

## The rainbow trout CMP-sialic acid synthetase utilises a nuclear localization signal different from that identified in the mouse enzyme

Joe Tiralongo<sup>2</sup>, Akiko Fujita<sup>3</sup>, Chihiro Sato<sup>3</sup>, Ken Kitajima<sup>3</sup>, Friederike Lehmann<sup>2</sup>, Melanie Oschlies<sup>4</sup>, Rita Gerardy-Schahn<sup>4</sup>, and Anja K Münster-Kühnel<sup>1,4</sup>

<sup>2</sup>Institute for Glycomics, Griffith University (Gold Coast Campus), PMB 50 Gold Coast Mail Centre 9726, Gold Coast, QLD, Australia; <sup>3</sup>Department of Applied Molecular Biosciences, Graduate School of Bioagricultural Sciences, Nagoya University, Nagoya 464–8601, Japan; and <sup>4</sup>Abteilung Zelluläre Chemie, Zentrum Biochemie, Medizinische Hochschule Hannover, Carl-Neuberg-Str. 1, D-30625 Hannover, Germany

Received on April 19, 2007; revised on June 7, 2007; accepted on June 7, 2007

**The terminal sugar sialic acid (Sia) plays a pivotal role in cell–cell interaction and recognition. A prerequisite for the biosynthesis of sialoglycoconjugates is the activation of Sia to cytidine monophosphate-Sia (CMP-Sia), by CMP-Sia synthetases (CMP-Sia-syn). CMP-Sia-syn are conserved from bacteria to man, and have been found to reside in the nucleus of all vertebrate species analysed to date. We previously cloned the CMP-Sia-syn from rainbow trout (rt) and identified three clusters of basic amino acids (BC) that might act as nuclear localization signals (NLS). Here, we utilised chimeric proteins and rt CMP-Sia-syn mutants in which putative NLS sequences were deleted, to identify the nuclear transport signal. Divergent from the mouse enzyme, where the crucial NLS is part of the enzyme's active site, in the rt CMP-Sia-syn the NLS and active site are disparate. The crucial NLS in the fish enzyme is bipartite and the functionality depends on a free N-terminus. Comparative analysis of all putative rt NLS in mouse and fish cells identified a second inferior motif (rtBC5–6), which was functional only in fish cells suggesting some differences in transport mechanism or folding variabilities in fish. Moreover, based on computational analyses of putative CMP-Sia-syn from distant deuterostomian organisms it was concluded that CMP-Sia-syn nuclear localization is a relatively recent invention, originating in echinoderms. In summary, our data describing structural differences in the NLS of vertebrate CMP-Sia-syn, and the independence of Sia activation from the subcellular localization of the enzyme, provide supporting evidence that nuclear localization is linked to a second yet unknown function.**

**Keywords:** CMP-sialic acid/CMP-sialic acid synthetase/nuclear localization signal/sialic acids

### Introduction

Sialic acids (Sia) are a family of 9-carbon  $\alpha$ -keto acids found predominantly at the non-reducing end of oligosaccharide chains

on glycoproteins and glycolipids. Due to their unique physico-chemical properties and exposed position, Sia have been implicated in a variety of vital biological processes, such as neural cell growth, embryogenesis and other growth processes including tumour cell metastasis (for review see Schauer and Kamerling 1997; Angata and Varki 2002). Sia functions are impressively documented in the immune system, where an evolutionary young family of Sia recognition molecules, the Siglecs, orchestrates cellular recognition and interaction (for review see Crocker 2002; Lehmann et al. 2006; Varki and Angata 2006).

Sia show remarkable structural diversity, with the family currently comprising over 50 naturally occurring members (for review see Schauer and Kamerling 1997; Angata and Varki 2002). While in mammals, 5-*N*-acetylneuraminic acid (Neu5Ac) is the predominant derivative, deaminoneuraminic acid (KDN) represent the major derivative in fish. The biosynthetic pathway leading to KDN seems to be very similar to that leading to Neu5Ac (Terada et al. 1993; Angata et al. 1999); however, the primary structures of many of the enzymes involved remain to be resolved (for review see Inoue and Kitajima, 2006).

A prerequisite for the incorporation of Sia into glycoconjugates is its activation to the cytidine monophosphate diester (CMP-Sia), a reaction catalysed by the enzyme CMP-Sia synthetase (CMP-Sia-syn), which is conserved from bacteria to man (for review see Münster-Kühnel et al. 2004). An unusual feature of all identified vertebrate CMP-Sia-syn is their localization in the nuclear compartment (for review see Kean 1991; Münster-Kühnel et al. 2004); all other nucleotide sugar synthetases are restricted to the cytoplasm. Translocation of mature proteins to the nuclear compartment often depends on the existence of short clusters of basic amino acids (BC) that act as nuclear localization signals (NLS) (for review see Mattaj and Englmeier 1998; Goldfarb et al. 2004). Recombinant expression of murine CMP-Sia-syn confirmed nuclear compartmentalization (Münster et al. 1998) and a basic cluster (K<sup>198</sup>RPRR) that fits well with the four residue motif K-K/R-X-K used in database searches has been identified as NLS. Interestingly, this monopartite NLS harbours elements of the enzyme's active site (Münster et al. 2002). Additionally, data obtained from X-ray crystallography showed that the NLS constitutes part of the catalytic and dimerization domain of the enzyme (Krapp et al. 2003). Mutation of each individual amino acid within the identified NLS showed that K<sup>198</sup> plays a crucial role in nuclear import. Importantly, nuclear transport could be prevented without affecting the enzyme's capacity to activate Sia in the cellular context. This observation that nuclear compartmentalization and synthetase activity are mutually exclusive functions, but mediated by the common sequence K<sup>198</sup>RPRR (Münster et al. 2002), led to the speculation that nuclear localization may be solely a consequence of the basic structure of the enzyme's active site.

<sup>1</sup>To whom correspondence should be addressed: Tel: +49-511-532-3367; Fax: +49-511-532-3956; e-mail: Muenster.Anja@mh-hannover.de

More recently, the CMP-Sia-syn from rainbow trout (rt) testis was isolated (Nakata et al. 2001). This enzyme, which due to the prevalence of KDN in the fish tissue forms mainly KDN cytidine-monophosphate diesters (CMP-KDN), shows 53.8% identity with the mouse enzyme, and like its mouse counterpart possesses clusters of BC that could act as NLS sequences. However, the sequence corresponding to the mouse NLS (K<sup>198</sup>RPRR) is altered to the sequence N<sup>187</sup>RPRR in fish, meaning that the lysine residue (K<sup>198</sup> in mouse) shown to be essential for nuclear transport of the murine enzyme is missing in the rt CMP-Sia-syn (Nakata et al. 2001).

With the objective of better understanding the biological significance of nuclear compartmentalization of the vertebrate CMP-Sia-syn, we show in this study that the recombinant rt enzyme is—as expected—localized in the nucleus. A systematic analysis of all identified putative NLS in the rt CMP-Sia-syn revealed that nuclear transport of the rt enzyme is mediated by an N-terminally located bipartite NLS that is not part of the enzyme's active site. Consequently, our data strongly suggest that nuclear transport of CMP-Sia-syn in vertebrate cells may have an autonomous biological function.

**Results**

*Identification of potential nuclear localization signals*

Nakata et al. (2001), based on amino acid sequence analyses, suggested that the rt CMP-Sia-syn resides in the cell nucleus. Computational analyses using PSORTII (Nakai and Horton 1999) identified a number of basic clusters within the rt CMP-KDN-syn (designated here as rtBC; Figure 1) that could potentially represent NLS sequences. The most N-terminal BC, rtBC1 (K<sup>5</sup>KRTQSDIEDVRDRKAK) represents a bipartite motif made up of two smaller BCs (rtBC2–rtBC3/4). The subsequent BC (rtBC6, N<sup>187</sup>RPRR) co-locates with the NLS of the murine enzyme, and comprises an amino acid stretch shown to be part of the murine enzyme's active site (Münster et al. 2002; Krapp et al. 2003). Since NLS sequences are often preceded by proline residues rtBC5 (P<sup>185</sup>CNRPRR) was defined to include P<sup>185</sup>. The C-terminal rtBC7 (K<sup>415</sup>KKAK) was identified by eye. Initial experiments were designed to investigate the ability of these sequences to direct the enhanced green fluorescent protein (eGFP) to the cell nucleus.

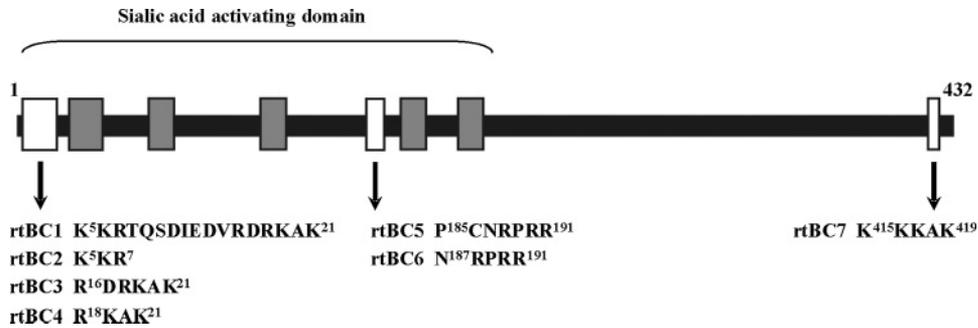
It is known from previous studies that recombinant eGFP can enter the nuclear compartment (Guo et al. 2000; Münster et al.

2002), leading to an even distribution of green fluorescence throughout the cell. In an attempt to reduce the resulting nuclear background staining, we increased the protein's size by generating an eGFP-eGFP chimera. Fusion of the two eGFP proteins was mediated via a flexible five alanine spacer. Transient expression of the chimera in NIH 3T3 cells (Figure 2, upper panels) however, demonstrated that this measure was not sufficient to prevent nuclear entry. Clearly, the overall protein structure, not merely molecular weight, determines whether a protein can access the nucleus. Further, it is also feasible that over-expression may lead to mis-expression depending on protein structure. However, fusion of the simian virus SV40 large T antigen (SV40-T) NLS (P<sup>126</sup>KKKRKV) to eGFP-eGFP provided the expected positive control upon transient expression in NIH 3T3 cells (Figure 2; bottom panel) (Kalderon et al. 1984; Münster et al. 2002). Like eGFP-eGFP-SV40-T, the fusion protein eGFP-eGFP-BC1, was transported to the nuclear compartment, while all other chimeras tested mirrored the staining obtained with eGFP-eGFP. The finding that rtBC5 and rtBC6 were unable to direct eGFP-eGFP to the mammalian nucleus is in perfect agreement with our previous study demonstrating that K<sup>198</sup> in the murine CMP-Sia-syn is indispensable for nuclear entry (Münster et al. 2002).

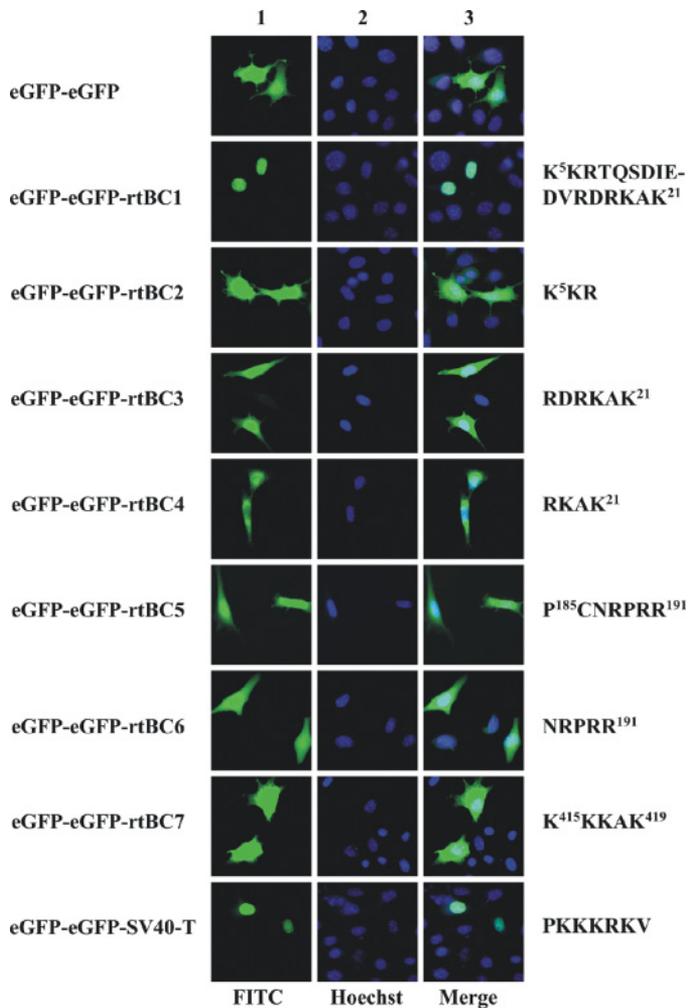
To ensure that in fish and mouse cells the same process mediates NLS-dependent nuclear import, localization studies were repeated in the rainbow trout cell line, RTG-2. Although not as clear-cut as the experiments performed in NIH 3T3 cells, Figure 3 shows that in addition to the positive control SV40-T and rtBC1, the basic clusters rtBC2, rtBC5 and rtBC6 were also able to accumulate eGFP-eGFP in the nuclear compartment. This accumulation was not observed in the negative control (eGFP-eGFP without fusion peptide) nor in cells transfected with vectors that allowed for the expression of eGFP-eGFP C-terminally extended with rtBC3, rtBC4 and rtBC7. The divergent data obtained in RTG-2 cells may either reflect differences between the two nuclear import systems, or could be due to variations in the folding and/or processing of proteins that may alter the presentation of the NLS motifs investigated.

*Complete rtBC1 sequence is necessary to target the rt CMP-Sia-syn to the nucleus*

To illustrate the intracellular localization of the rt CMP-Sia-syn, and to unequivocally determine the importance of rtBC1 through rtBC7 in nuclear transport, wild-type and deletion mutants of the rt CMP-Sia-syn were analysed in RTG-2 cells. As predicted,



**Fig. 1.** A schematic representation of the rt CMP-Sia-syn. The positions of the basic clusters (BC) are shown as open boxes with the corresponding amino acid sequences indicated. Individual monopartite segments (rtBC2–4) of the bipartite BC (rtBC1) used in subsequent experiments are shown. The five primary sequence motifs conserved in bacterial and vertebrate CMP-Sia synthetases are represented by gray boxes, the Sia-activating domain is depicted in parenthesis.

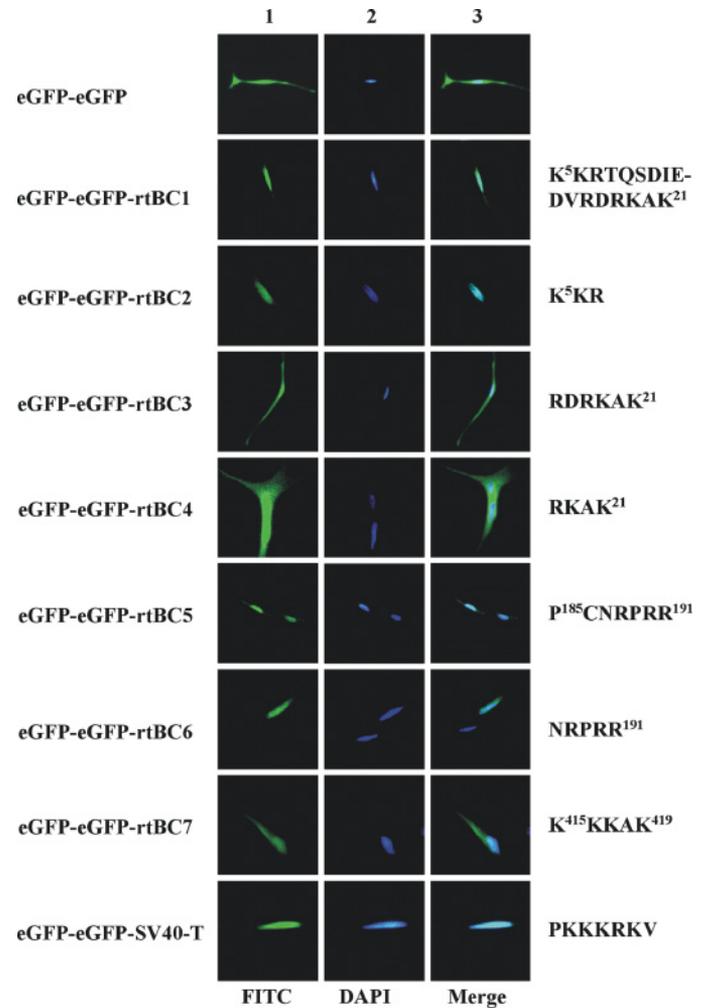


**Fig. 2.** The intracellular localization of eGFP-eGFP fusion proteins in NIH 3T3 cells. BCs of the rt CMP-Sia-syn or SV40-T antigen NLS were fused individually to the C-terminus of eGFP-eGFP (Table I). NIH 3T3 cells were transiently transfected with cDNA of either eGFP-eGFP alone or eGFP-eGFP fusion proteins (20% transfection efficiency). Intracellular localization was analysed by direct immunofluorescence microscopy at 400 $\times$  magnification (panel 1). Cell nuclei were stained with Hoechst 33258 (panel 2). The images are merged in panel 3. Representative images are shown.

the wild-type rt CMP-Sia-syn was localized to the nucleus upon transient expression in RTG-2 cells (Figure 4, upper panel). The same patterns were observed for the mutants,  $\Delta$ rtBC5,  $\Delta$ rtBC6 and  $\Delta$ rtBC7. In contrast, nuclear transport was abolished in the mutants,  $\Delta$ rtBC1,  $\Delta$ rtBC2,  $\Delta$ rtBC3 and  $\Delta$ rtBC4. This clearly shows that the intact rtBC1 is necessary for nuclear targeting of the rt CMP-Sia-syn, with deletion of any portion of the bipartite sequence resulting in a cytoplasmic protein. The deletion of  $\Delta$ rtBC5 and  $\Delta$ rtBC6, as well as the deletion of  $\Delta$ rtBC7, had no effect on the nuclear transport of the rt CMP-Sia-syn.

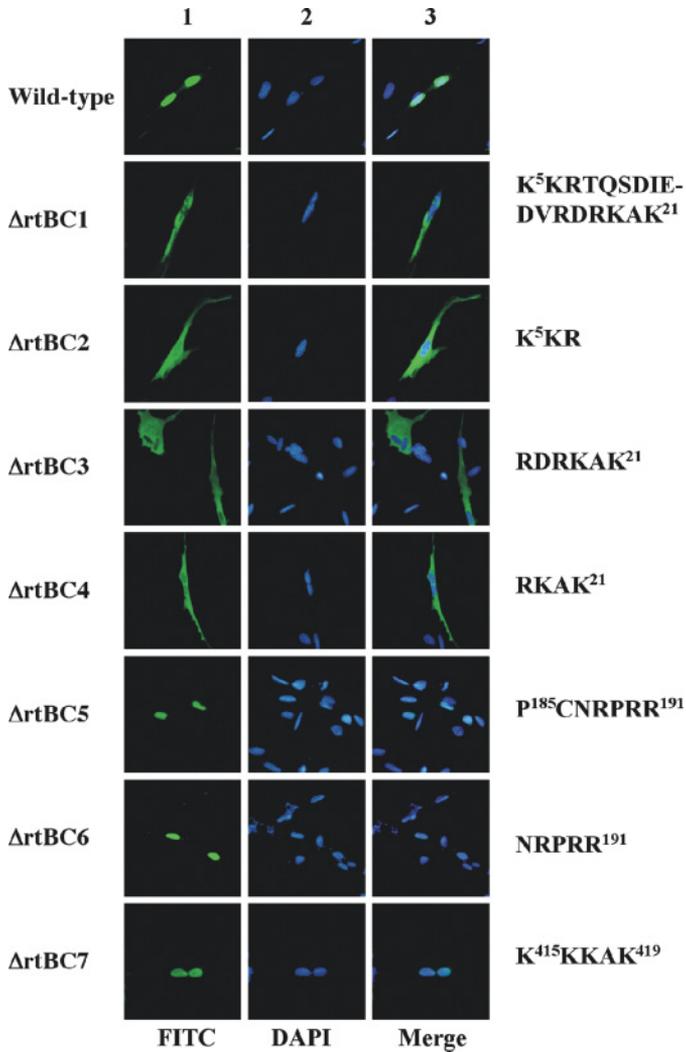
#### *Cytoplasmic forms of the rt CMP-Sia-syn are catalytically active*

To analyse whether sugar activation is linked to an intact NLS in rt CMP-Sia-syn, the enzymatic activity of cytoplasmic and nuclear localized mutants (Figure 4) was analysed in *Escherichia*



**Fig. 3.** The intracellular localization of eGFP-eGFP fusion proteins in RTG-2 cells. BCs of the rt CMP-Sia-syn or SV40-T antigen NLS were fused individually to the C-terminus of eGFP-eGFP (Table I). RTG-2 cells were transiently transfected with cDNAs of either eGFP-eGFP alone or eGFP-eGFP fusion proteins (10% transfection efficiency). Intracellular localization was analysed by direct immunofluorescence microscopy at 400 $\times$  magnification (panel 1). Cell nuclei were stained with DAPI (panel 2). The images are merged in panel 3. Representative images are shown.

*coli* K1 strain EV5. In this cellular system, a genetic defect inactivates endogenous CMP-Sia-syn, leading to an asialophenotype (Vimr et al. 1989). The ability of rt CMP-Sia-syn mutants to complement this phenotype was determined using the mAb 735 that specifically binds  $\alpha$ 2,8-linked polysialic acid (polySia) (Frosch et al. 1985). In *E. coli* K1, polySia is bound to a lipid anchor (Bliss and Silver, 1996) and migrates as a typical smear between 50 and 100 kDa (Figure 5). Endosialidase (endoNE) treatment, which removes polySia, was used to demonstrate the specificity of the mAb 735. As seen in Figure 5A, the wild-type rt CMP-Sia-syn rescues the asialophenotype. The deletion mutants that resulted in cytoplasmic localization of the rt CMP-Sia-syn in fish cells (Figure 4) were shown to be fully active in EV5 cells (Figure 5). However, deletion of rtBC5 and rtBC6, which did not alter the enzymes localization, rendered the rt CMP-Sia-syn inactive and unable to complement the asialophenotype. Deletion of rtBC7, in contrast, had no effect



**Fig. 4.** Intracellular localization in RTG-2 cells of rt CMP-Sia-syn mutants lacking individual basic clusters ( $\Delta$ rtBC, Table II). Wild-type and mutant constructs were transiently expressed in RTG-2 cells (10% transfection efficiency). C-terminally Myc/ V5-tagged CMP-Sia-syn and the different mutant forms were localized by immunostaining with anti-Myc mAb 9E10 and a FITC-conjugated anti-mouse secondary antibody (*panel 1*). Nuclear staining was performed with DAPI (*panel 2*). The images are merged in *panel 3* (400 $\times$  magnification). Representative images are shown.

on synthetase activity. All mutants were expressed at the level of the wild-type rt CMP-Sia-syn (Figure 5B).

*A free N-terminus is essential for the nuclear import of the rt CMP-Sia-syn*

In addition to the deletion of rtBC1, the fusion of an N-terminal Myc/Flag tag also prevented nuclear transport of rt CMP-Sia-syn in NIH 3T3 cells (Figure 6A). Nuclear transport was restored when the N-terminal tag was substituted by a C-terminal Myc/V5 tag (Figure 6A), indicating that recognition of rtBC1 by the cells nuclear transport machinery depends on a freely accessible N-terminus. Functionality of both epitope tagged rt CMP-Sia-syn constructs was confirmed *in vivo* by complementation of LEC29.Lec32 cells (Figure 6B) as previously described (Münster et al. 2002). As with EV5 cells, the mutation *lec32* in-

activates the CMP-Sia-syn gene in CHO cells leading to an asialophenotype (Potvin et al. 1995). Again, the reappearance of polySia (diffuse band above 100 kDa) that in vertebrate cells is added to NCAM, was used to monitor activity in transfected LEC29.Lec32 cells. As shown in Figure 6 both, nuclear and cytoplasmic forms of the rt CMP-Sia-syn were active in mammalian cells.

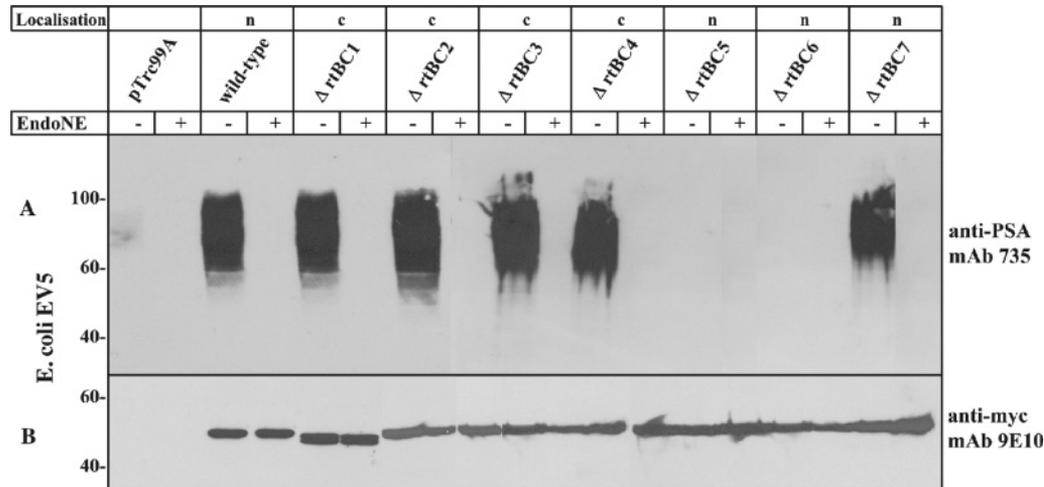
*Evolutionary conservation of CMP-Sia-syn nuclear localization in deuterostomes*

The identification of an alternative NLS in rt CMP-Sia-syn prompts the question; do putative CMP-Sia-syn from more distant deuterostomes also possess BC sequences that may be utilized as NLS and if so, are these sequences more closely related to the verified NLS in mouse or rt CMP-Sia-syn? This question is particularly important because the recently isolated CMP-Sia-syn from *Drosophila melanogaster* showed Golgi localization following recombinant expression in mammalian and insect cells (Viswanathan et al. 2006), implying that nuclear localization of CMP-Sia-syn may be a relatively recent invention.

Figure 7A summarizes verified (underlined) and putative CMP-Sia-syn NLS sequences from various deuterostomian organisms and the protostome, *D. melanogaster*. The basic clusters were identified using computational analysis with PSORTII (Nakai and Horton 1999). Amino acids shown to be critical for NLS function in the mouse and rt enzyme are highlighted in bold. All sequences listed in Figure 7 conserve five structural motifs and amino acid residues known to be critical for enzymatic activity (Münster et al. 2002). Interestingly, sequence selection on this basis listed an orthologue from *Ciona intestinalis* (sea squirt; Acc. No. Q0E670); an organism where Sia has never been detected (Warren 1963; Kawamura et al. 1991).

The multi-sequence alignment shows that elements of the murine NLS (K<sup>198</sup>RPRR) are strictly conserved from humans to sea urchins (*Strongylocentrotus purpuratus*) (Figure 7A). These include the catalytically active Arg residues 199 and 202 (shaded), as well as 201 that are also involved in nuclear transport (residues are boxed). In contrast, K<sup>198</sup> required for nuclear targeting in mammalian cells, shows variability in more distant species and is conservatively substituted in *Xenopus tropicalis* (frog) and *C. intestinalis*. Moreover, the aligned sequences show that the N-terminal BC is not present in species with a functional central BC, occurring only (with the single exception of *X. tropicalis*) in sequences where the position corresponding to K<sup>198</sup> is non-conservatively exchanged (*Oncorhynchus mykiss*, *Takifugu rubripes*, *Danio rerio* and *D. melanogaster*). These variations therefore, argue for differences in the nuclear import machinery between species belonging to the phylum Teleostei (fish) and other deuterostomes.

Another interesting feature evident from Figure 7A is that the di-amino acid motif KR (amino acid positions 6 and 7 in rtBC1) is highly conserved in all sequences harbouring the N-terminal BC (*T. rubripes*, *D. rerio* and *X. tropicalis*), and with only two exceptions (*C. intestinalis* and *D. melanogaster*) also forms part of the central BC in sequences not harbouring the N-terminal BC. Based on the observation that the *D. melanogaster* enzyme does not enter the nuclear compartment (Viswanathan et al. 2006) one can speculate that this di-amino acid motif is an indicator for nuclear transport. The phylogenetic tree illustrated in Figure 7B further supports the importance of



**Fig. 5.** *In vivo* analysis of wildtype and mutant rt CMP-Sia-syn activity in *E. coli* EV5. Wildtype and mutant constructs were expressed in *E. coli* EV5. The intracellular localization of the mutants is indicated in the top row ('n', nucleus; 'c', cytoplasm). Mock transformations were carried out with the empty vector pTrc99A. Whole cell lysates were separated by 10% SDS-PAGE (A), and the expression of polySia was monitored by Western blot analysis using mAb 735 (A). Specificity of the anti-polySia mAb was controlled in a second aliquot of the cell lysate by endoNE treatment prior to SDS-PAGE analysis (*endoNE*+). Expression levels of the recombinant proteins were analysed by 12% SDS-PAGE followed by Western blot analysis using anti-Myc mAb 9E10 (B).

the KR motif by demonstrating, that the sequences not containing this motif split off as separate branches.

## Discussion

Unlike other sugar activating enzymes that are localized in the cytoplasm, CMP-Sia-syn are found in the cell nucleus, a phenomenon that still awaits elucidation (for review see Kean 1991; Kean et al. 2004). Thus far, the primary sequence elements responsible for nuclear targeting of the CMP-Sia-syn have been identified only for the murine enzyme. This NLS (K<sup>198</sup>RPRR) not only targets the murine CMP-Sia-syn to the nucleus but also harbours amino acids important for catalytic activity (Münster et al. 2002; Krapp et al. 2003). Based on this data the possibility that nuclear targeting occurs by chance, as a consequence of the basic structure of the active site, could not be ruled out. The isolation of the human (Lawrence et al. 2001) and rt CMP-Sia-syn (Nakata et al. 2001) cDNAs now permits the analysis of nuclear localization in these species. However, because the human cDNA encodes a protein with 94% identity to the murine homologue, in which all structural and functional motifs are perfectly conserved, the intracellular localization of the more distant fish enzyme was analysed.

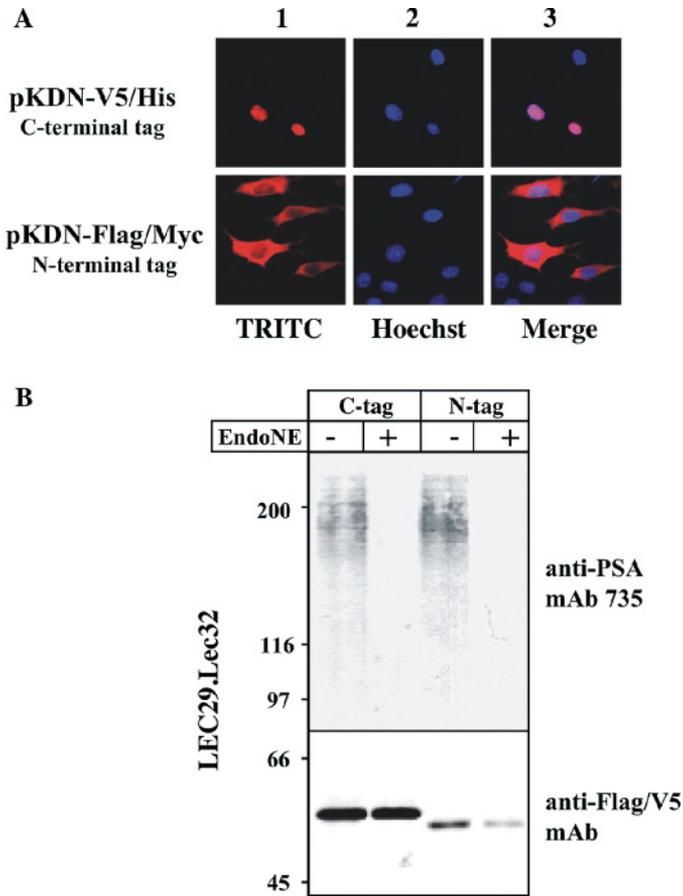
Indirect immunofluorescence analysis of rt CMP-Sia-syn deletion mutants clearly showed that the intact most N-terminal BC (rtBC1: K<sup>5</sup>KRTQSDIEDVRDRKAK) was necessary for nuclear targeting of the rt CMP-Sia-syn in fish cells (Figure 4). This result was mirrored in mammalian cells, where rtBC1 recognition and subsequent nuclear import was hindered by the addition of an N-terminal tag (Figure 6A). rtBC1 consists of two smaller BCs (rtBC2 and rtBC3/4) separated by a short spacer (Figure 1), and the observation that an intact bipartite BC is required for nuclear targeting has been made for a number of NLS sequences (Mirski et al. 1997; Munoz-Fontela et al. 2003; Hahn and Marsh 2005). In Figure 8, four confirmed bipartite NLSs are depicted, together with the consensus basic-type bipartite sequence ((K/R)<sub>2</sub>-X<sub>10-12</sub>-(K/R)<sub>3-5</sub>) (Christophe et al. 2000).

Even though some variability among known bipartite NLS sequences exists—mainly with respect to the number of residues separating the two basic clusters—the rt CMP-Sia-syn NLS identified here (rtBC1) fits well with the consensus sequence with 10 amino acids separating rtBC2 and rtBC4.

The results outlined in this report clearly show that the rt CMP-Sia-syn utilises an alternative, structurally different NLS to that previously identified for the murine CMP-Sia-syn. *In vivo* analysis of enzymatic activity of mutants lacking individual BCs also revealed that nuclear transport and enzymatic activity are independent functions (Figure 5). However, in contrast to the NLS of the murine CMP-Sia-syn (Münster et al. 2002), the fish NLS does not harbour elements of the active site. Instead, the strictly conserved arginine residues, Arg<sup>199</sup> and Arg<sup>201</sup> shown to be important for enzymatic activity and nuclear localization in the murine CMP-Sia-syn (Münster et al. 2002) (Figure 7) also constitute part of the fish enzyme's active site.

Interestingly, only the most N-terminal portion, specifically the KR residues within rtBC2 (K<sup>5</sup>KR) of the rt CMP-Sia-syn bipartite NLS, is conserved in other fish species and *X. tropicalis*, with this N-terminal motif being completely absent in other animals. The finding that rtBC2 is sufficient to direct eGFP to the nucleus, and is necessary for the nuclear import of the rt CMP-Sia-syn in fish cells suggests that the corresponding sequence in *T. rubripes* and *D. rerio* could be sufficient for nuclear targeting. However, experimental data are needed to verify this hypothesis, particularly because the possibility that the central BC supports nuclear transport must also be considered. Examples where nuclear transport is enhanced by two or more NLS sequences within a given protein have been reported (Richardson et al. 1986; Shaalsky et al. 1990).

Due to its complete conservation in the putative sea urchin CMP-Sia-syn, the minimum NLS sequence utilized by mammalian cells to target the CMP-Sia-syn to the nucleus (K<sup>198</sup>RPR) appears to have arisen following the deuterostome/protostome split, probably in echinoderms, and remained highly conserved (except for K<sup>198</sup>) in all deuterostomian species. The consequence of the conservative substitution seen in the *X. tropicalis* and



**Fig. 6.** The intracellular localization and *in vivo* activity of C- and N-terminal tagged rt CMP-Sia-syn. (A) pKDN-V5/His (C-terminal tag) and pKDN-Flag/Myc (N-terminal tag) were transiently expressed in NIH 3T3 cells. The C-terminal and N-terminal tagged rt CMP-Sia-syn were localized by immunostaining using the primary antibodies anti-V5 and anti-Flag, respectively, and a Cy3-conjugated secondary antibody (panel 1). Nuclear staining was performed with Hoechst 33258 (panel 2). The images are merged in panel 3 (400 $\times$  magnification). (B) Both constructs were transiently expressed in LEC29.Lec32 cells. Whole cell lysates were separated by 8% SDS-PAGE, and the expression of polySia was monitored by Western blot analysis using mAb 735 (upper panel). Specificity of the anti-polySia mAb was controlled in a second aliquot of the cell lysate by endoNE treatment prior to SDS-PAGE analysis (endoNE+). Expression levels of the recombinant proteins were analysed by 12% SDS-PAGE followed by Western blot analysis using both anti-Flag and anti-Myc mAbs (lower panel).

*C. intestinalis* CMP-Sia-syn central BC (Lys to Arg and His, respectively) on nuclear localization is difficult to predict. Characterization of the *C. intestinalis* enzyme, where the central BC represents the only putative NLS (see Figure 7) would be of great interest, particularly because Sia has never been detected in this organism (Warren 1963; Kawamura et al. 1991). Expression of the gene may, therefore, argue for a secondary function of the gene product inside or outside the cell nucleus not related to Sia expression.

The *E. coli* and *Streptococcus agalactiae* serotype V CMP-Sia-syn have recently been described as bifunctional enzymes, possessing both CMP-Sia-syn and acetylhydrolase activity, with the former function localized in the N-terminal domain and the latter localized in the C-terminal domain (Liu and Jin 2004; Yu et al. 2006). Although its physiological function is still unclear,

it has been proposed that the second function for the *E. coli* enzyme may be to aid in the traversal of the blood – brain barrier (Liu and Jin 2004; Yu et al. 2006). However, a closer examination of *Streptococcus* CMP-Sia-syn from different serotypes revealed a correlation between CMP-Sia-syn over-expression and down regulation of Sia O-acetylation. In fact, deacetylation of O-acetylated Sia has been observed *in vitro*, indicating that the *Streptococcus* CMP-Sia-syn may also modulate virulence and immunogenicity of the capsular polysaccharide (Lewis et al. 2006). Further, tertiary structure prediction based on the homologous search of the secondary structure (Kelley et al. 2000), identified the C-terminal half of confirmed deuterostomian CMP-Sia-syn cDNAs as putative phosphatases belonging to the haloacid dehalogenase (HAD) superfamily of hydrolases. However, trials carried out in our laboratory to identify phosphatase activity for the murine enzyme were not successful either with the wild-type CMP-Sia-syn or with mutant proteins engineered to improve phosphatase activity (unpublished data).

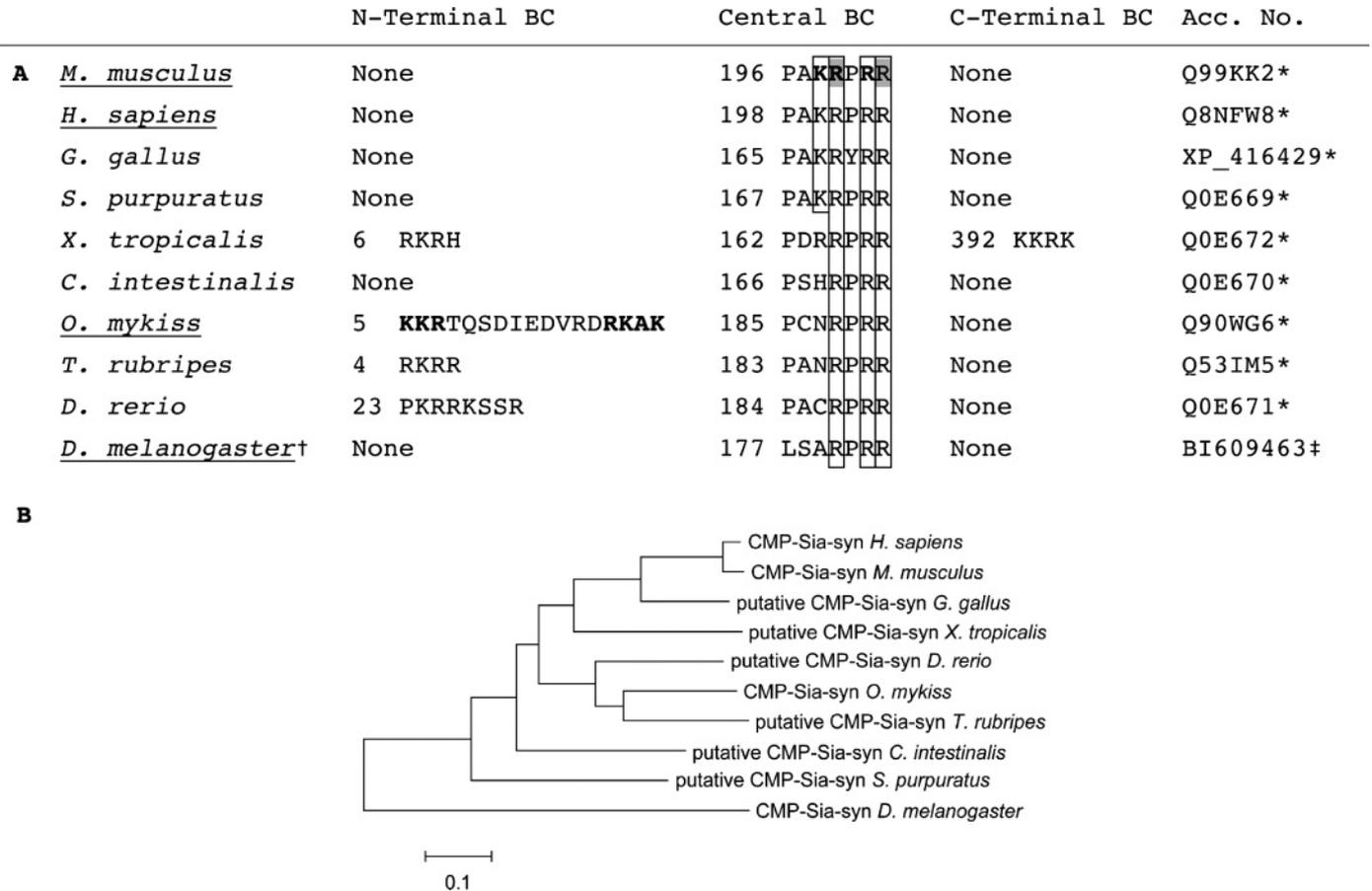
The data reported here strongly suggest that nuclear localization of the CMP-Sia-syn in higher deuterostomes is of physiological importance. Experimental data and amino acid sequence analyses of representative CMP-Sia-syn sequences from each deuterostomian phylum (Figure 7) shows that nuclear compartmentalization appears to be an evolutionary conserved characteristic of CMP-Sia-syn. Nevertheless, the delineation of the enzyme's function in the nuclear compartment remains unresolved.

Evolutionary divergent NLS sequences have been identified in other nuclear proteins. For example, the NLS mediating the nuclear transport of the *Xenopus laevis* ADAR1 (adenosine deaminase that acts on RNA) is disparate to that utilized for the import of the human ADAR1, even though both proteins share an overall identity of 51.2% (Eckmann et al. 2001). Similarly, the RNA binding La protein has evolved divergent NLS sequences (Rosenblum et al. 1998). In both cases, the divergence in NLS utilized for nuclear import appears to coincide with the attainment of additional functionality. In the case of the rt CMP-Sia-syn, the presence of a divergent NLS to that identified in the mouse CMP-Sia-syn may be more closely related to the maintenance of an additional rather than the attainment of a new function.

## Materials and methods

### Materials

Endoneuraminidase NE (endoNE), which specifically degrades  $\alpha$ 2,8-linked polysialic acid (polySia) was purified from the *E. coli* K1 bacteriophage, PK1E, as described previously (Gerardy-Schahn et al. 1995). Monoclonal antibody (mAb) 735 directed against  $\alpha$ 2,8-linked polySia (Frosch et al. 1985) was used after purification on protein A-Sepharose (GE Healthcare, Buckinghamshire, UK). Anti-Myc mAb 9E10 directed against the Myc epitope (EQKLISEEDLN) and the anti-V5 (GKPIP-PLLGLDST) antibody were purchased from Invitrogen (Carlsbad, CA). Anti-Flag mAb M5 directed against the Flag epitope (MDYKDDDDK) was purchased from Sigma-Aldrich (St Louis, MO). Horseradish peroxidase (HRP)-conjugated anti-mouse Ig was obtained from Bio-Rad (Hercules, CA).



**Fig. 7. (A)** A multi-sequence alignment was carried out with confirmed (underlined) and putative CMP-Sia-syn identified from various deuterostomian species and *D. melanogaster*. Basic clusters identified using PSORTII were grouped into N-terminal, Central and C-terminal blocks and aligned to display maximum homology. Numbers indicate the first amino acid residue of the displayed sequence. The shaded boxes highlight residues required for full enzymatic activity of the mouse CMP-Sia-syn. Both Arg residues (R<sup>199</sup> and R<sup>202</sup>) are strictly conserved, indicated by open boxes. The alignment clearly shows that two of the three residues shown to be essential for nuclear targeting (bold letters, R<sup>199</sup> and R<sup>201</sup>), are also highly conserved (open boxes). However, conservation of K<sup>198</sup> was not observed in all CMP-Sia-syn analysed. Alternative BCs, predominantly at the N-terminus, were identified by PSORTII only for those enzymes that do not possess the K residue within the Central BC. The exception to this was the *D. melanogaster* CMP-Sia-syn that was recently been found to be Golgi-localized in insect cells (Viswanathan et al. 2006), and the putative CMP-Sia-syn from the *C. intestinalis*. Thus far, the rt CMP-Sia syn NLS is the only bipartite motif identified, with the basic residues important for nuclear targeting shown in bold. European Molecular Biology Laboratory database (\*); Golgi localized (Viswanathan et al. 2006) (†); GenBank database (‡). **(B)** Phylogenetic analysis of various functionally confirmed and a selection of predicted CMP-Sia-syn. The complete protein sequences have been aligned using ClustalW algorithm (Thompson et al. 1994) and the phylogenetic tree was constructed by a minimum evolution analysis using MEGA 3.0 software (Kumar et al. 2004).

### Expression plasmids

The eukaryotic expression plasmid pJT011 comprises the full-length wild-type rt CMP-Sia-syn cDNA with C-terminal Myc- and V5-tag in a pCDNA3 vector (Invitrogen). The eukaryotic expression plasmid pKDN-Flag/Myc consists of the full-length rt CMP-Sia-syn cDNA with N-terminal Myc- and Flag-tag in pCDNA3. The eukaryotic expression plasmid pKDN-V5/His consists of the full-length rt CMP-Sia-syn cDNA with C-terminal V5- and polyHis-tag in pCDNA3.

For prokaryotic expression, the C-terminally Myc/V5-tagged CMP-Sia-syn cDNA was subcloned into pTrc99A (GE Healthcare), resulting in the plasmid pTrcJT011.

Plasmids for eukaryotic expression of C-terminally extended eGFP-eGFP fusion proteins were generated using the eukaryotic expression vector pJTeGFPx2. This vector was generated by ligation of the DNA coding for eGFP, which was amplified by PCR using the primers: 5'-GCATCCGGAGCAGCAGCAGCAATGGTGAGGAAG-

GGCGAGGAG-3' and 5'-TGCTCCGGATCTAGATCCGGTG-GATCCCGG-3', into the *Bsp*E1 site of pEGFP-C1 (Clontech, Mountain View, CA). Sense and antisense oligonucleotides that encode rtBC1 through rtBC7 of the rt CMP-Sia-syn or the SV40-T NLS were synthesized. Oligonucleotides were generated with overhanging sequences to allow for the directed cloning into *Bgl*II and *Eco*RI restriction sites as shown in Table I. Matching oligonucleotide pairs were annealed and ligated into the *Bgl*II and *Eco*RI sites of pJTeGFPx2, resulting in 3'-extended eGFP-eGFP cDNAs. The integrity of all plasmids was confirmed by sequencing.

### Construction of mutants

Site-directed mutations into pJT011 were introduced using the QuikChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) according to the manufacturer's instructions, using *Pfu* polymerase (Stratagene). Mutation primers were designed to delete nucleotide triplets encoding selected amino

Protein	NLS sequence	Reference
rt CMP-Sia-syn	<b>KKRTQSDIEDVRDRKAK</b>	
Nucleoplasmin	<b>KRPAATKKAGQAKKKK</b>	(Makkerh, et al., 1996)
SRC-1	<b>KRKGSPCDTLASSTEKRRR</b>	(Amazit, et al., 2003)
p53	<b>KRALPNNTSSSPQPKKKP</b>	(Liang and Clarke, 1999)
Parafibromin	<b>KRAADEVLAEAKKPR</b>	(Hahn and Marsh, 2005)
Consensus	<b>(K/R)<sub>2</sub>-X<sub>10-12</sub>-(K/R)<sub>3-5</sub></b>	(Christophe, et al., 2000)

**Fig. 8.** Comparison of the NLS of the rt CMP-Sia-syn with known bipartite NLS sequences found in other nuclear proteins. Basic residues required for nuclear import are indicated in bold letters. rtBC1 closely resembles the classical nucleoplasmin NLS and fits well with the consensus bipartite sequence shown.

acids. Sequences of the mutation primers together with the plasmid names are given in Table II. The integrity of all C-terminally Myc/V5-tagged mutants was confirmed by automated sequencing both, before and after subcloning. For prokaryotic expression of mutant C-terminally Myc/V5-tagged CMP-Sia-syn, cDNA of each mutant was subcloned into pTrc99A.

*Cell culture*

CHO cells of the complementation group LEC29.Lec32 were cultured in  $\alpha$ -MEM (Invitrogen). NIH 3T3 cells (ATCC CRL 1685) were maintained in Dulbecco's modified Eagle's medium (Seromed, Vienna, Austria) and rt gonadal tissue-derived RTG-2 cells (RIKEN Biosciences Center, Tsukuba, Japan) were maintained in MEM (Sigma-Aldrich) containing 100 units/mL

**Table I.** Plasmids for eukaryotic expression of eGFP-eGFP fusion proteins were generated as described in Materials and Methods. The resulting plasmid, the amino acid sequence and the corresponding oligonucleotide encoding the basic clusters are shown. The *Bgl*III and *Eco*RI extensions used for direct cloning into pJTeGFPx2 are italicised

Plasmid	aa sequence	Oligonucleotides (5'-3')
peGFP-eGFP-rtBC1	K <sup>5</sup> KRTQSDIEDVRDRKAK <sup>21</sup>	<i>GATCTAAAAAACGTACGCAATCGGACATCGAAGATGTGAGAGACAGAAAAGCAAAGTAGG</i> <i>AATTCCTACTTTGCTTTTCTGTCTCTCACATCTTCGATGTCCGATTGCGTACGTTTTTTA</i>
peGFP-eGFP-rtBC2	K <sup>5</sup> KR <sup>7</sup>	<i>GATCTAAAAAACGTTAGG</i> <i>AATTCCTAACGTTTTTT</i>
peGFP-eGFP-rtBC3	R <sup>16</sup> DRKAK <sup>21</sup>	<i>GATCTAGAGACAGAAAAGCAAAGTAGG</i> <i>AATTCCTACTTTGCTTTTCTGTCTCTA</i>
peGFP-eGFP-rtBC4	R <sup>18</sup> KAK <sup>21</sup>	<i>GATCTAGAAAAGCAAAGTAGG</i> <i>AATTCCTACTTTGCTTTTCT</i>
peGFP-eGFP-rtBC5	P <sup>185</sup> CNRPRR <sup>191</sup>	<i>GATCTCCGTGTAACAGGCCCGGAGATAGG</i> <i>AATTCCTATCTCCGGGGCCTGTTACACGGA</i>
peGFP-eGFP-rtBC6	N <sup>187</sup> RPRR <sup>191</sup>	<i>GATCTAACAGGCCCGGAGATAGG</i> <i>AATTCCTATCTCCGGGGCCTGTTA</i>
peGFP-eGFP-rtBC7	K <sup>415</sup> KKAK <sup>419</sup>	<i>GATCTAAGAAGAAGCCAAGTAGG</i> <i>AATTCCTACTTGGCCTTCTTCTTA</i>
peGFP-eGFP-SV40-T	PKKKRKV	<i>GATCTCCGAAGAAGAAGCGAAAGGTAG</i> <i>AATTCCTACTTTCGGTTCTTCTTCGGG</i>

**Table II.** Basic clusters in the rt CMP-Sia-syn were deleted using the primers indicated. The arrows shown in the primer sequence (column 3) indicate the nucleotide position where the deletion was introduced

Deletion	Deleted aa sequence	Primers (5'-3')
$\Delta$ rtBC1	A <sup>2</sup> AAKRTQSDIEDVRDRKAK <sup>21</sup>	GGGGTACCGCCACCATG↓GTTATTAAGACAGCGGTGAG T TTTGCTTTTGCGGCCGCGAATGTATTCTGTGGATGCG
$\Delta$ rtBC2	K <sup>5</sup> KR <sup>7</sup>	ACCATGGCCGCGGCA↓ACGCAATCGGACATC GTCCGATTGCGT↓TGCCGCGGCCATGGTGGC
$\Delta$ rtBC3	R <sup>16</sup> DRKAK <sup>21</sup>	GACATCGAAGATGTG↓GTTATAAAAGACAGCGGTGAG CACCGCTGTCTTTTATAAC↓CACATCTTCGATGTCC
$\Delta$ rtBC4	R <sup>18</sup> KAK <sup>21</sup>	GATGTGAGAGAC↓GTTATAAAAGACAGCGGTGAG CACCGCTGTCTTTTATAAC↓GTCTCTCACATCTTC
$\Delta$ rtBC5	P <sup>185</sup> CNRPRR <sup>191</sup>	GCCTCTGAACTTAGAC↓CAGGACTGGGATGGAGAG CTCCATCCCAGTCCCTG↓GTCTAAGTTCAGAGGCTG
$\Delta$ rtBC6	N <sup>187</sup> RPRR <sup>191</sup>	CTGAACTTAGACCCGTG↓CAGGACTTGGATGGAGAG CTCTCCATCCCAGTCCCTG↓ACACGGGTCTAAGTTCAG
$\Delta$ rtBC7	K <sup>415</sup> KKAK <sup>419</sup>	CATCCTGTGCTC↓TCTCAGATGGAACAGGAC CCTGTTCCATCTGAGA↓GAGCAGCAGGATGTG

penicillin and 0.1 mg/mL streptomycin. All media were supplemented with 1 mM sodium pyruvate and 10% fetal calf serum. All cells were maintained in a humidified 5% CO<sub>2</sub> atmosphere at 37°C.

#### *Expression of rt CMP-Sia-syn and Western blot analysis*

The functionality of wild-type and mutant rt CMP-Sia-syn was analysed in complementation studies using CHO LEC29.Lec32 and *E. coli* EV5 as previously described (Münster et al. 2002). The analysis of  $\alpha$ 2,8-linked polySia was performed by Western blotting (10% SDS-PAGE) using the mAb 735 (5  $\mu$ g/mL) and HRP-conjugated anti-mouse Ig (diluted 1:10,000). The C-terminal Myc epitope of the wild-type and mutant protein was detected by Western blotting (10% SDS-PAGE) using the anti-Myc mAb 9E10 (diluted 1:2000) and HRP-conjugated anti-mouse Ig (diluted 1:10,000). ECL (Pierce, Rockford, IL) was used as the substrate for horseradish peroxidase.

#### *Transfection of NIH 3T3 and RTG-2 cells and immunofluorescence*

Transfection of NIH 3T3 cells and subsequent direct fluorescence and indirect immunofluorescence analysis was performed essentially as described (Münster et al. 2002), except the anti-Myc mAb 9E10 (diluted 1:100) was used for indirect immunofluorescence staining.

Transfection of RTG-2 cells for direct fluorescence analysis was performed by mixing a cell suspension (1  $\times$  10<sup>6</sup>/0.2 mL PBS) with 5  $\mu$ g DNA and undergoing electroporation by 30  $\mu$ s pulse at 1.5 kV/cm on Gene Pulser Transfection Apparatus (Bio-Rad). The cells (1  $\times$  10<sup>5</sup>) were seeded onto coverslips coated with 0.01% poly-L-lysine and incubated under 5% CO<sub>2</sub> at 20°C for 3 days. Fluorescence was observed under an Olympus DP70 microscope.

Indirect immunofluorescence using RTG-2 cells was performed by culturing 2.2–3.8  $\times$  10<sup>6</sup> cells on glass coverslips, and transfecting with between 5 and 20  $\mu$ g DNA for the wild-type and mutant constructs,  $\Delta$ rtBC1,  $\Delta$ rtBC2,  $\Delta$ rtBC3,  $\Delta$ rtBC4 and  $\Delta$ rtBC7, using the Gene Pulser Xcell electroporation system (Bio-Rad). For the other mutant constructs,  $\Delta$ rtBC5 and  $\Delta$ rtBC6, 7  $\times$  10<sup>5</sup> cells were transfected with approximately 2.5  $\mu$ g of DNA using Fugene6 (Roche Diagnostics, Basel, Switzerland). After 63–72 h, cells were fixed in methanol at –20°C for 5 min. The Myc epitope was detected by sequential incubation with anti-Myc mAb 9E10 (diluted 1:100) at 37°C for 2 h and with fluorescein (FITC)-conjugated anti-mouse Ig (Seikagaku Corp., Tokyo, diluted 1:500 in PBS) at room temperature for 30 min. Nuclei were stained with DAPI at 37°C for 15 min. Fluorescence was observed under an Olympus DP70 microscope.

#### *Computational analysis*

PSORTII was performed for the prediction of protein localization sites in cells [psort.nibb.ac.jp (Nakai and Horton 1999)]. Phylogenetic analysis was performed with the complete protein sequences aligned by Vector NTI using ClustalW algorithm (Thompson et al. 1994). The phylogenetic tree was constructed by a minimum evolution analysis (CNI search level 2, gapped sites were pair-wise deleted) using MEGA 3.0 software (Kumar et al. 2004).

## Acknowledgements

J.T. gratefully acknowledges the Australian Research Council for the award of an Australian Post-Doctoral Fellowship. Prof. Eric Vimr is kindly acknowledged for providing the K1 mutant strain EV5. This research was supported by the Deutsche Forschungsgemeinschaft (DFG; GE 801/5–2 and MU 1849/1–1).

## Conflict of interest statement

None declared.

## Abbreviations

BC, basic cluster; CMP, cytidine monophosphate; CMP-Sia, CMP diester sialic acid; CMP-Sia-syn, CMP-sialic acid synthetase; eGFP, enhanced green fluorescent protein; KDN, deaminoneuraminic acid; Neu5Ac, 5-*N*-acetylneuraminic acid; NLS, nuclear localization signal; polySia, polysialic acid; rt, rainbow trout; Sia, sialic acid.

## References

- Amazit L, Alj Y, Tyagi RK, Chauchereau A, Loosfelt H, Pichon C, Pantel J, Foulon-Guinard E, Leclerc P, Milgrom E, Guiochon-Mantel A. 2003. Subcellular localization and mechanisms of nucleocytoplasmic trafficking of steroid receptor coactivator-1. *J Biol Chem.* 278:32195–32203.
- Angata T, Nakata D, Matsuda T, Kitajima K. 1999. Elevated expression of free deaminoneuraminic acid in mammalian cells cultured in mannose-rich media. *Biochem Biophys Res Commun.* 261:326–331.
- Angata T, Nakata D, Matsuda T, Kitajima K, Troy FA. 1999. Biosynthesis of KDN (2-keto-3-deoxy-D-glycero-D-galacto-nononic acid). Identification and characterization of a KDN-9-phosphate synthetase activity from trout testis. *J Biol Chem.* 274:22949–22956.
- Angata T, Varki A. 2002. Chemical diversity in the sialic acids and related  $\alpha$ -keto acids: An evolutionary perspective. *Chem Rev.* 102:439–469.
- Bliss JM, Silver RP. 1996. Coating the surface: A model for expression of capsular polysialic acid in *Escherichia coli* K1. *Mol Microbiol.* 21:221–231.
- Christophe D, Christophe-Hobertus C, Pichon B. 2000. Nuclear targeting of proteins: How many different signals? *Cell Signal.* 12:337–341.
- Crocker PR. 2002. Siglecs: Sialic-acid-binding immunoglobulin-like lectins in cell-cell interactions and signalling. *Curr Opin Struct Biol.* 12:609–615.
- Eckmann CR, Neunteufl A, Pfaffstetter L, Jantsch MF. 2001. The human but not the *Xenopus* RNA-editing enzyme ADAR1 has an atypical nuclear localization signal and displays the characteristics of a shuttling protein. *Mol Biol Cell.* 12:1911–1924.
- Frosch M, Gorgen I, Boulnois GJ, Timmis KN, 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 USA.* 82:1194–1198.
- Gerardy-Schahn R, Bethé A, Brennecke T, Mühlenhoff M, Eckhardt M, Zising S, Lottspeich F, Frosch, M. 1995. Molecular cloning and functional expression of bacteriophage PK1E-encoded endoneuraminidase. *Endo NE Mol Microbiol.* 16:441–450.
- Goldfarb DS, Corbett AH, Mason DA, Harreman MT, Adam SA. 2004. Importin alpha: A multipurpose nuclear-transport receptor. *Trends Cell Biol.* 14:505–514.
- Guo N, Faller DV, Denis GV. 2000. Activation-induced nuclear translocation of RING3. *J Cell Sci.* 113 (Pt 17), 3085–3091.
- Hahn MA, Marsh DJ. 2005. Identification of a functional bipartite nuclear localization signal in the tumor suppressor parafibromin. *Oncogene.* 24:6241–6248.
- Inoue S, Kitajima K. 2006. KDN (Deaminated neuraminic acid): Dreamful past and exciting future of the newest member of the sialic acid family. *Glycoconj J.* 23:277–290.

- Kalderon D, Richardson WD, Markham AF, Smith AE. 1984. Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature*. 311:33–38.
- Kawamura K, Nomura M, Kameda T, Shimamoto H, Nakauchi M. 1991. Self-nonsel recognition activity extracted from self-sterile eggs of the ascidian, *Ciona intestinalis*. *Develop Growth Differ*. 33:139–148.
- Kean EL. 1991. Sialic acid activation. *Glycobiology*. 1:441–447.
- Kean EL, Münster-Kühnel AK, Gerardy-Schahn R. 2004. CMP-sialic acid synthetase of the nucleus. *Biochim Biophys Acta*. 1673:56–65.
- Kelley LA, MacCallum RM, Sternberg MJ. 2000. Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J Mol Biol*. 299:499–520.
- Krapp S, Münster-Kühnel AK, Kaiser JT, Huber R, Tiralongo J, Gerardy-Schahn R, Jacob U. 2003. The crystal structure of murine CMP-5-*N*-acetylneuraminic acid synthetase. *J Mol Biol*. 334:625–637.
- Kumar S, Tamura K, Nei M. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform*. 5:150–163.
- Lawrence SM, Huddleston KA, Tomiya N, Nguyen N, Lee YC, Vann WF, Coleman TA, Betenbaugh MJ. 2001. Cloning and expression of human sialic acid pathway genes to generate CMP-sialic acids in insect cells. *Glycoconj J*. 18:205–213.
- Lehmann F, Tiralongo J. 2006. Sialic acid-specific lectins: Occurrence, specificity and function. *Cell Mol Life Sci*. 63:1331–1354.
- Lewis AL, Patel S, Diaz SL, Ryan W, Chen X, Lin FYC, Nizet V, Varki A. 2006. CMP-sialic acid synthase of group B *Streptococcus* has sialic acid *O*-acetyltransferase activity and reduces *O*-acetylation of sialic acids on the capsular polysaccharide. *Trends Glycosci Glycotech*. 18:S36–S48.
- Liang SH, Clarke MF. 1999. A bipartite nuclear localization signal is required for p53 nuclear import regulated by a carboxyl-terminal domain. *J Biol Chem*. 274:32699–32703.
- Liu G, 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*. 279:17738–17749.
- Makkerh JP, Dingwall C, Laskey RA. 1996. Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids. *Curr Biol*. 6:1025–1027.
- Mattaj JW, Englmeier L. 1998. Nucleocytoplasmic transport: The soluble phase. *Annu Rev Biochem*. 67:265–306.
- Mirski SE, Gerlach JH, Cummings HJ, Zirngibl R, Greer PA, Cole SP. 1997. Bipartite nuclear localization signals in the C terminus of human topoisomerase II alpha. *Exp Cell Res*. 237:452–455.
- Münster AK, Eckhardt M, Potvin B, Mühlenhoff M, Stanley P, Gerardy-Schahn R. 1998. Mammalian cytidine 5'-monophosphate *N*-acetylneuraminic acid synthetase: A nuclear protein with evolutionarily conserved structural motifs. *Proc Natl Acad Sci USA*. 95:9140–9145.
- Münster AK, Weinhold B, Gotza B, Mühlenhoff M, Frosch M, Gerardy-Schahn R. 2002. Nuclear localization signal of murine CMP-Neu5Ac synthetase includes residues for both nuclear targeting and enzymatic activity. *J Biol Chem*. 277:19688–19696.
- Münster-Kühnel AK, Tiralongo J, Krapp S, Weinhold B, Ritz-Sedlacek V, Jacob U, Gerardy-Schahn R. 2004. Structure and function of vertebrate CMP-sialic acid synthetases. *Glycobiology*. 14:43R–51R.
- Munoz-Fontela C, Rodríguez E, Nombela C, Arroyo J, Rivas C. 2003. Characterization of the bipartite nuclear localization signal of protein LANA2 from Kaposi's sarcoma-associated herpesvirus. *Biochem J*. 374:545–550.
- Nakai K, Horton P. 1999. PSORT: A program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci*. 24:34–36.
- Nakata D, Münster AK, Gerardy-Schahn R, Aoki N, Matsuda T, 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*. 11:685–692.
- Potvin B, Raju TS, Stanley P. 1995. Lec32 is a new mutation in Chinese hamster ovary cells that essentially abrogates CMP-*N*-acetylneuraminic acid synthetase activity. *J Biol Chem*. 270:30415–30421.
- Richardson WD, Roberts BL, Smith AE. 1986. Nuclear location signals in polyoma virus large-T. *Cell*. 44:77–85.
- Rosenblum JS, Pemberton LF, Bonifaci N, Blobel G. 1998. Nuclear import and the evolution of a multifunctional RNA-binding protein. *J Cell Biol*. 143:887–899.
- Schauer R, Kamerling JP. 1997. Chemistry, biochemistry and biology of sialic acids. In: Montreuil J, Vliegthart JFG, Schachter H, editors. *Glycoproteins II*. Amsterdam: Elsevier, pp. 243–402.
- Shaulsky G, Goldfinger N, Ben-Ze'ev A, Rotter V. 1990. Nuclear accumulation of p53 protein is mediated by several nuclear localization signals and plays a role in tumorigenesis. *Mol Cell Biol*. 10:6565–6577.
- Terada T, Kitazume S, Kitajima K, Inoue S, Ito F, Troy FA, Inoue Y. 1993. Synthesis of CMP-deaminoneuraminic acid (CMP-KDN) using the CTP: CMP-3-deoxynonulosonate cytidyltransferase from rainbow trout testis. *J Biol Chem*. 268:2640–2648.
- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 22:4673–4680.
- Varki A, Angata T. 2006. Siglecs: The major subfamily of I-type lectins. *Glycobiology*. 16:1R–27R.
- Vimr ER, Aaronson W, Silver RP. 1989. Genetic analysis of chromosomal mutations in the polysialic acid gene cluster of *Escherichia coli* K1. *J Bacteriol*. 171:1106–1117.
- Viswanathan K, Tomiya N, Singh S, Park J, Lee YC, Palter K, Betenbaugh MJ. 2006. Expression of a functional *Drosophila melanogaster* CMP-sialic acid synthetase: Differential localization of the *Drosophila* and human enzymes. *J Biol Chem*. 281:15929–15940.
- Warren L. 1963. The distribution of sialic acids in nature. *Comp Biochem Physiol*. 10:153–171.
- Yu H, Ryan W, Chen X. 2006. Characterization of a bifunctional cytidine 5'-monophosphate *N*-acetylneuraminic acid synthetase cloned from *Streptococcus agalactiae*. *Biotechnol Lett*. 28:107–113.