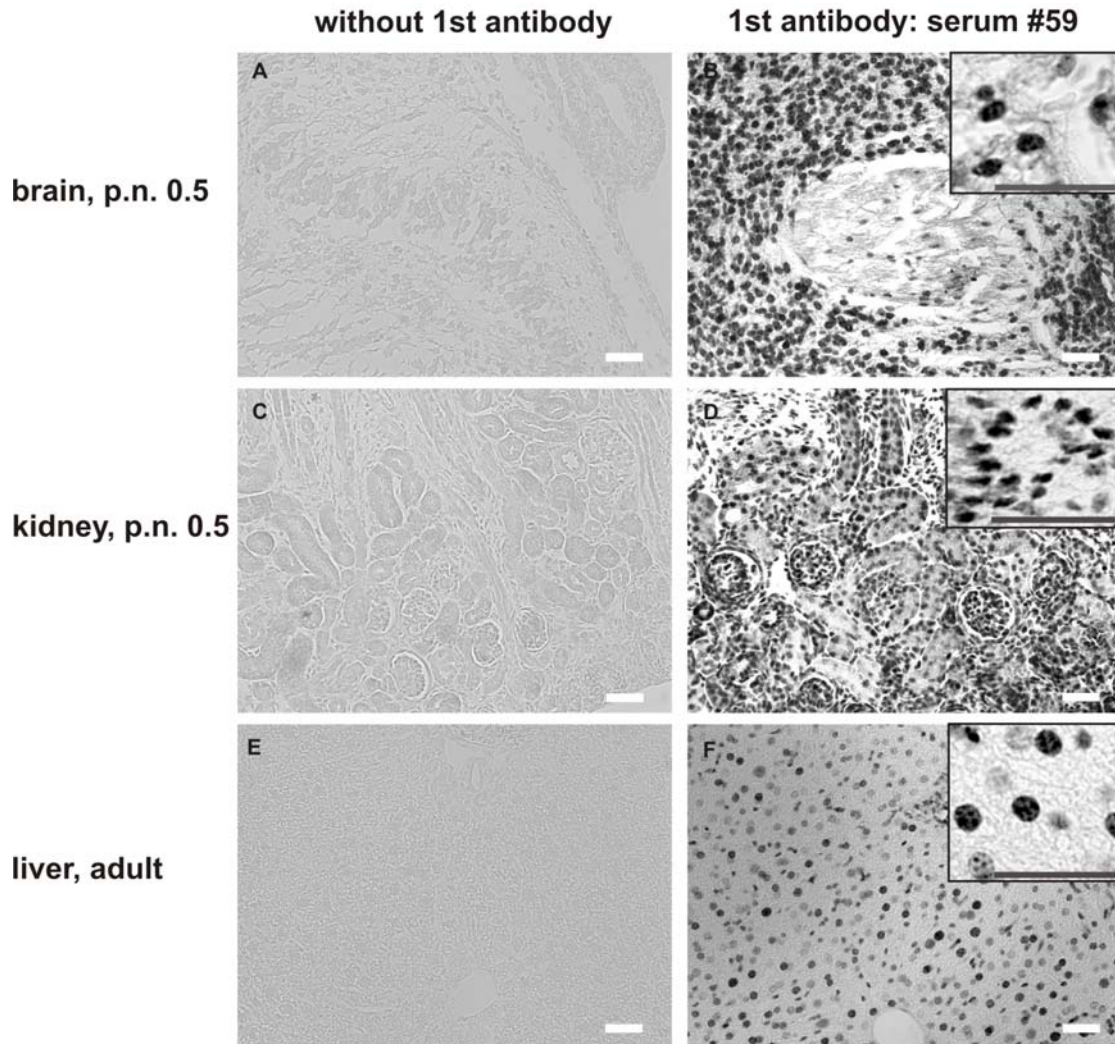

Fig. 17 Immunohistological detection of endogenous CSS in mouse brain, kidney and liver. (A, B) Mouse brain, postnatal day 0.5, (C, D) mouse kidney, post natal day 0.5 and (E, F) adult mouse liver were fixed in 4% PFA and sectioned after dehydration and paraffin embedding. Staining was performed with anti-CSS-CT serum #59 (1:4000 dilution) and anti-rabbit IgG-biotin conjugate as secondary antibody, developed with ABC-complex (Vector Laboratories, Inc.). Nuclear localisation of endogenous CMP-Sialic acid synthetase (CSS) was confirmed in all three tissues (B, D, F, right column). Negative controls without first antibody are shown in the left column (A, C, E). The white bar indicates a length of 100 µm (200 x magnifications). Zoomed pictures are inserted top right.

Table 6. Specificity and application of anti-mCSS antibodies

Application in	Western Blotting				Histology
CSS from	mouse		hamster	other species	
	recombinant	endogenous	endogenous		
α-NT					
#62	+	+	+	*H. sapiens*	
				D. melanogaster	
#69	++	++	++	*H. sapiens*	
				D. melanogaster	
				N. meningitidis [w]	
#82	+	++	+	*H. sapiens*	
				D. melanogaster	
				N. meningitidis	
α-CT					
#58	+	+[a]	+[a]		
#59	+	+	+		+++
#86	+	++	+[a]		
α-full-length					
#51	+/-[b]	+/-[b]	+/-[b]	*H. sapiens*	
#56	+++	+++	+++	*H. sapiens*	
				O. mykiss	
#61	++	++	++	*H. sapiens*	
mAB LF6	+++		++	*H. sapiens*	

[a]additonal bands

[b]high background

[w]weak signal

2.3.4. Monoclonal antibody LF6 recognizes an epitope located within the N-terminal domain of murine CMP-sialic acid synthetase

With the aim to identify the CSS epitope that is recognized by mAB LF6, several in the laboratory available CSS deletion mutants were expressed in *E.coli* and the purified proteins or soluble and insoluble extracts, respectively, were analysed by Western blot analysis. Deletion mutants are shown schematically in figure 18A. Figure 18B (left part) illustrates that LF6 specifically recognizes the N-terminal domain of murine CSS (residues 39-267, CSS-NT), whereas CSS-CT (residues 267-432) was not detected. As a positive control, an identical blot with purified

recombinant NusA-StrepII-tagged proteins was stained with Strep-Tactin-AP conjugate (Fig. 18B, right part). Soluble and insoluble fractions of bacterial lysates expressing a truncated CSS-NT (residues 83-267, N-terminal myc tag) or an extended CSS-CT (residues 234-432, NusA-StrepII-tag) were analysed by Western blotting developed with mAB LF6. A parallel approach was performed for detection of epitope tags (Fig 18C, centre and right). While the extended C-terminal domain (aa 234-432) with the calculated molecular weight of 74 kDa was not detected by mAB LF6, residues 83-267 with the calculated size of 25 kDa was recognized (Fig 18C, left part). Experiments shown in figure 18C restrict the epitope recognized by mAB LF6 to residues 83 to 234 of murine CSS. Furthermore, a short amino acid stretch was identified that seems to be required for LF6 binding. The deletion mutant lacking residues Pro^{196} to Arg^{202} was not recognized by LF6 (Fig 18D). To further identify crucial amino acids within this motif several point mutants were investigated. LF6 recognized all point mutants analysed in figure 18D (K198A/R201A, R202A, P200A, Q203A), and only the deletion mutant lacking residues Pro^{196} to Arg^{202} was not detected, indicating that these residues are crucial factors of the LF6 epitope. As loading control, a parallel blot was developed with an anti-myc antibody, corresponding to the used epitope-tags. Double bands found with LF6 development may be due to sample preparation or SDS-PAGE and were also found – less intense – in the control blot.
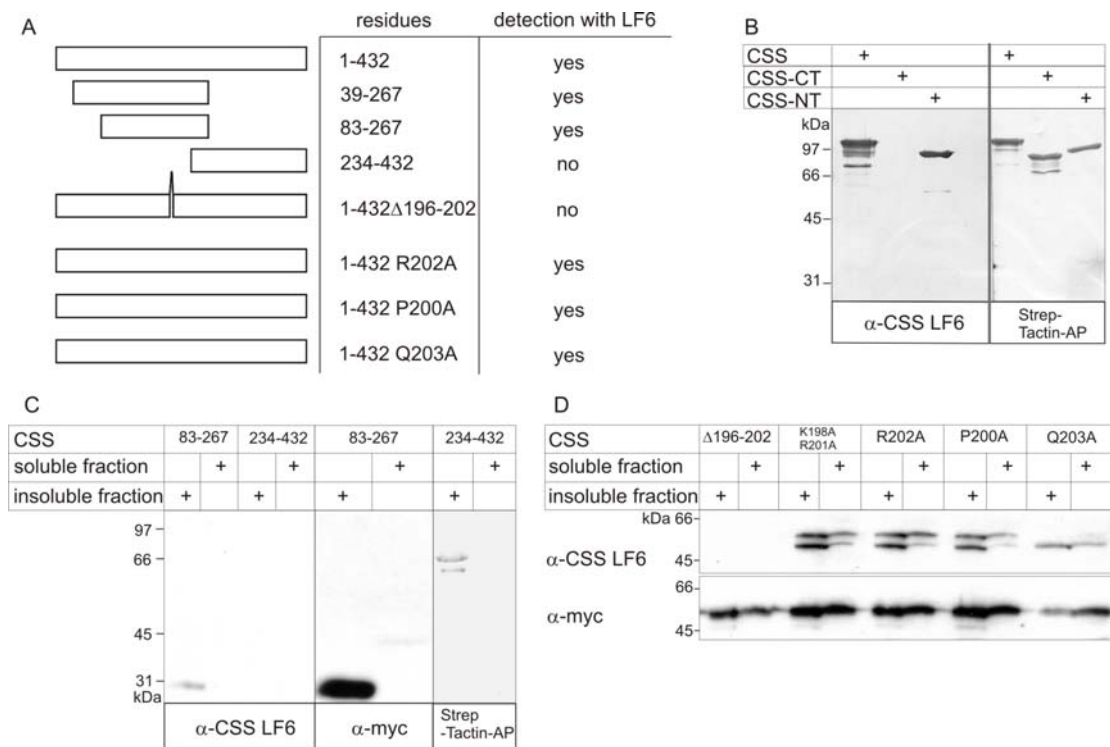
Fig. 18 Epitope mapping of anti-CSS monoclonal antibody LF6

(A) Schematic representation of mCSS constructs used for LF6 epitope mapping and results obtained by Western blotting. (B) 1.5 µg of affinity purified StrepII tagged full-length CSS (39-432), CSS-CT (267-432) or CSS-NT (39-267), respectively, were separated by 10% SDS-Page and blotted onto nitrocellulose. The Western blotting was developed with cell culture supernatant of hybridoma clone LF6 and AP-coupled secondary antibody using BCIP/NBT as substrate (left side). StrepTactin-AP conjugate was used as positive control for development of an identical approach (right side). Due to the N-terminal NusA-StrepII-Tag (56 kDa) the calculated molecular weights of the CSS, CSS-CT, and CSS-NT fusion proteins are 100 kDa, 74 kDa and 82 kDa, respectively. (C, D) *E.coli* BL21 were used to express the truncated proteins encoded in vectors pTrc-FM-GB17 (aa 83-267), pTrAM22-Del2l (aa1-432, Δ196-202), pTrAM22-Mut2a (aa1-432, K198A, R201A), pTrAM22-Mut2f (aa1-432, R202A), pTrAM22-Mut2g (aa1-432, P200A) and pTrAM22-Q203 (aa1-432, Q203A), driving the expression of N-terminal myc-tagged proteins. Expression was induced at an OD600 of 0.6 with 1 mM IPTG for 5 h at 37°C. Cells of 1 ml culture volume were lysed with Bug Buster (Novagen) and separated into soluble and insoluble fractions by centrifugation. Extracts (according to 0.1 and 0.4 ml culture volume for soluble and insoluble extracts, respectively) were separated on 15% SDS-PAGE and blotted onto nitrocellulose. Western blot was analysed with the mAB LF6 and anti-myc mAB, respectively, followed by ECL detection. For analysis of NusA-Strep tagged CSS-CT (234-432), encoded in pET43a vector, the same protocol was used, but expression was performed with transfected *E.coli* BL21 (DE3). As positive control development with Strep-Tactin-AP conjugate was performed using BCIP/NBT as substrate (C, right).

2.3.5. The 70 kDa protein detected in cytoplasmic extracts of murine brain is developmentally regulated and not Hsp70

Beside the expected CSS protein band at 48 kDa in nuclear extracts of mouse brains, LF6 additionally detects a protein of about 70 kDa in cytoplasmic extracts by Western Blot analysis (Fig 16D). Interestingly, the expression of this protein appeared to be age dependent in mice. Cytoplasmic and nuclear extracts from mouse brain at embryonic day 17.5 (E 17.5) up to postnatal day 3.5 (P 3.5) have been analysed by Western blotting with mAB LF6. Figure 19A illustrates the strong expression of the 70 kDa protein in mouse embryos of day 17.5 (E 17.5). After birth, signal intensity constantly reduces up to day 2.5 (E2.5). One day later the protein band disappeared completely. To identify the nature of the 70 kDa protein, cytoplasmic extracts of two mouse brains (P0.5) were isolated and proteins were concentrated by TCA precipitation. Proteins were further separated by a two-dimensional SDS-PAGE analysis (2D-SDS-PAGE) with pI 3-11 in the first dimension. 2D-SDS-PAGE and MS analysis were done by Dr. Büttner in cooperation with the department of microbiology at TiHo Hannover in the group of Prof. Gerlach. The corresponding spot was excised (Fig 19B, circle) and the extracted protein analysed by Mass Spectrometry. Isolated peptides correspond to a heat shock protein of 70 kDa (hsp70). However, so far this result could not be verified by Western blot analyses using anti-hsp70 mAB as shown in figure 19C. Cytoplasmic (c) or nuclear (n) mouse brain extracts from newborn mice (P0.5) or adult mice (adult), respectively, were analysed by Western blotting, developed with anti-CSS mAB LF6 (Fig 19C, left), anti-hsp70 mAB (Fig 19C, middle) or secondary antibody as negative control (Fig. 19C, right). Whereas the 70 kDa signal detected by LF6 in newborn mice is not visible in adult mice, the anti-Hsp70 mAB detected a band with constant intensity in new born and adult mice (Fig 19C). In addition, nothing is known so far about the identity of the additional 40 kDa protein band, obtained by LF6 in cytplasmic and nuclear extracts of adult mouse brain extracts (Fig 19C).
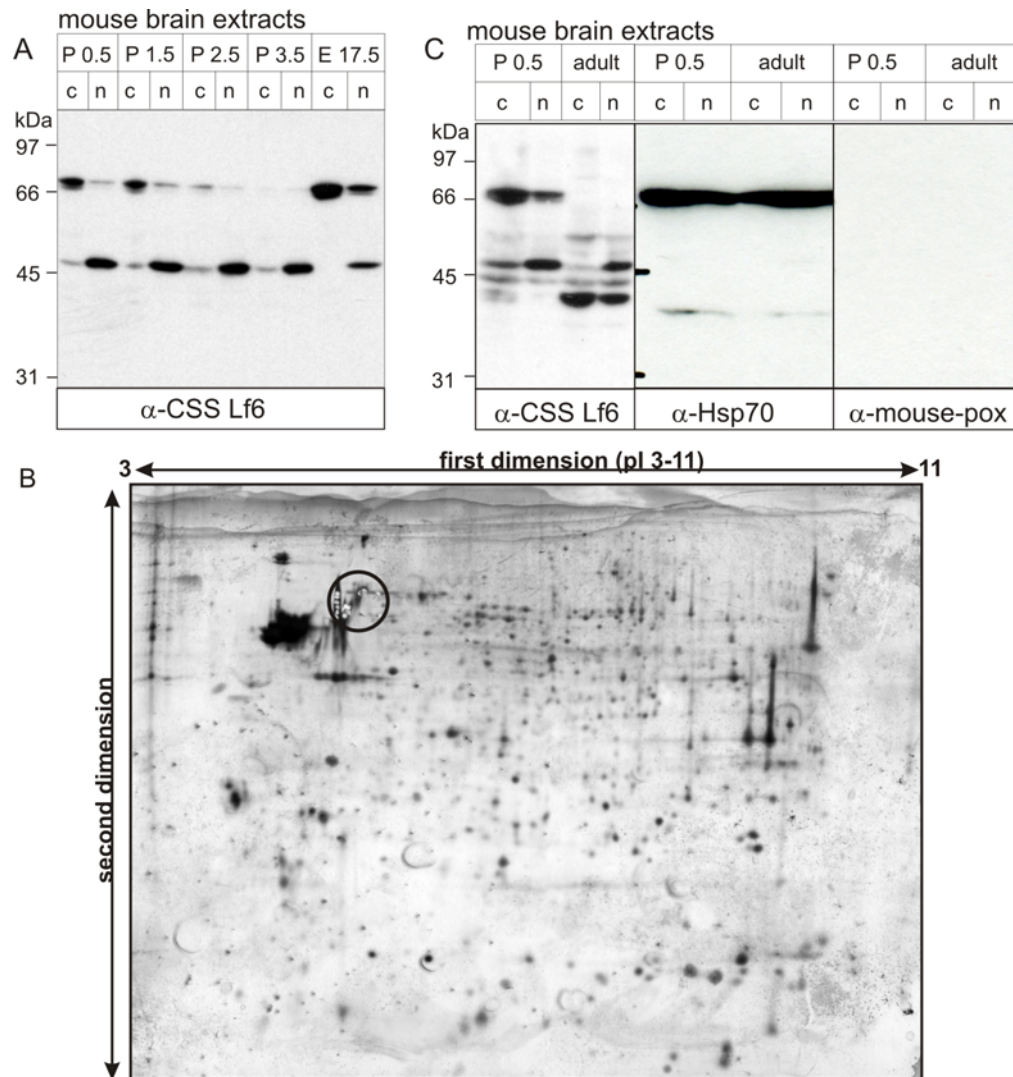
Fig19 Western Blot analysis of the 70 kDa protein, detected by mAB LF6 in cytoplasmic extracts of mouse brain. (**A**) age dependent expression of 70 kDa protein. Cytoplasmic extracts (c) and nuclear (n) extracts of mouse brains from embryos of day 17.5 (E 17.5) and postnatal mice of day 1.5, 2.5 and 3.5 (P1.5, P2.5, P3.5) were separated by 10% SDS-PAGE and blotted onto nitrocellulose. Western Blot was developed with anti-CSS mAB LF6 (5 µg/ml) and peroxidase-coupled secondary antibody. Staining was performed with ECL detection. (**B**) Two-dimensional gel electrophoresis of cytoplasmic extracts from mouse brain (P0.5) after TCA precipitation. In the first dimension, proteins were separated according to a non-linear gradient of pI 3-11, the second dimension separated proteins of different molecular weights. In order to pick spots that were identified by Western blottings developed with mAB LF6 (not shown), the 2D-PAGE was directly stained with colloidal Coomassie blue G-250 (Candiano et al., 2004). Circle indicates excised spots, corresponding to LF6 detection and surrounding. (**C**) Expression profile of the 70 kDa protein recognized by mAB LF6 is different from hsp 70 expression profile. Cytoplasmic (c) and nuclear (n) extracts of P0.5 and adult mouse brains (20 µg/lane) were separated by 10% SDS-PAGE and blotted onto nitrocellulose. Development was performed with mAB LF6 and peroxidase-conjugated secondary antibody and ECL detection (left). Identical Western blotting was developed with anti Hsp70 mAB (middle) and only peroxidase-coupled secondary antibody, as a negative control (right).

Nuclear sequestration of endogenous mCSS could be confirmed with the generated antibodies. Polyclonal serum #56 displayed high specificity for CSS in Western blot analysis; serum #59 appeared to be of particular applicability for histological staining. Monoclonal antibody LF6 is highly specific for murine CSS and gives low background signals on Western blotting.

2.4. Discussion

One of the most enigmatic characteristics of the CMP-sialic acid-activating enzyme CSS is its location in the nucleus. Successful generation of anti-CSS antibodies and their ability to specifically detect endogenous CSS against the background of total cell lysate allowed further investigation of the subcellular distribution in defined cellular systems. The presented antibody based detection of endogenous CSS revealed nuclear localisation of CSS in several tumour cell lines and organs of postnatal and adult mice, confirming earlier localisation studies, which were done by cell fractionation studies or fluorescent labelling of epitope tagged recombinant CSS. Minor proportions of CSS found in cytoplasmatic extracts

For the first time, antibodies specific for a vertebrate CSS were produced successfully. Because of its weak antigenic properties, immunisation of rabbits with murine CSS was performed for an extremely long time period (37 weeks) with booster injections in markedly long intervals (6 weeks), assuring a complete drop of antibody responses. Eight out of nine obtained polyclonal sera specifically recognizes murine CSS in a cellular background (Fig 16). Nonetheless, polyclonal sera had to be purified before application in Western blot analysis of cell lysates. In addition, sera have to be utilised in dilutions ranging from 1:500 to 1:1000, mirroring the weak immune response to CSS. Immunogenicity is determined both by the intrinsic chemical structure of the antigen and by whether or not the host animal can recognize the compound. Since CSS is a highly conserved protein, intrinsic immunogenicity is low.

In addition to purified soluble protein, solubilised inclusion bodies were used as antigens, since particulate antigens are known to be excellent immunogens. However, the immune response to CSS and CSS-NT inclusion bodies was comparable to the response to soluble CSS-CT. Remarkably, quality of the antisera seem to mainly depend on the individual immune responses of the immunised animal. For example, three rabbits were immunised with full-length CSS according to the same protocol and the resulting sera, #51, #56 and #61 showed remarkable differences. Serum #51, so far is useless for Western blot analyses, because of high background and very faint CSS specific signals. In contrast, serum #56 and #61 are

both able to detect full-length CSS as well as the separate domains CSS-NT and CSS-CT from mouse in fairly-well signal to noise ratios. Furthermore, both antisera recognize human but not *Drosophila melanogaster* CSS. Notably, only serum #56 is able to detect the CSS derived from trout (Fig 15B).

Vertebrate CSS are comprised of two domains, with both Sia-activating enzymatic activity and nuclear localisation signal localised in the N-terminal domain (residues 39-267 in mCSS) (Krapp et al., 2003; Münster et al., 2002). So far, the biological significance of the C-terminal domain in vertebrate CSS is unknown and no tool was available for separate investigation of the two domains. The generated polyclonal antibodies #62, #69 and #82 and the monoclonal antibody LF6 are specific for the N-terminal domain of the murine enzyme, whereas sera # 58, #59, #86 specifically detect the C-terminal domain (CSS-CT). Interestingly, analyses of mouse brain extracts shows a 110 kDa band detectable only with anti-CSS-CT sera. However, specificity of these bands remains to be elucidated. The generated antibodies significantly recognizes a band with the apparent molecular weight of 50 kDa in nuclear extracts (48 kDa is expected for full-length mCSS), indicating that endogenous mature CSS is not cleaved inside the cell. Western Blot analysis of CHOC6 nuclear extracts (Fig 16A) showed the expected 48 kDa band, representing endogenous hamster CSS, with sera specific for CSS-NT as well as for CSS-CT. The CSS-CT specific antibodies detected further protein bands at around 35 and 40 kDa. However, so far it can not be judged if these signals were due to CSS-specific reactions, representing degraded or cleaved enzyme, or if they were background reactions with non-CSS proteins. A specific cleavage of CSS in the two separate domains can be excluded, because the obtained bands do not fit to one of the two domains (CSS-NT is about 30 kDa and CSS-CT 18 kDa). N-terminal degradation seems likely, because also in case of recombinantly expressed CSS degradation can be observed, if the first 38 residues are present. Expression of proteins starting with residue 39 resulted in much more stable CSS protein. Similar observations were reported by Krapp et al. (2003) for the murine and by Lawrence et al. (2001) for the human enzyme.

Nuclear localisation of endogenous CSS was confirmed by Western blot analyses and immunohistochemical staining of various mouse tissues, using the generated antibodies. This raises the question about relevance of nuclear export signals (NES) *in vivo*, recently observed for mCSS by Fujita *et al.* (2007). However, studies carried out so far only give fragmentary insight and do not provide a global picture of the natural situation. It would be reasonable to observe CSS localisation time- and

tissuedependent. In addition, a heterokaryon assay could provide evidence whether CSS shuttles inside the cell.

The epitope recognized by anti-CSS mAB LF6 could be narrowed down to residues 83-234 of murine CSS. Usability of mAB LF6 in Western Blot analysis argues for a predominantly linear epitope, since proteins are denatured and unfolded by SDS-PAGE. Mouse CSS, comprised of residues 63-267 gave a more intense signal on Western blotting, than the previous mentioned CSS (residues 83-234). If this observation is due to an already impaired LF6 epitope or caused by technical reasons needs to be further investigated. In case of internal point- and deletion mutants, the only constructs not detected by LF6 was the full-length CSS with Pro^{196}-Arg^{202} deletion. On the other hand detection of single and double mutants within this amino acid stretch (K198A/R201A, R202A, P200A, Q203A) was not impaired. Possibly single mutations have no significant influence on LF6 avidity, due to compensatory effects of the neighbouring residues. Furthermore it is likely that also residues adjacent to Pro^{196}-Arg^{202} are part of LF6 epitope. So far it can be concluded, that the amino acid stretch Pro^{196}-Arg^{202} in mCSS at least partially belongs to the LF6 epitope. These results are consistent with the inability of mAB LF6 to detect purified recombinant trout or NmB CSS in Western blottings (Fig. 15F). Amino acids sequences at the equivalent positions to the murine $P^{196}AKRPRR$ motif are significantly different in trout and NmB CSS, with $P^{185}CNRPRR$ and $S^{159}DLEQPR$, respectively (Fig 12). Further experiments could comprise deletion or mutation of residues N^{193} to W^{205} neighbouring the $P^{196}AKRPRR$ motif. Additionally residues P196-R202 could be exchanged to poly-alanin, to exclude that impaired sensitivity is just due to conformational changes. Another possibility to determine the LF6 epitope would be proteolytic fingerprinting. As known from 3D structure of the CSS N-terminal domain, residues Pro^{196}-Arg^{202} are part of the dimerisation domain and positioned at the protein surface. They include the nuclear localisation signal, and residues of the active site. Success in initial immunoprecipitation experiments supports a freely accessible LF6 epitope of the native, non-denatured protein.

Initial trials to immunoprecipitate the 70 kDa protein detected by anti-CSS mAB LF6 in cytoplasmic extracts of brain lysates of newborn mice, failed, indicating that the LF6 epitope is not freely accessible on the surface of the native 70 kDa protein. Identicalness of this protein with the heat shock protein 70, as indicated by 2D gel analysis und mass spectrometry, seems unlikely after Western blot analysis of mouse brain extracts with anti-hsp70 mAB. In contrast to the 70 kDa protein recognized by mAB LF6, the protein detected with the anti-hsp70 mAB is also present in brain extracts of adult mice (Fig 19C). Incorrect identification of proteins

after separation of total cell lysates by 2D-PAGE may be due to superposition of proteins with similar properties. Therefore, further experiments will include repetition of 2D/MS experiments with better disintegration of the corresponding pI region.

Conclusively, the successful production of specific antibodies directed against the murine CMP-sialic acid synthetase opens floodgates to further investigate the endogenous protein *in vivo*. Furthermore, it might be possible to identify interacting proteins by co-immunoprecipitation, due to promising initial experiments. Moreover, the antibodies are indispensable tools for the investigation of the biological function of nuclear sequestration in gene-targeted mice.

General Discussion

CMP Sialic acid synthetases (CSS) catalyse the activation of sialic acid (Sia) to CMP-Sia which then serves as substrate for Golgi resident sialyltransferases. Vertebrate CSS reside inside the cell nucleus, which is a very unusual and so far unexplained localisation, because all other sugar activating enzymes are restricted to the cytoplasm. About 40 years ago CSS enzymatic activity has been found to be associated with the crude nuclear fraction of preparations of hog retina (Kean, 1969). Nuclear localisation was further proven by a series of studies carried out by preparation of subcellular fractions from a variety of tissues in combination with detection of CSS enzymatic activity (for reviews see Kean et al., 1991 and 2004). Immunofluorescent labelling of the recombinant murine, human and trout enzymes in cell culture further confirmed nuclear sequestration (Münster et al., 1998; Lawrence et al., 2001; Nakata et al., 2001). However, the biological function of vertebrate CSS in the cell nucleus is still an enigma, since nuclear localisation is no prerequisite for the sugar activating function (Münster et al., 2002). The second enigmatic aspect of vertebrate CSS is its additional C-terminal domain, which is conserved among all vertebrate CMP-sialic acid synthetases, but neither required for nuclear targeting (Münster, pers. communication) nor Sia activation (Krapp et al., 2003).

In the present study both questions were addressed using the murine CSS as model structure. First, the relevance of the C-terminal domain (CSS-CT) was addressed, by studying enzymatic activity and crystal structure of the murine CSS-CT. Second, endogenous vertebrate CSS was analysed in terms of its intracellular localisation by generation and application of mono- and polyclonal antibodies directed against the murine enzyme.

C-terminal domain of vertebrate CMP-sialic acid synthetase

The first part of this work focussed on the biological and structural characterisation of the C-terminal domain of murine CMP-sialic acid synthetase (CSS). Primary sequence analysis revealed homology to bacterial KDO8P phosphatases belonging to the HAD superfamily (chapter 1.3.1.). However, by careful comparison of proposed active site residues with equivalent residues of functional phosphatases of this family (Collet et al., 1998), two striking amino acid exchanges were obtained, which are conserved within all known vertebrate CSS-CT sequences. Thus, phosphatase activity assays were performed with native and mutated mouse CSS-CT, whose active site composition was adapted to functional phosphatases. Although various substrates were tested, no activity could be observed *in vitro* (1.3.2.). Nevertheless,

after having solved the 3D structure of the CSS, outstanding similarity to HAD phosphatases, as implied before by primary sequence analysis, was obtained. With the benefit of this new crystal structure the lack of phosphatase activity could be further explained by discovery of amino acids occluding the active site entry. Reflecting flexibility of the native enzyme, the crystal structure revealed two different protein conformations of CSS-CT, with a less pronounced closure of the active site in monomer B than in monomer A (Fig. 9). Possible but unlikely, CSS-CT phosphatase activity might be dependent on (i) cofactor binding that induce opening of the CSS-CT active site, or (ii) CSS-CT might be highly specific for its physiological substrate as has been shown for *Escherichia coli* KDO8P phosphatase (Wu and Woodard, 2003), and (iii) might use an alternative catalytic mechanism, different from HAD phosphatases (Wang et al., 2002). The findings in amino acid sequence as well as in 3D structure tempt to speculate that vertebrate CSS-CTs originate from HAD-phosphatases, but in course of evolution adopted a physiological function different from phosphatases.

However, domain structure of CSS still suggests a possible bifunctionality of the entire protein. It is reasonable to assume that each of the separately expressed domains, CSS-NT as well as CSS-CT, is correctly folded: 3D structure of murine CSS-NT displays obvious similarity to the structural arrangement of *Neisseria meningitidis* CSS (Fig. 5; Mosimann et al., 2001; Krapp et al., 2003) and enzymatic activity of the recombinant domain is comparable with that of recombinant full length CSS (data not shown). Furthermore, CSS-CT crystal structure shows undeniable similarity with 3D coordinates of functional phosphatases (Parsons et al., 2002; Allen and Dunaway-Mariano, 2004). Recombinant expression of the two separate CSS domains provides a higher degree of soluble protein than expression of the full-length protein, indicating independent folding mechanisms (data not shown and Krapp et al., 2003). Interestingly, gene fusion events seem to be quite common for CSS, since bifunctionality has already been described for CSS from *E. coli* K1 and *S. agalactiae* serogroup V (Liu et al., 2004; Yu et al., 2006). Acetylhydrolase activity has been shown for the C-terminal domains of both CSS *in vitro* (Steenbergen et al., 2006; Lewis et al., 2007). Interestingly, deacetylation of O-acetylated Sia monomers is closely linked to Sia activation. Modulation of the degree of O-acetylation in the capsule polysaccharide influences the virulences and immunogenicity of the bacteria (Frasa et al., 1993; Carlin et al., 2007). However, while the Sia activating domain is highly conserved from bacteria to man, no homology of CSS-CT was found at primary sequence level (Fig. 3).

The lack of activity points towards a non-enzymatic function of vertebrate CSS-CT. In the present study the induction of CSS tetramer formation could be shown to be dependent on the presence of the C-terminal domain. As investigated by size exclusion chromatography experiments, the N-terminal domain of murine CSS exclusively forms dimers in solution, whereas the separately expressed C-terminal domain as well as the entire protein forms tetramers (Fig. 11). These observations are in line with multimer formation of other vertebrate CSS, obtained by size exclusion chromatography, by Schauer et al., Rodriguez-Aparicio et al. and Vionnet et al. for CSS purified from *Rana esculenta*, *rattus norvegicus* and *bos taurus*, respectively (Schauer et al., 1980; Rodriguez-Aparicio et al., 1992; Vionnet et al., 1999). Probably multimer formation *per se* is not sufficient to explain the evolutionary driving force sustaining this domain, but it might be the basis of future explanations. Possibly C-terminal domains of vertebrate CSS originate from functional enzymes but then transformed to regulatory or stabilising domains, without own catalytic activities. Conservation of CSS-CT among all vertebrate CSS strongly suggests that its presence is somehow important for animal life; otherwise a more inhomogeneous picture would be expected. Interestingly, genome analyses revealed the possibility of splice variants for human and murine CSS, possibly resulting in separate expression of CSS-NT and CSS-CT, respectively (Gerhard et al., 2004; Carninci et al., 2005). However, no experimental confirmation for CSS isoform expression is available yet. Experiments performed in this work only provide evidence for the presence of full-length CSS in the analysed tissues. In addition, endogenous CSS seems to be stable as full length protein.

Assuming that CSS nuclear localisation is an essential feature in vertebrates, the present C-terminal domain could simplify coordination of CSS localisation by increasing molecular weight of the monomer from 30 kDa to 48 kDa. Smaller proteins, like the 30 kDa CSS-NT, are able to freely diffuse through the nuclear pore complex (NPC), whereas proteins with a molecular weight above 40 kDa require active transport (Christophe et al., 2000; Terry et al., 2007). However, size exclusion chromatography experiments showed that CSS-NT alone forms stable dimers in solution, which also sufficiently increases molecular weight. Furthermore, indirect immunofluorescent labelling of transfected NIH 3T3 cells revealed nuclear localisation for the recombinant Flag-tagged CSS-NT in NIH 3T3 cells (Münster, pers. communication).

Structural similarity of CSS-CT to proteins derived from thermostable bacteria (Table 3), indicate, that CSS-CT might enhance stability of the entire CSS. *In vitro* thermostability assays or pulse chase experiments will be performed to answer the

question of an altered half life of mCSS in presence or absence of the C-terminal domain.

CMP-sialic acid synthetase in the nucleus

The purpose of generating specific anti-CSS antibodies described in chapter 2 was to pave the way for future studies addressing the nuclear localisation of vertebrate CSS in more detail. Specific mono- and polyclonal antibodies were produced directed against the separate domains of murine CSS as well as the full-length enzyme. Thus it was made possible for the first time to directly monitor localisation of endogenous CSS, and first histological as well as cell fractionation studies were carried out in this PhD thesis, confirming nuclear sequestration of murine CSS.

Former hypotheses, assuming that the nuclear environment is somehow essential for Sia activation were disproven by mutated murine CSS exhibiting full enzymatic activity and exclusive localisation in the cytoplasm in NIH 3T3 cells (Münster et al., 2002). In these experiments a cytoplasmic CSS mutant was used in which the wild type nuclear localisation signal (NLS) sequence $K^{198}RPRR$ was replaced by $A^{198}RPAR$, and intracellular localisation was investigated by indirect immunofluorescence analysis. Both the observation of CSS activity independent from nuclear localisation (Münster et al., 2002) in NIH 3T3 cells and the fact that cytoplasmic CSS is able to restore wild-type sialylation pattern in LEC29.Lec32 cells (Münster et al., 1998), argue against (i) a limiting CTP concentration in the cytoplasm, (ii) premature feedback inhibition of UDP-GlcNAc 2-epimerase (GNE) (Hinderlich et al., 1997; Kornfeld et al., 1964), and (iii) increased degradation of CMP-Sia by cytoplasmic hydrolases (Kean and Bighouse, 1974). However, it must be kept in mind that studies available to date do not go beyond the cellular level and that the natural situation of multicellular organisms might differ completely. Interestingly, the NLS of murine CSS comprises basic residues, which are also essential for enzymatic activity, in particular Arg^{199} and Arg^{202}. This offers the possibility of a random event, connecting enzymatic activity with nuclear localisation. However, separation of sequence motifs responsible for nuclear localisation and activity in rainbow trout CSS underlines the importance of nuclear import (Tiralongo et al., 2007) and indicate that nuclear localisation of vertebrate CSS is not fortuitous. Furthermore, basic clusters that may serve as nuclear localisation signals have been found in predicted CSS sequences from echinoderms on (Tiralongo et al., 2007).

An analogous role to CSS in the nucleus might be accomplished by 3'-phosphoadenosine 5'-phosphosulfate (PAPS) synthetase 1 (Venkatachalam, 2003), producing the activated sulfate donor PAPS, which then serves as substrate for Golgi

resident sulfotransferases. The human PAPS synthetase 1 is a nuclear protein (Besset et al., 2000). Remarkably, sulfation and sialylation are both critical terminal modifications of many cell surface glycoconjugates (Bowman and Bertozzi, 1999; Varki, 1997) that result in highly anionic structures, such as the 6-sulfo-sialyl-LewisX [Sia(α2–3)Gal(β1–4)(Fuc(α1–3))(6-SO$_3$-) GlcNAc] motif, controlling lymphocyte recruitment into lymph nodes (Hemmerich et al., 1995; Mitsuoka et al., 1998).

On the other hand, it seems likely, that also trace amounts of CSS in the cytosol might be sufficient for adequate generation of sialylated structures. Interestingly, investigation of CSS expression of mouse embryo fibroblasts (MEFs) revealed expression levels below the detection limit as assayed by Western blot analysis (Fig. 16D) and indirect immunofluorescent labelling experiments (not shown), whereas CSS expression in CHO C6 cells could be detected with both methods in the nucleus (Fig. 16AB and data not shown). Nevertheless, recent analyses of protein bound Sia, carried out by DMB/HPLC analysis (Lin et al., 2000; Inoue et al., 2001), revealed similar amounts for MEF and CHO C6 cells, indicating the presence of activated sialic acids and thus of functional CSS (personal communication with Prof. Geyer, JLU Gießen, Germany). The recent identification of nuclear export signals might be involved in tight intracellular regulation of vertebrate CSS (Fujita et al., 2007).

One last hypothesis, which can only be disproven by revealing the true reason for nuclear localisation, is the possibility of a second yet unknown function within the cell nucleus. This second function of CSS might be accomplished by the C-terminal domain, which is, as discussed above, dispensable for the sugar activating function (Krapp et al., 2003). Hence, nuclear localisation could be a prerequisite for the so far unknown function of vertebrate CSS-CT, or CSS might perform a highly regulated function in the cytoplasm and can there only be tolerated in trace amounts.

Taking into account that sialic acids are present at the cell surface and have unique physicochemical properties directly modulating cellular contacts and motility (Varki, 1997, Schauer, 2000), it seems reasonable to tightly control Sia metabolism. Lack of sialylation is embryonic lethal (Schwarzkopf et al., 2002) and reduced sialylation reveals dramatic phenotypes. Hyposialylation in kidney causes severe proteinurea and renal failure at postnatal day 3.5 (Galeano et al., 2007); the lack of polysialic acid cause severe neurodevelopmental defects, postnatal growth retardation, and precocious death (Weinhold et al., 2005). Altered sialylation patterns, like enhanced expression of terminal α2,6-linked Sia, typically found on N-linked glycans of mucins on the cell surface, often correlates with poor prognosis of human malignancies (Varki and Varki, 2007). So far the only known regulatory mechanism of the

sialylation pathway is the feedback inhibition of UDP-GlcNAc 2-epimerase (GNE) by CMP-Neu5Ac, the product of CSS (Kornfeld et al., 1964, and 'general introduction'). Interestingly, a recent study suggests expression of GNE, which is the key enzyme in Sia biosynthesis, not only in the cytosol but also in the nucleus (Krause et al., 2005). Possibly, nuclear localisation of CSS might be explained by an additional regulatory mechanism, directly influencing CSS activity or specificity. Regulation of CSS activity would influence the production of all sialic acid derivatives in one step, including KDN, which is synthesised in a parallel pathway to Neu5Ac. However, activation and transfer of all sialic acid derivatives to glycoconjugates is catalysed by CSS and sialyltransferases, respectively (Inoue and Kitajima, 2006). CSS localisation in the cell nucleus directly connects the sialylation pathway with the controlling body of the cell. However, so far it is not known in which sub-nuclear compartment CSS is positioned, but proteins of common pathways are often concentrated into specific areas of the nucleus. For example (i) production and assembly of ribosome components take place in nucleoli, (ii) proteins of PML bodies are connected to various functions from transcription, DNA repair and tumour suppression up to apoptosis induction and (iii) factors for snRNP biogenesis and preRNA processing accumulate in cajal bodies (Sutherland et al., 2001; Spector, 2001; Bickmore and Sutherland, 2002). Identification of the precise localisation possibly could be used as guide to nuclear function of CSS *in vivo*. For this purpose, the generated antibodies are of tremendous value, e.g. for immunofluorescent labelling of endogenous CSS, comparing localisation with known proteins of the different nuclear compartments as well as identifying putative interacting proteins by co-immunoprecipitation. Besides a possible second function of CSS itself in the cell nucleus, the ratio of Sia and CMP-Sia may have regulatory effects on other nuclear proteins, like transcription factors, which might then influence protein, e.g. sialyltransferase expression. By localisation of CSS in the nucleus, all Sia molecules, which are destined for glycoconjugates, have to pass the nucleus first. Thus a representative picture of Sia / CMP-Sia ratio is found in the nucleus. Interestingly, analysis of intracellular Sia concentration revealed a 1.2fold higher Sia concentration in the nucleus than in the cytosol (Bork et al., 2005). Further experiments indicated that Sia is enriched and trapped in the nucleus, since diffusion through the nuclear pore complexes was not observed (Bork et al., 2005). In addition, recent studies from Kontou et al. (2008) provided evidence that Sia metabolism is closely associated with signal transduction and regulation of neuronal differentiation. Treatment of PC12 cells with Sia precursors led to neurite outgrowth and erk1/2 phosphorylation as well as accumulation in the nucleus.

However, two enigmas surround vertebrate CSS: nuclear sequestration and a conserved protein domain of unknown function. This study presented the first 3D structure of a C-terminal domain of a CMP-sialic acid synthetase and provided detailed insights into enzymatic organisation. Furthermore a variety of antibodies specific for the separate domains of mCSS as well as CSS of other species were generated, and enabled localisation studies of endogenous CSS. With these tools on hand, two mouse models will be investigated in future studies. First, gene-targeted mice expressing an active CSS localised in the cytoplasm should provide insights to relevance of nuclear localisation. Second, mice expressing an active but truncated CSS lacking the C-terminal domain will give insights into the biological function of CSS-CT.

References

Allen, K.N. and Dunaway-Mariano, D. (2004) Phosphoryl group transfer: evolution of a catalytic scaffold. *Trends Biochem.Sci.* **29**(9), 495-503.

Angata, T. (2006) Molecular diversity and evolution of the Siglec family of cell-surface lectins. *Mol.Divers.* **10**(4), 555-566.

Angata, T. and Varki, A. (2002) Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. *Chem.Rev.* **102**(2), 439-469.

Aravind, L. and Koonin, E.V. (1998) The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem.Sci.* **23**(12), 469-472.

Besset, S., Vincourt, J.B., Amalric, F., and Girard, J.P. (2000) Nuclear localization of PAPS synthetase 1: a sulfate activation pathway in the nucleus of eukaryotic cells. *FASEB J.* **14**(2), 345-354.

Bickmore, W.A. and Sutherland, H.G. (2002) Addressing protein localization within the nucleus. *EMBO J.* **21**(6), 1248-1254.

Blume, A., Weidemann, W., Stelzl, U., Wanker, E.E., Lucka, L., Donner, P., Reutter, W., Horstkorte, R., and Hinderlich, S. (2004) Domain-specific characteristics of the bifunctional key enzyme of sialic acid biosynthesis, UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase. *Biochem.J.* **384**(Pt 3), 599-607.

Bork, K., Reutter, W., Gerardy-Schahn, R., and Horstkorte, R. (2005) The intracellular concentration of sialic acid regulates the polysialylation of the neural cell adhesion molecule. *FEBS Lett.* **579**(22), 5079-5083.

Bowman, K.G. and Bertozzi, C.R. (1999) Carbohydrate sulfotransferases: mediators of extracellular communication. *Chem.Biol.* **6**(1), R9-R22.

Brilli, M. and Fani, R. (2004) Molecular evolution of hisB genes. *J.Mol.Evol.* **58**(2), 225-237.

Brunger, A.T. (1993) Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Crystallogr.D.Biol.Crystallogr.* **49**(Pt 1), 24-36.

Bruses, J.L. and Rutishauser, U. (2001) Roles, regulation, and mechanism of polysialic acid function during neural development. *Biochimie* **83**(7), 635-643.

Candiano, G., Bruschi, M., Musante, L., Santucci, L., Ghiggeri, G.M., Carnemolla, B., Orecchia, P., Zardi, L., and Righetti, P.G. (2004) Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis* **25**(9), 1327-1333.

Carlin, A.F., Lewis, A.L., Varki, A., and Nizet, V. (2007) Group B streptococcal capsular sialic acids interact with siglecs (immunoglobulin-like lectins) on human leukocytes. *J.Bacteriol.* **189**(4), 1231-1237.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, et al (2005) The transcriptional landscape of the mammalian genome. *Science* **309**(5740), 1559-1563.

CCP4: The CCP4 suite: programs for protein crystallography. (1994) *Acta Crystallogr.D.Biol.Crystallogr.* **50**(Pt 5), 760-763

Christophe, D., Christophe-Hobertus, C., and Pichon, B. (2000) Nuclear targeting of proteins: how many different signals? *Cell Signal.* **12**(5), 337-341.

Claude, J.B., Suhre, K., Notredame, C., Claverie, J.M., and Abergel, C. (2004) CaspR: a web server for automated molecular replacement using homology modelling. *Nucleic Acids Res.* **32**(Web Server issue), W606-W609.

Collet, J.F., Stroobant, V., Pirard, M., Delpierre, G., and Van Schaftingen, E. (1998) A new class of phosphotransferases phosphorylated on an aspartate residue in an amino-terminal DXDX(T/V) motif. *J.Biol.Chem.* **273**(23), 14107-14112.

Comeau, S.R., Gatchell, D.W., Vajda, S., and Camacho, C.J. (2004) ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics.* **20**(1), 45-50.

Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**(22), 10881-10890.

Cowtan, K.D. and Zhang, K.Y. (1999) Density modification for macromolecular phase improvement. *Prog.Biophys.Mol.Biol.* **72**(3), 245-270.

Crocker, P.R. (2002) Siglecs: sialic-acid-binding immunoglobulin-like lectins in cell-cell interactions and signalling. *Curr.Opin.Struct.Biol.* **12**(5), 609-615.

Deszo, E.L., Steenbergen, S.M., Freedberg, D.I., and Vimr, E.R. (2005) Escherichia coli K1 polysialic acid O-acetyltransferase gene, neuO, and the mechanism of capsule form variation involving a mobile contingency locus. *Proc.Natl.Acad.Sci.U.S.A* **102**(15), 5564-5569.

Eckhardt, M., Gotza, B., and Gerardy-Schahn, R. (1998) Mutants of the CMP-sialic acid transporter causing the Lec2 phenotype. *J.Biol.Chem.* **273**(32), 20189-20195.

Edwards, U. and Frosch, M. (1992) Sequence and functional analysis of the cloned Neisseria meningitidis CMP-NeuNAc synthetase. *FEMS Microbiol.Lett.* **75**(2-3), 161-166.

Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr.D.Biol.Crystallogr.* **60**(Pt 12 Pt 1), 2126-2132.

Frasa, H., Procee, J., Torensma, R., Verbruggen, A., Algra, A., Rozenberg-Arska, M., Kraaijeveld, K., and Verhoef, J. (1993) Escherichia coli in bacteremia: O-acetylated K1 strains appear to be more virulent than non-O-acetylated K1 strains. *J.Clin.Microbiol.* **31**(12), 3174-3178.

Frosch, M., Gorgen, I., Boulnois, G.J., Timmis, K.N., and Bitter-Suermann, D. (1985) NZB mouse system for production of monoclonal antibodies to weak bacterial antigens: isolation of an IgG antibody to the polysaccharide capsules of Escherichia coli K1 and group B meningococci. *Proc.Natl.Acad.Sci.U.S.A* **82**(4), 1194-1198.

Fujita, A., Sato, C., and Kitajima, K. (2007) Identification of the nuclear export signals that regulate the intracellular localization of the mouse CMP-sialic acid synthetase. *Biochem.Biophys.Res.Commun.* **355**(1), 174-180.

Gagiannis, D., Orthmann, A., Danßmann, I., Scharzkopf, M., Weidemann, W. and Horstkorte, R. (2007) Reduced sialylation status in UDP-*N*-acetylglucosamine-2-epimerase/*N*-acetylmannosamine kinase (GNE)-deficient mice. *Glycoconj.J* **24**, 125-130

Galeano, B., Klootwijk, R., Manoli, I., Sun, M., Ciccone, C., Darvish, D., Starost, M.F., Zerfas, P.M., Hoffmann, V.J., Hoogstraten-Miller, S., Krasnewich, D.M., Gahl, W.A., and Huizing, M. (2007) Mutation in the key enzyme of sialic acid biosynthesis causes severe glomerular proteinuria and is rescued by N-acetylmannosamine. *J.Clin.Invest* **117**(6), 1585-1594.

Galfre, G. and Milstein, C. (1981) Preparation of monoclonal antibodies: strategies and procedures. *Methods Enzymol.* **73**(Pt B), 3-46.

Gerhard, D.S., Wagner, L., Feingold, E.A., Shenmen, C.M., Grouse, L.H., et al. (2004) The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome Res.* **14**(10B), 2121-2127.

Gielen, W., Schaper, R., and Pink, H. (1971) [Subcellular distribution and activity of cytidinemonophospho-N-acetylneuraminate-synthetase in the young rat brain]. *Hoppe Seylers.Z.Physiol Chem.* **352**(10), 1291-1296.

Haft, R.F. and Wessels, M.R. (1994) Characterization of CMP-N-acetylneuraminic acid synthetase of group B streptococci. *J.Bacteriol.* **176**(23), 7372-7374.

Harvey, H.A., Swords, W.E., and Apicella, M.A. (2001) The mimicry of human glycolipids and glycosphingolipids by the lipooligosaccharides of pathogenic neisseria and haemophilus. *J.Autoimmun.* **16**(3), 257-262.

Hedlund, M., Tangvoranuntakul, P., Takematsu, H., Long, J.M., Housley, G.D., Kozutsumi, Y., Suzuki, A., Wynshaw-Boris, A., Ryan, A.F., Gallo, R.L., Varki, N., and Varki, A. (2007) N-glycolylneuraminic acid deficiency in mice: implications for human biology and evolution. *Mol.Cell Biol.* **27**(12), 4340-4346.

Hemmerich, S., Leffler, H., and Rosen, S.D. (1995) Structure of the O-glycans in GlyCAM-1, an endothelial-derived ligand for L-selectin. *J.Biol.Chem.* **270**(20), 12035-12047.

Henderson, B.R. and Eleftheriou, A. (2000) A comparison of the activity, sequence specificity, and CRM1-dependence of different nuclear export signals. *Exp.Cell Res.* **256**(1), 213-224.

Hildebrandt, H., Becker, C., Gluer, S., Rosner, H., Gerardy-Schahn, R., and Rahmann, H. (1998) Polysialic acid on the neural cell adhesion molecule correlates with expression of polysialyltransferases and promotes neuroblastoma cell growth. *Cancer Res.* **58**(4), 779-784.

Hildebrandt, H., Muhlenhoff, M., Weinhold, B., and Gerardy-Schahn, R. (2007) Dissecting polysialic acid and NCAM functions in brain development. *J.Neurochem.* **103 Suppl 1**56-64.

Hinderlich, S., Stasche, R., Zeitler, R., and Reutter, W. (1997) A bifunctional enzyme catalyzes the first two steps in N-acetylneuraminic acid biosynthesis of rat liver. Purification and characterization of UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase. *J.Biol.Chem.* **272**(39), 24313-24318.

Hisano, T., Hata, Y., Fujii, T., Liu, J.Q., Kurihara, T., Eskai, N., and Soda, K. (1996) Crystal structure of L2-haloacid dehalogenase from *Pseudomonas ap.* YL. An α/β hydrolase structure that is different from the α/β hydrolase fold. *J.Biol.Chem.* **271**, 20322-20330

Holm, L. and Sander, C. (1996) Mapping the protein universe. *Science* **273**(5275), 595-603.

Holm, L. and Sander, C. (1998) Touring protein fold space with Dali/FSSP. *Nucleic Acids Res.* **26**(1), 316-319.

Inoue, S. and Kitajima, K. (2006) KDN (deaminated neuraminic acid): dreamful past and exciting future of the newest member of the sialic acid family. *Glycoconj.J.* **23**(5-6), 277-290.

Inoue, S., Lin, S.L., Lee, Y.C., and Inoue, Y. (2001) An ultrasensitive chemical method for polysialic acid analysis. *Glycobiology* **11**(9), 759-767.

Ishige, K., Hamamoto, T., Shiba, T., and Noguchi, T. (2001) Novel method for enzymatic synthesis of CMP-NeuAc. *Biosci.Biotechnol.Biochem.* **65**(8), 1736-1740.

Jelakovic, S. and Schulz, G.E. (2002) Catalytic mechanism of CMP:2-keto-3-deoxy-manno-octonic acid synthetase as derived from complexes with reaction educt and product. *Biochemistry* **41**(4), 1174-1181.

Jones, S. and Thornton, J.M. (1996) Principles of protein-protein interactions. *Proc.Natl.Acad.Sci.U.S.A* **93**(1), 13-20.

Kabsch, W. *J.* (1993) *Appl. Cryst.* **26**, 795-800

Kageshita, T., Hirai, S., Kimura, T., Hanai, N., Ohta, S., and Ono, T. (1995) Association between sialyl Lewis(a) expression and tumor progression in melanoma. *Cancer Res.* **55**(8), 1748-1751.

Kean, E.L. (1969) Sialic acid activating enzyme in ocular tissue. *Exp.Eye Res.* **8**(1), 44-54.

Kean, E.L. (1991) Sialic acid activation. *Glycobiology* **1**(5), 441-447.

Kean, E.L. and Bighouse, K.J. (1974) Cytidine 5'-monophosphosialic acid hydrolase. Subcellular location and properties. *J.Biol.Chem.* **249**(24), 7813-7823.

Kean, E.L., Münster-Kühnel, A.K., and Gerardy-Schahn, R. (2004) CMP-sialic acid synthetase of the nucleus. *Biochim.Biophys.Acta* **1673**(1-2), 56-65.

Kearney, J.F., Radbruch, A., Liesegang, B., and Rajewsky, K. (1979) A new mouse myeloma cell line that has lost immunoglobulin expression but permits the construction of antibody-secreting hybrid cell lines. *J.Immunol.* **123**(4), 1548-1550.

Kelm, S. and Schauer, R. (1997) Sialic acids in molecular and cellular interactions. *Int.Rev.Cytol.* **175** 137-240.

Keppler, O.T., Hinderlich, S., Langner, J., Schwartz-Albiez, R., Reutter, W., and Pawlita, M. (1999) UDP-GlcNAc 2-epimerase: a regulator of cell surface sialylation. *Science* **284**(5418), 1372-1376.

Kim, K.S. (2003) Pathogenesis of bacterial meningitis: from bacteraemia to neuronal injury. *Nat.Rev.Neurosci.* **4**(5), 376-385.

Kontou, M., Bauer, C., Reutter, W., and Horstkorte, R. (2008) Sialic acid metabolism is involved in the regulation of gene expression during neuronal differentiation of PC12 cells. *Glycoconj.J.* **25**(3), 237-244.

Kornfeld, S., Kornfeld, R., Neufeld, E.F., and O'Brien, P.J. (1964) The feedback control of sugar nucleotide biosynthesis in liver. *Proc.Natl.Acad.Sci.U.S.A* **52,** 371-379.

Krapp, S., Münster-Kühnel, A.K., Kaiser, J.T., Huber, R., Tiralongo, J., Gerardy-Schahn, R., and Jacob, U. (2003) The crystal structure of murine CMP-5-N-acetylneuraminic acid synthetase. *J.Mol.Biol.* **334**(4), 625-637.

Krause, S., Hinderlich, S., Amsili, S., Horstkorte, R., Wiendl, H., Argov, Z., Mitrani-Rosenbaum, S., and Lochmuller, H. (2005) Localization of UDP-GlcNAc 2-epimerase/ManAc kinase (GNE) in the Golgi complex and the nucleus of mammalian cells. *Exp.Cell Res.* **304**(2), 365-379.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**(5259), 680-685.

Lahiri, S.D., Zhang, G., Dunaway-Mariano, D., and Allen, K.N. (2006) Diversification of function in the haloacid dehalogenase enzyme superfamily: The role of the cap domain in hydrolytic phosphoruscarbon bond cleavage. *Bioorg.Chem.* **34**(6), 394-409.

Laskowski R A, MacArthur M W, Moss D S & Thornton J M (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, **26**, 283-291.

Lawrence, S.M., Huddleston, K.A., Tomiya, N., Nguyen, N., Lee, Y.C., Vann, W.F., Coleman, T.A., and Betenbaugh, M.J. (2001) Cloning and expression of human sialic acid pathway genes to generate CMP-sialic acids in insect cells. *Glycoconj.J.* **18**(3), 205-213.

Lee, B. and Richards, F.M. (1971) The interpretation of protein structures: estimation of static accessibility. *J.Mol.Biol.* **55**(3), 379-400.

Lehmann, F., Tiralongo, E., and Tiralongo, J. (2006) Sialic acid-specific lectins: occurrence, specificity and function. *Cell Mol.Life Sci.* **63**(12), 1331-1354.

Lepers, A., Shaw, L., Schneckenburger, P., Cacan, R., Verbert, A., and Schauer, R. (1990) A study on the regulation of N-glycoloylneuraminic acid biosynthesis and utilization in rat and mouse liver. *Eur.J.Biochem.* **193**(3), 715-723.

Lewis, A.L., Cao, H., Patel, S.K., Diaz, S., Ryan, W., Carlin, A.F., Thon, V., Lewis, W.G., Varki, A., Chen, X., and Nizet, V. (2007) NeuA sialic acid O-acetylesterase activity

modulates O-acetylation of capsular polysaccharide in group B Streptococcus. *J.Biol.Chem.* **282**(38), 27562-27571.

Lewis, A.L., Hensler, M.E., Varki, A., and Nizet, V. (2006) The group B streptococcal sialic acid O-acetyltransferase is encoded by neuD, a conserved component of bacterial sialic acid biosynthetic gene clusters. *J.Biol.Chem.* **281**(16), 11186-11192.

Li, Y.F., Hata, Y., Fujii, T., Hisano, T., Nishihara, M., Kurihara, T., and Esaki, N. (1998) Crystal structures of reaction intermediates of L-2-haloacid dehalogenase and implications for the reaction mechanism. *J.Biol.Chem.* **273**(24), 15035-15044.

Lin, S.L., Inoue, S., and Inoue, Y. (2000) Acid-base properties of the reaction product of sialic acid with fluorogenic reagent, 1,2-diamino-4,5-methylenedioxybenzene (DMB). *Carbohydr.Res.* **329**(2), 447-451.

Liu, G., Jin, C., and Jin, C. (2004) CMP-N-acetylneuraminic acid synthetase from Escherichia coli K1 is a bifunctional enzyme: identification of minimal catalytic domain for synthetase activity and novel functional domain for platelet-activating factor acetylhydrolase activity. *J.Biol.Chem.* **279**(17), 17738-17749.

Liu, M.Z. and Lee, Y.C. (2001) Comparison of chemical and enzymatic synthesis of 2-acetamido-2-deoxy-D-mannose 6-phosphate: a new approach. *Carbohydr.Res.* **330**(3), 413-419.

Maliekal, P., Vertommen, D., Delpierre, G., and Van Schaftingen, E. (2006) Identification of the sequence encoding N-acetylneuraminate-9-phosphate phosphatase. *Glycobiology* **16**(2), 165-172.

Mitsuoka, C., Sawada-Kasugai, M., Ando-Furui, K., Izawa, M., Nakanishi, H., Nakamura, S., Ishida, H., Kiso, M., and Kannagi, R. (1998) Identification of a major carbohydrate capping group of the L-selectin ligand on high endothelial venules in human lymph nodes as 6-sulfo sialyl Lewis X. *J.Biol.Chem.* **273**(18), 11225-11233.

Miyagi, T., Wada, T., Yamaguchi, K., and Hata, K. (2004) Sialidase and malignancy: a minireview. *Glycoconj.J.* **20**(3), 189-198.

Mizanur, R.M. and Pohl, N.L. (2007) Cloning and characterization of a heat-stable CMP-N-acylneuraminic acid synthetase from Clostridium thermocellum. *Appl.Microbiol.Biotechnol.* **76**(4), 827-834.

Morais, M.C., Zhang, W., Baker, A.S., Zhang, G., Dunaway-Mariano, D., and Allen, K.N. (2000) The crystal structure of bacillus cereus phosphonoacetaldehyde hydrolase: insight into catalysis of phosphorus bond cleavage and catalytic diversification within the HAD enzyme superfamily. *Biochemistry* **39**(34), 10385-10396.

Morell, A.G., Gregoriadis, G., Scheinberg, I.H., Hickman, J., and Ashwell, G. (1971) The role of sialic acid in determining the survival of glycoproteins in the circulation. *J.Biol.Chem.* **246**(5), 1461-1467.

Morris, A.L., MacArthur, M.W., Hutchinson, E.G., and Thornton, J.M. (1992) Stereochemical quality of protein structure coordinates. *Proteins* **12**(4), 345-364.

Mosimann, S.C., Gilbert, M., Dombroswki, D., To, R., Wakarchuk, W., and Strynadka, N.C. (2001) Structure of a sialic acid-activating synthetase, CMP-acylneuraminate synthetase in the presence and absence of CDP. *J.Biol.Chem.* **276**(11), 8190-8196.

Mühlenhoff, M., Eckhardt, M., and Gerardy-Schahn, R. (1998) Polysialic acid: three-dimensional structure, biosynthesis and function. *Curr.Opin.Struct.Biol.* **8**(5), 558-564.

Münster, A.K., Eckhardt, M., Potvin, B., Mühlenhoff, M., Stanley, P., and Gerardy-Schahn, R. (1998) Mammalian cytidine 5'-monophosphate N-acetylneuraminic acid synthetase: a nuclear protein with evolutionarily conserved structural motifs. *Proc.Natl.Acad.Sci.U.S.A* **95**(16), 9140-9145.

Münster, A.K., Weinhold, B., Gotza, B., Mühlenhoff, M., Frosch, M., and Gerardy-Schahn, R. (2002) Nuclear localization signal of murine CMP-Neu5Ac synthetase includes residues required for both nuclear targeting and enzymatic activity. *J.Biol.Chem.* **277**(22), 19688-19696.

Münster-Kühnel, A.K., Tiralongo, J., Krapp, S., Weinhold, B., Ritz-Sedlacek, V., Jacob, U., and Gerardy-Schahn, R. (2004) Structure and function of vertebrate CMP-sialic acid synthetases. *Glycobiology* **14**(10), 43R-51R.

Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S., and Dodson, E.J. (1999b) Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr.D.Biol.Crystallogr.* **55**(Pt 1), 247-255.

Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S., and Dodson, E.J. (1999a) Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr.D.Biol.Crystallogr.* **55**(Pt 1), 247-255.

Nadano, D., Iwasaki, M., Endo, S., Kitajima, K., Inoue, S., and Inoue, Y. (1986) A naturally occurring deaminated neuraminic acid, 3-deoxy-D-glycero-D-galacto-nonulosonic acid (KDN). Its unique occurrence at the nonreducing ends of oligosialyl chains in polysialoglycoprotein of rainbow trout eggs. *J.Biol.Chem.* **261**(25), 11550-11557.

Nagradova, N. (2003) Interdomain communications in bifunctional enzymes: how are different activities coordinated? *IUBMB.Life* **55**(8), 459-466.

Nakata, D., Close, B.E., Colley, K.J., Matsuda, T., and Kitajima, K. (2000) Molecular cloning and expression of the mouse N-acetylneuraminic acid 9-phosphate synthase which does not have deaminoneuraminic acid (KDN) 9-phosphate synthase activity. *Biochem.Biophys.Res.Commun.* **273**(2), 642-648.

Nakata, D., Münster, A.K., Gerardy-Schahn, R., Aoki, N., Matsuda, T., and Kitajima, K. (2001) Molecular cloning of a unique CMP-sialic acid synthetase that effectively utilizes both deaminoneuraminic acid (KDN) and N-acetylneuraminic acid (Neu5Ac) as substrates. *Glycobiology* **11**(8), 685-692.

Nigg, E.A. (1997) Nucleocytoplasmic transport: signals, mechanisms and regulation. *Nature* **386**(6627), 779-787.

Parsons, J.F., Lim, K., Tempczyk, A., Krajewski, W., Eisenstein, E., and Herzberg, O. (2002) From structure to function: YrbI from Haemophilus influenzae (HI1679) is a phosphatase. *Proteins* **46**(4), 393-404.

Perrakis, A., Morris, R., and Lamzin, V.S. (1999) Automated protein model building combined with iterative structure refinement. *Nat.Struct.Biol.* **6**(5), 458-463.

Ridder, I. S., Rozenboom, H.J., Kalk, K. H., Janssen, D. B., and Dijkstra, B. W. (1997) Three-dimensional structure of L-2-haloacid dehalogenase from *Xanthobacter autotrophicus* GJ10 complexed with the substrate-analogue formate, *J.Biol.Che*m. **272**, 33025-33022.

Reglero, A., Rodriguez-Aparicio, L.B., and Luengo, J.M. (1993) Polysialic acids. *Int.J.Biochem.* **25**(11), 1517-1527.

Roberts, A., Lee, S.Y., McCullagh, E., Silversmith, R.E., and Wemmer, D.E. (2005) YbiV from Escherichia coli K12 is a HAD phosphatase. *Proteins* **58**(4), 790-801.

Rodriguez-Aparicio, L.B., Luengo, J.M., Gonzalez-Clemente, C., and Reglero, A. (1992) Purification and characterization of the nuclear cytidine 5'-monophosphate N-acetylneuraminic acid synthetase from rat liver. *J.Biol.Chem.* **267**(13), 9257-9263.

Roseman, S. (1962) Enzymatic synthesis of cytidine 5'-mono-phospho-sialic acids. *Proc.Natl.Acad.Sci.U.S.A* **48**437-441.

Royo, J., Gomez, E., and Hueros, G. (2000) A maize homologue of the bacterial CMP-3-deoxy-D-manno-2-octulosonate (KDO) synthetases. Similar pathways operate in plants and bacteria for the activation of KDO prior to its incorporation into outer cellular envelopes. *J.Biol.Chem.* **275**(32), 24993-24999.

Rutishauser, U. (2008) Polysialic acid in the plasticity of the developing and adult vertebrate nervous system. *Nat.Rev.Neurosci.* **9**(1), 26-35.

Schauer, R., Haverkamp, J., and Ehrlich, K. (1980) Isolation and characterization of acylneuraminate cytidylyltransferase from frog liver. *Hoppe Seylers.Z.Physiol Chem.* **361**(5), 641-648.

Schauer R., Kelm S., Reuter G., Roggentin P., and Shaw L. 1995. Biochemistry and role of sialic acids. In Biology of the sialic acids (ed. Rosenberg A.), pp. 7–67. Plenum Press, New York

Schauer, R. and Kamerling, J. P. (1997) Chemistry, Biochemistry and Biology of Sialic Acids. (29B), 243-402

Schauer, R. (2000) Achievements and challenges of sialic acid research. *Glycoconj.J.* **17**(7-9), 485-499.

Schlenzka, W., Shaw, L., Schneckenburger, P., and Schauer, R. (1994) Purification and characterization of CMP-N-acetylneuraminic acid hydroxylase from pig submandibular glands. *Glycobiology* **4**(5), 675-683.

Schwarzkopf, M., Knobeloch, K.P., Rohde, E., Hinderlich, S., Wiechens, N., Lucka, L., Horak, I., Reutter, W., and Horstkorte, R. (2002) Sialylation is essential for early development in mice. *Proc.Natl.Acad.Sci.U.S.A* **99**(8), 5267-5270.

Seidenfaden, R., Krauter, A., Schertzinger, F., Gerardy-Schahn, R., and Hildebrandt, H. (2003) Polysialic acid directs tumor cell growth by controlling heterophilic neural cell adhesion molecule interactions. *Mol.Cell Biol.* **23**(16), 5908-5918.

Selengut, J.D. (2001) MDP-1 is a new and distinct member of the haloacid dehalogenase family of aspartate-dependent phosphohydrolases. *Biochemistry* **40**(42), 12704-12711.

Severi, E., Hood, D.W., and Thomas, G.H. (2007) Sialic acid utilization by bacterial pathogens. *Microbiology* **153**(Pt 9), 2817-2822.

Sheldrick, G. M., Hauptman, H. A., Weeks, C. M., Miller, R. & Usón, I. (2001). International Tables for Crystallography, vol. F. Edited by E. Arnold, & M. Rossmann, pp. 333-351. Dordrecht: Kluwer Academic Publishers.

Shevchenko, A., Wilm, M., Vorm, O., Jensen, O.N., Podtelejnikov, A.V., Neubauer, G., Shevchenko, A., Mortensen, P., and Mann, M. (1996) A strategy for identifying gel-separated proteins in sequence databases by MS alone. *Biochem.Soc.Trans.* **24**(3), 893-896.

Spector, D.L. (2001) Nuclear domains. *J.Cell Sci.* **114**(Pt 16), 2891-2893.

Steenbergen, S.M., Lee, Y.C., Vann, W.F., Vionnet, J., Wright, L.F., and Vimr, E.R. (2006) Separate pathways for O acetylation of polymeric and monomeric sialic acids and identification of sialyl O-acetyl esterase in Escherichia coli K1. *J.Bacteriol.* **188**(17), 6195-6206.

Suck, D. and Ficner, R. (1996) Structure and function of PDC/DCoH, an enzyme with regulatory properties. FEBS Letters **389**, 35-39

Sutherland, H.G., Mumford, G.K., Newton, K., Ford, L.V., Farrall, R., Dellaire, G., Caceres, J.F., and Bickmore, W.A. (2001) Large-scale identification of mammalian proteins localized to nuclear sub-compartments. *Hum.Mol.Genet.* **10**(18), 1995-2011.

Tangvoranuntakul, P., Gagneux, P., Diaz, S., Bardor, M., Varki, N., Varki, A., and Muchmore, E. (2003) Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proc.Natl.Acad.Sci.U.S.A* **100**(21), 12045-12050.

Terry, L.J., Shows, E.B., and Wente, S.R. (2007) Crossing the nuclear envelope: hierarchical regulation of nucleocytoplasmic transport. *Science* **318**(5855), 1412-1416.

Terwilliger, T.C. (2000) Maximum-likelihood density modification. *Acta Crystallogr.D.Biol.Crystallogr.* **56**(Pt 8), 965-972.

Terwilliger, T.C. (2003) SOLVE and RESOLVE: automated structure solution and density modification. *Methods Enzymol.* **374**22-37.

Tiralongo, J., Fujita, A., Sato, C., Kitajima, K., Lehmann, F., Oschlies, M., Gerardy-Schahn, R., and Münster-Kühnel, A.K. (2007) The rainbow trout CMP-sialic acid synthetase utilises a nuclear localization signal different from that identified in the mouse enzyme. *Glycobiology* **17**(9), 945-954.

Traving, C. and Schauer, R. (1998) Structure, function and metabolism of sialic acids. *Cell Mol.Life Sci.* **54**(12), 1330-1349.

Tullius, M.V., Munson, R.S., Jr., Wang, J., and Gibson, B.W. (1996) Purification, cloning, and expression of a cytidine 5'-monophosphate N-acetylneuraminic acid synthetase from Haemophilus ducreyi. *J.Biol.Chem.* **271**(26), 15373-15380.

van Dijk, W., Ferwerda, W., and van den Eijnden, D.H. (1973) Subcellular and regional distribution of CMP-N-acetylneuraminic acid synthetase in the calf kidney. *Biochim.Biophys.Acta* **315**(1), 162-175.

Varki, A. (1997) Sialic acids as ligands in recognition phenomena. *FASEB J.* **11**(4), 248-255.

Varki, A. (2001) N-glycolylneuraminic acid deficiency in humans. *Biochimie* **83**(7), 615-622.

Varki, A. (2007) Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. *Nature* **446**(7139), 1023-1029.

Varki, N.M. and Varki, A. (2007) Diversity in cell surface sialic acid presentations: implications for biology and disease. *Lab Invest* **87**(9), 851-857.

Venkatachalam, K.V. (2003) Human 3'-phosphoadenosine 5'-phosphosulfate (PAPS) synthase: biochemistry, molecular biology and genetic deficiency. *IUBMB.Life* **55**(1), 1-11.

Vimr, E.R., Kalivoda, K.A., Deszo, E.L., and Steenbergen, S.M. (2004) Diversity of microbial sialic acid metabolism. *Microbiol.Mol.Biol.Rev.* **68**(1), 132-153.

Vionnet, J., Concepcion, N., Warner, T., Zapata, G., Hanover, J., and Vann, W.F. (1999) Purification of CMP-N-acetylneuraminic acid synthetase from bovine anterior pituitary glands. *Glycobiology* **9**(5), 481-487.

Viswanathan, K., Tomiya, N., Park, J., Singh, S., Lee, Y.C., Palter, K., and Betenbaugh, M.J. (2006) Expression of a functional Drosophila melanogaster CMP-sialic acid synthetase. Differential localization of the Drosophila and human enzymes. *J.Biol.Chem.* **281**(23), 15929-15940.

Wang, W., Cho, H.S., Kim, R., Jancarik, J., Yokota, H., Nguyen, H.H., Grigoriev, I.V., Wemmer, D.E., and Kim, S.H. (2002) Structural characterization of the reaction pathway in phosphoserine phosphatase: crystallographic "snapshots" of intermediate states. *J.Mol.Biol.* **319**(2), 421-431.

Warren, L. and Blacklow, R.S. (1962) The biosynthesis of cytidine 5'-monophospho-n-acetylneuraminic acid by an enzyme from Neisseria meningitidis. *J.Biol.Chem.* **237**3527-3534.

Weinhold, B., Seidenfaden, R., Rockle, I., Mühlenhoff, M., Schertzinger, F., Conzelmann, S., Marth, J.D., Gerardy-Schahn, R., and Hildebrandt, H. (2005) Genetic ablation of polysialic acid causes severe neurodevelopmental defects rescued by deletion of the neural cell adhesion molecule. *J.Biol.Chem.* **280**(52), 42971-42977.

Wu, J. and Woodard, R.W. (2003) Escherichia coli YrbI is 3-deoxy-D-manno-octulosonate 8-phosphate phosphatase. *J.Biol.Chem.* **278**(20), 18117-18123.

Yu, H., Ryan, W., Yu, H., and Chen, X. (2006) Characterization of a bifunctional cytidine 5'-monophosphate N-acetylneuraminic acid synthetase cloned from Streptococcus agalactiae. *Biotechnol.Lett.* **28**(2), 107-113.

Zapata, G., Vann, W.F., Aaronson, W., Lewis, M.S., and Moos, M. (1989) Sequence of the cloned Escherichia coli K1 CMP-N-acetylneuraminic acid synthetase gene. *J.Biol.Chem.* **264**(25), 14769-14774.

Zeleny, R., Kolarich, D., Strasser, R., and Altmann, F. (2006) Sialic acid concentrations in plants are in the range of inadvertent contamination. *Planta* **224**(1), 222-227.

Ziak, M., Kerjaschki, D., Farquhar, M.G., and Roth, J. (1999) Identification of megalin as the sole rat kidney sialoglycoprotein containing poly alpha2,8 deaminoneuraminic acid. *J.Am.Soc.Nephrol.* **10**(2), 203-209.

Appendix 1 – Abbreviations

α-	Anti-
aa	Amino acid
AP	Alkaline phosphasase
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
CE	Cytoplasmic extract
CIP	Calf intestinal alkaline phosphatase
CMP	Cytidine 5'-monophosphat
CSS	CMP-Sialic acid synthetase
CSS-CT	C-terminal domain of CSS
CSS-NT	N-terminal domain of CSS
C-terminal	Carboxy-terminal
CTP	Cytidine 5'-triphosphat
Da	Dalton; 1 Da =1 g/mol= 1,6605402*10-27 kg
DMEM	'Dulbecco's Modified Eage's medium'
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ELISA	Enzyme linked immuno sorbant assay
EtOH	Ethanol
GNE	UDP-*N*-acetylglucosamine epimerase
HAD	Haloacid dehalogenase
IPTG	Isopropyl-beta-D-thiogalactopyranoside
KDN	2-keto-3-deoxynonulosonic acid; deaminated sialic acid
KDO	3-deoxy-D-manno-octulosonic acid
mAB	Monoclonal antibody
MALDI-TOF	Matrix assisted laser desorption/ionisation – time of flight
MEF	Mouse embryo fibroblasts
MNK	*N*-acetylmannosamine kinase
Mr	Relative molecular weight
MS	Mass spectrometry
NE	Nuclear extract
Neu5Ac	5-*N*-acetyl-neuraminic acid
Neu5Gc	5-*N*-glycolyl-neuraminic acid
NPS	Neu5Ac-9P synthase
N-terminal	Amino-terminal
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEP	Phosphoenolpyruvate
PMSF	Phenylmethansulfonylflouride
pNPP	4-Nitrophenylphosphate
polySia	Polysialic acid
pox	Horseradish peroxidise
PPi	Pyrophosphate
SDS	Sodium dodecyl-sulphate
SeMet	Seleno-L-methionine
Sia	Sialic acid
U	Unit
v	Volume
w	Weight

Danksagung