CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Elsevier - Publisher Connector

FEBS Letters 581 (2007) 3327–3331

Prediction of three different isoforms of the human UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase

Stefan O. Reinke, Stephan Hinderlich*

Charité – Universitätsmedizin Berlin, Campus Benjamin Franklin, Institut für Biochemie und Molekularbiologie, Arnimallee 22, 14195 Berlin-Dahlem, Germany

Received 14 May 2007; revised 6 June 2007; accepted 13 June 2007

Available online 21 June 2007

Edited by Peter Brzezinski

Abstract The bifunctional enzyme UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase (GNE) is the key enzyme of the biosynthesis of sialic acids, terminal components of glycoconjugates associated with a variety of cellular processes. Two novel isoforms of human GNE, namely GNE2 and GNE3, which possess extended and deleted N-termini, respectively, were characterized. GNE2 was also found in other species like apes, rodents, chicken or fish, whereas GNE3 seems to be restricted to primates. Both, GNE2 and GNE3, displayed tissue specific expression patterns, therefore may contribute to the complex regulation of sialic acid metabolism.
© 2007 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: GNE; Isoforms; Sialic acid; Splice variants; UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase

1. Introduction

Sialylation of glycoproteins and glycolipids on eukaryotic cell surfaces plays an important role during development and regeneration and in the pathogenesis of diseases [1]. By their high expression on cell surfaces, terminal sialic acids are involved in a variety of cell–cell interactions [2]. They are also known to be involved in the formation and masking of recognition determinants [3], and the biological stability of glycoproteins [4].

N-Acetylneuraminic acid (Neu5Ac) is the biosynthetic precursor of virtually all of the naturally occurring sialic acids [5]. In mammals, Neu5Ac and its activated nucleotide sugar CMP-Neu5Ac are synthesized from UDP-*N*-acetylglucosamine (UDP-GlcNAc) by five consecutive reactions [6]. The first two steps in this biosynthesis are catalyzed by the bifunctional enzyme UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase (UDP-GlcNAc 2-epimerase/ManNAc kinase; GNE). GNE has been recognized as the key enzyme in the biosynthetic pathway of sialic acids, as it is rate-limiting for the whole pathway and feedback inhibited by CMP-Neu5Ac [7]. The biological importance of the enzyme is further reflected in a drastic reduction of cellular sialylation upon loss of enzyme activity [8]. Furthermore the knockout of the gene in mice is embryonic lethal at day 8.5 [9]. Two human diseases are referred to point mutations in the GNE gene. Sialuria is characterized by a massive production of free Neu5Ac due to loss of the feedback control of the UDP-GlcNAc 2-epimerase activity [10]. Hereditary inclusion body myopathy is an autosomal recessive neuromuscular disorder, caused by more than 40 different mutations spreading over both functional domains of GNE [11].

Recently, four different mRNA splice variants of human GNE were described [12]. The original GNE gene [13] is complemented by an additional 90 bp exon, named Al, which is located about 20 000 bp upstream of exon 1 [12]. The four splice variants result from alternative splicing of the exons Al, 1 and 2. In this study we analyzed the splice variants for the encoded GNE proteins, and predicted three different isoforms, GNE1, GNE2 and GNE3. Furthermore, we investigated the tissue distribution of the splice variants and analyzed non-human species for the presence of the protein isoforms.

2. Materials and methods

2.1. Tissue specific reverse transcriptase (RT)-PCR

Commercial QUICK-Clone™ human placenta cDNA (BD Biosciences; Heidelberg, Germany) was used as a template for cDNA amplification of splice variant II and III. PCR Ready First Strand cDNA panels of human and mouse tissues for splicing pattern analysis were obtained from BioCat (Heidelberg, Germany). PCR reactions were performed with 2.5 ng cDNA, 5 U Taq DNA polymerase (Fermentas; St. Leon-Rot, Germany), 2.5 μl 10× Taq buffer, 0.4 mM dNTP, 1 mM $MgCl_2$, 1 μM forward primer, 1 μM reverse primer, and filled up to 25 μl with nuclease-free H_2O. Thermocycling was done in a Mastercycler EP Gradient S (Eppendorf; Hamburg, Germany) with the program: 5 min 95 °C; indicated numbers of cycles of 30 s 95 °C/30 s 60 °C/1 min 72 °C. As primers for human cDNA amplification were used: Forward-Primer hGNEl (5′-ATGGAGAAGAATGGAAA-TAACCGAAAG-3′), Forward-Primer hGNE2 (5′-AGGGTACA-GAGCTCGTGCTTCGGG-3′) and Reverse-Primer hGNE (5′-GGCAGCCTGCCAAAAGGATGC-3′). Note, that Forward-Primer hGNE2 binds in the non-coding part of exon Al, which is not displayed in Fig. 1, but in the database entry of the complete cDNA sequence. For mouse cDNA amplification Forward-Primer mGNEl (5′-ATGGAGAAGAACGGGAACAACCGAAAGCTCCGG-3′), Forward-Primer mGNE2 (5′-ATGGAAACACACGCGCATCTCC-3′) and Reverse-Primer mGNE (5′-TGACCTCGCCTCCTT-CAATG-3′) were used. The PCR products were separated by agarose gel electrophoresis and corresponding bands were excised. After gel extraction with the "QIAquick Gel Extraction Kit" (QIAGEN;

*Corresponding author. Fax: +49 30 84451541.
E-mail address: stephan.hinderlich@charite.de (S. Hinderlich).

Abbreviations: GNE, UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase; UDP-GlcNAc, UDP-*N*-acetylglucosamine; ManNAc, *N*-acetylmannosamine; Neu5Ac, *N*-acetylneuraminic acid; RT, Reverse transcriptase

0014-5793/$32.00 © 2007 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.febslet.2007.06.026

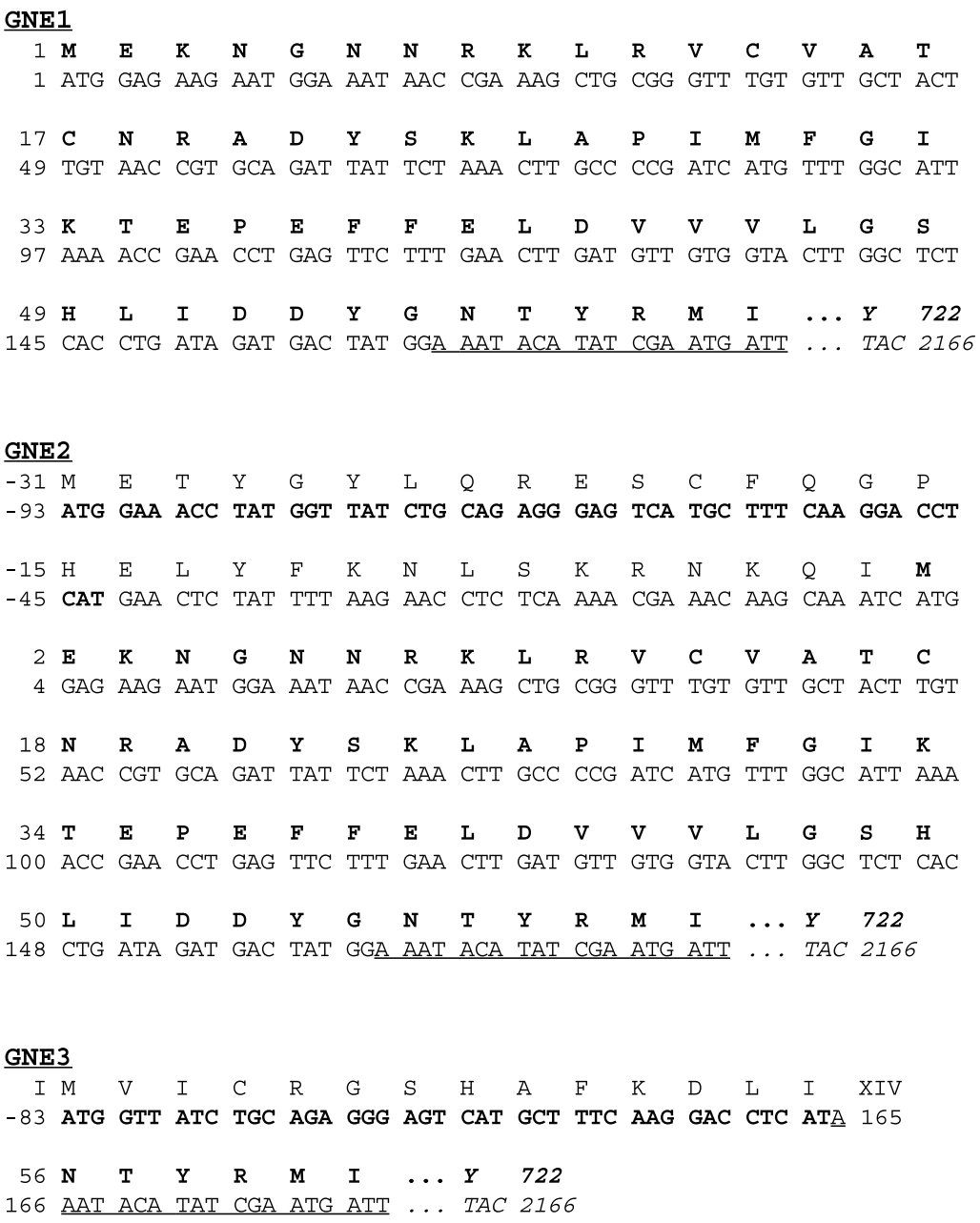

Fig. 1. N-terminal sequences of human GNE isoforms. Upper panel, GNE1; middle panel, GNE2; lower panel, GNE3. In each panel upper rows show amino acid sequences of the N-termini of proteins, lower rows show nucleotide sequences of the open reading frame 5′-ends of corresponding cDNAs. Bold amino acids are common for GNE1 and GNE2. Plain amino acids are specific for GNE2 and GNE3, respectively; GNE2 specific amino acids are numbered with negative arabic numerals, GNE3 specific amino acids are numbered with Latin numerals. Bold nucleotides derive from exon Al, plain nucleotides derive from exon 2, underlined nucleotides derive from exon 3. Amino acids and nucleotides in italics indicate the C-termini of the proteins and the 3′-ends of the open reading frames, respectively.

Hilden, Germany) the respective cDNAs were ligated into the pCR®2.1-TOPO vector (Invitrogen; Karlsruhe, Germany). Then, competent *Escherichia coli* TOP10 cells (Invitrogen) were transformed. Finally, the plasmids were isolated and the cDNA inserts were sequenced.

2.2. DNA sequencing

For the sequencing reaction the "Thermo Sequenase™ Primer Cycle Sequencing" Kit (Amersham Bioscience; Buckinghamshire, UK) was used following the manufacturers instructions. As labeled primers were used: T7-Forward (IRD 800 5′-TAATACGACTCACTATAGGG-3′) and M13-Reverse (IRD 700 5′-CAGGAAACAGCTATGACCA-TGA-3′). The following PCR thermocycling program was used: 25× (20 s 95 °C/20 s 60 °C/10 s 72 °C). The sequences were obtained by a LI-COR 4200 dual-dye DNA sequencer (MWG-Biotech; Ebersberg, Germany).

2.3. Bioinformatics

NCBI GenBank searches for genomic and cDNA sequences of diverse species were performed using BLAST (http://www.ncbi.nlm.nih.gov/BLAST). The UCSC genome server was browsed with the program Blat (http://genome.ucsc.edu/cgi-bin/hgBlat). Sequence alignments were done using MacMolly software (Softgene; Berlin, Germany).

3. Results and discussion

3.1. Identification of human GNE isoform sequences

In order to verify the novel GNE splice variants [12] independently, first of all we browsed genomic and cDNA databases. Indeed entries of cDNA sequences of the splice variants were found, which were either predicted from genomic sequences or truly isolated from cDNA libraries (data not shown). Two splice variants encode for the well-known GNE protein consisting of 722 amino acids [13,14], which was denominated GNE1. Two other splice variants contain open reading frames encoding for two GNE proteins with modified N-termini, namely GNE2 and GNE3. However, slight discrepancies in the nucleotide sequences of the database entries compared to the data published by Watts et al. [12] were observed, in particular at the exon–exon transitions. Therefore the 5′-endings of the cDNAs encoding for GNE2 and GNE3 were amplified from human placenta cDNA and ligated into the pCR®2.1-TOPO vector (data not shown). In both cases sequencing of the PCR products confirmed the sequences of the database entries.

The 5′-sequences of the open reading frames of the cDNAs derived from the splice variants, and the N-termini of GNE1, GNE2 and GNE3 are presented in Fig. 1. The complete cDNA and protein sequences can be found as database entries with the accession numbers AM697708 and AM697709. Compared to GNE1, the N-terminus of GNE2 was prolonged by 31 additional amino acids. The first 17 amino acids were encoded by exon Al, the residual 14 amino acids were encoded by the nucleotides of exon 2. The lack of exon 2 in the cDNA encoding for GNE3 leads to loss of the first 55 amino acids of GNE1. These amino acids were replaced by 14 new amino acids encoded by exon Al. Note, that for GNE2 the start codon is at position −93, but for GNE3 an alternative start codon at position −83 is used. The N-terminal amino acid sequences of GNE2 and GNE3 are therefore different.

3.2. Analysis of GNE isoforms of non-human species

Having identified GNE isoforms in humans, we were interested in knowing whether other species also express mRNAs encoding for GNE isoforms. We therefore browsed available genomic and cDNA databases of vertebrates. Proteins homologous to human GNE1 were detected in all mammals investigated and also in chicken and fish (Table 1). The amino acid sequences of GNE1 proteins have similarities higher than 95% among mammals. Chicken GNE1 reveals a similarity of 93% compared to human GNE1, whereas the sequence similarity of GNE1 of zebrafish and pufferfishes to human GNE1 is only about 80%. cDNAs encoding for GNE2 were found for all investigated apes and for mouse, rat and chicken. It is likely that such cDNAs are also present in the other mammals listed in Table 1, as sequencing and analysis of their genomes is not completed. The presence of GNE2 therefore seems to be evolutionary conserved amongst mammals. Furthermore, the existence of GNE2 in chicken argues for the emergence of GNE2 in a common ancestor. However, there is no evidence for GNE2 in fishes so far. Consequently GNE2 have occurred after separation of the mammal/bird ancestor from fishes.

The total amino acid sequence similarities of GNE2 proteins are within the same high range as for GNE1. However, comparing only the additional N-termini of GNE2, the sequence similarities of mouse, rat and chicken GNE2 compared to human GNE2 drop to less than 60%. For humans an isolated and sequenced cDNA encoding for GNE3 was found. This was not the case for other species. Nevertheless, GNE3-encoding cDNAs could be predicted from ape genomic data (Table 1).

Table 1
Sequence homology of GNE1, GNE2 and GNE3 from several species

Species	Taxonomy	GNE1	GNE2	GNE2 N-terminus	GNE3
Human	*Homo sapiens*	722/722 aa (100%)	753/753 aa (100%)	31/31 aa (100%)	681/681 aa (100%)
Chimpanzee	*Pan troglodytes*	722/722 aa (100%)	752/753 aa (99.9%)	30/31 aa (96.8%)	680/681 aa (99.9%)
Rhesus monkey	*Macaca mulatta*	721/722 aa (99.9%)	751/753 aa (99.7%)	30/31 aa 96.8%	679/681 aa (99.7%)
Orang utan	*Pongo pygmaeus*	719/722 aa (99.6%)	750/753 aa (99.6%)	31/31 aa (100%)	679/681 aa (99.7%)
Pig	*Sus scrofa*	714/722 aa (98.9%)			
Mouse	*Mus musculus*	711/722 aa (98.5%)	729/753 aa (96.8%)	18/31 aa (58.1%)	
Rat	*Rattus norvegicus*	710/722 aa (98.3%)	728/753 aa (96.7%)	18/31 aa (58.1%)	
Hamster	*Cricetulus griseus*	708/722 aa (98.1%)			
Opossum	*Monodelphis domestica*	689/722 aa (95.4%)			
Chicken	*Gallus gallus*	671/722 aa (92.9%)	688/754 aa (91.2%)	17/32 aa (53.1%)	
Zebrafish	*Danio rerio*	582/725 aa (80.3%)			
Green pufferfish	*Tetraodon nigroviridis*	570/725 aa (78.6%)			
Japanese pufferfish	*Takifugu rubripes*	552/695 aa (79.4%)			

Sequences were obtained from cDNA and genomic databases. All sequences were compared to the sequences of human GNE isoforms. Absolute and relative numbers of identical amino acids are given. GNE2 N-terminus denotes the additional amino acids compared to GNE1.

Homologies of the GNE3 proteins were as high as for the other isoforms. Although an exon similar to human exon Al was found also for mouse, rat and chicken, all these exons lack the GNE3-specific start codon (Fig. 1), and no mRNA encoding for GNE3 could be postulated. It is therefore likely, that GNE3 occurred from a point mutation after separation of primates from the other mammals, followed by generation of a separate splice variant.

3.3. Tissue distribution of mRNAs encoding for human and mouse GNE isoforms

Although GNE activity was detected in almost every mammalian tissue investigated so far, gross variations are found in the expression levels [13,14]. Therefore, RT-PCR of a cDNA panel including ten different human tissues was performed. Specific primers distinguishing between the open reading frames encoding for GNE1, GNE2 and GNE3 were used (Fig. 2). cDNAs encoding for GNE1 were found in all investigated tissues, including heart with a weak, but significant signal. GNE2-encoding cDNAs were only detected in brain, kidney, liver, placenta, lung, pancreas and colon. GNE3-encoding cDNAs, which were amplified simultaneously in the same RT-PCR, could only be observed by increased number of PCR cycles, indicating a lower expression level than GNE2-encoding cDNAs. However, kidney, liver, placenta and colon revealed significant amounts of GNE3-encoding cDNAs, suggesting a tissue-specific role of this isoform. Finally, the presence of mRNAs encoding for mouse GNE1 and GNE2 (for sequence details use accession number AM697710) were investigated by RT-PCR with a cDNA panel of five tissues (Fig. 3). As found for human tissues, all investigated samples display the presence of GNE1. Brain, colon, kidney and skeletal muscle samples also possess GNE2-encoding cDNAs. Multi-tissue Northern blot analysis revealed only one 5.2 kb band for human and one 3.2 kb band for mouse, respectively (data not shown), in accordance to former studies [13,15]. Although discrimination of mRNAs with differences of less than 100 bases in size, as it is the case for GNE1, GNE2 and GNE3 mRNAs, is tricky in Northern blots, carefully inspection of the blots gave no evidences for a prominent second band. We therefore conclude, that in all investigated tissues there is only one major mRNA, most likely for GNE1, and the other two can only be visualized by RT-PCR.

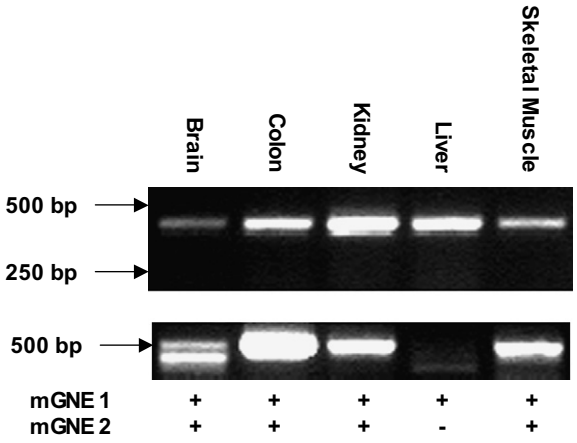

Fig. 3. Tissue distribution of mouse GNE isoforms. cDNAs from a panel of five different mouse tissues were used as templates for RT-PCR. Thirty PCR cycles amplified a 415 bp GNEl-cDNA fragment (upper panel). Forty PCR cycles amplified a 508 bp GNE2-cDNA fragment (lower panel; the prominent lower band in the brain sample was identified as syntaxin binding protein 1 by sequencing of the PCR product).

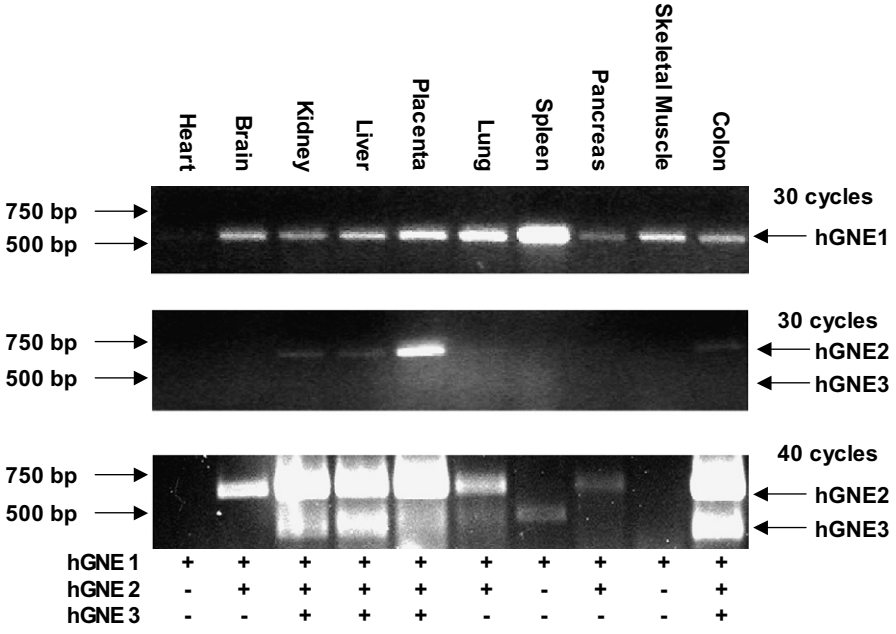

Fig. 2. Tissue distribution of human GNE isoforms. cDNAs from a panel of 10 different human tissues were used as templates for RT-PCR. Specific primers were used for amplification of GNE1 cDNA (upper panel), and GNE2 and GNE3 cDNAs (lower panels). Note, that the two lower panels only differ in the number of PCR cycles in order to visualize the GNE3 specific band at 500 bp. In the lowest panel the band for the spleen sample is larger than 500 bp and therefore does not represent GNE3. For clarity, presence or absence of the isoform-specific bands are depicted below the panels.

GNE1 was unequivocally identified as the bifunctional enzyme catalyzing the first two steps of sialic acid biosynthesis, and should therefore be sufficient to maintain the production of sialic acids in all mammalian tissues, including liver which is the major organ for biosynthesis of sialylated serum glycoproteins [14]. The exclusive role of GNE1 in sialic acid production is underlined by the lack of GNE2 mRNA in mouse liver. The isoform GNE2 therefore cannot be necessary for sialic acid production. Additionally, mouse skeletal muscle expresses GNE2 mRNA, which is not the case for human skeletal muscle. This point should also be kept in mind in particular when mouse models of human diseases caused by GNE mutations, e.g. hereditary inclusion body myopathy, are generated. Finally, the same conclusions are also true for GNE3, which is restricted to only some tissues. It should therefore taken into account, that the mRNAs for GNE2 and GNE3 are not translated to proteins, and potential functions of the mRNAs remain speculative. However, preliminary attempts of recombinant expression of GNE2 and GNE3 result in stable proteins (S.O.R., unpublished results). If this is also true in vivo, GNE2 and GNE3 may have a function in fine-tuning of the sialic acid pathway, which is involved in regulation of the sialylation of particular glycoconjugates via variation of the intracellular concentration mainly of CMP-sialic acid [16]. The extended N-terminus of GNE2, which is linked directly to the epimerase domain, likely influences the activity of this domain. This is underlined by preliminary data of recombinant GNE2, which still possesses both enzymatic activities of GNE. On the other hand, the deletion of the first 55 GNEl-amino acids in GNE3 seems to cause a total loss of the epimerase activity in recombinant GNE3. This is in agreement with a rat GNE deletion mutant lacking the first 39 amino acids [17], and the new 14 amino acids of GNE3, encoded by exon Al, are most likely unable to rescue this effect. Consequently GNE3 would possess only ManNAc kinase activity in vivo, since this function of GNE only needs an intact kinase domain [17]. It may then participate in a salvage pathway of the sialic acid metabolism, converting ManNAc derived by the action of sialate-pyruvate lyase [6,18]. Another possibility how GNE2 and GNE3 may regulate GNE function is the formation of mixed oligomers. GNE1 was found to assemble as a homotetramer [19]. Co-expression of different isoforms in one cell may lead to the formation of heterooligomers of diverse functions in a tissue specific manner. Further studies, including recombinant expression, biochemical characterization and knock-out experiments, will give more insights into the functional roles of GNE isoforms.

Acknowledgements: This work was supported by the German-Israeli Foundation for Scientific Research and Development, Jerusalem, Israel, and the Sonnenfeld-Stiftung, Berlin, Germany. We thank Felista Tansi for improving the English style of the manuscript.

References

[1] Corfield, A.P. and Schauer, R. (1982) Occurrence of sialic acids in: Sialic Acids. Chemistry, Metabolism, and Function (Schauer, R., Ed.), pp. 5–50, Springer, Wien, New York.
[2] Varki, A. (2007) Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. Nature 446, 1023–1029.
[3] Schauer, R., Kelm, S., Reuter, G., Roggentin, P. and Shaw, L. (1995) Biochemistry and role of sialic acids in: Biology of the Sialic Acids (Rosenberg, A., Ed.), pp. 7–67, Plenum Press, New York.
[4] Rens-Domiano, S. and Reisine, T.J. (1991) Structural analysis and functional role of the carbohydrate component of somatostatin receptor. Biol. Chem. 266, 20094–20102.
[5] Angata, T. and Varki, A. (2002) Chemical diversity in the sialic acids and related α-keto acids: an evolutionary perspective. Chem. Rev. 102, 439–469.
[6] Hinderlich, S., Oetke, C. and Pawlita, M. (2005) Biochemical engineering of sialic acids in: Handbook of Carbohydrate Engineering (Yarema, K.J., Ed.), pp. 387–405, Taylor & Francis, Boca Raton.
[7] 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. USA 52, 371–379.
[8] 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, 1372–1376.
[9] 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. USA 99, 5267–5270.
[10] Seppala, R., Lehto, V.P. and Gahl, W.A. (1999) Mutations in the human UDP-*N*-acetylglucosamine 2-epimerase gene define the disease sialuria and the allosteric site of the enzyme. Am. J. Hum. Genet. 64, 1563–1569.
[11] Eisenberg, I., Grabov-Nardini, G., Hochner, H., Korner, M., Sadeh, M., Bertorini, T., Bushby, K., Castellan, C., Felice, K., Mendell, J., Merlini, L., Shilling, C., Wirguin, I., Argov, Z. and Mitrani-Rosenbaum, S. (2002) Mutations spectrum of GNE in hereditary inclusion body myopathy sparing the quadriceps. Hum. Mutat. 21, 99–105.
[12] Watts, G.D.J., Thorne, M., Kovach, M.J., Pestronk, A. and Kimonis, V.E. (2003) Clinical and genetic heterogeneity in chromosome 9p associated hereditary inclusion body myopathy: exclusion of GNE and three other candidate genes. Neuromuscular Disord. 13, 559–567.
[13] Lucka, L., Krause, M., Danker, K., Reutter, W. and Horstkorte, R. (1999) Primary structure and expression analysis of human UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase, the bifunctional enzyme in neuraminic acid biosynthesis. FEBS Lett. 454, 341–344.
[14] Stäsche, R., Hinderlich, S., Weise, C., Effertz, K., Lucka, L., Moormann, P. and Reutter, W. (1997) A bifunctional enzyme catalyzes the first two steps in *N*-acetylneuraminic acid biosynthesis of rat liver: molecular cloning and functional expression of UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase. J. Biol. Chem. 272, 24319–24324.
[15] Horstkorte, R., Nöhring, S., Wiechens, N., Schwarzkopf, M., Danker, K., Reutter, W. and Lucka, L. (1999) Tissue expression and amino acid sequence of murine UDP-*N*-acetylglucosamine-2-epimerase/*N*-acetylmannosamine kinase. Eur. J. Biochem. 260, 923–927.
[16] 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, 5079–5083.
[17] 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, 599–607.
[18] Traving, C., Bruse, P., Wächter, A. and Schauer, R. (2001) The sialate-pyruvate lyase from pig kidney. Elucidation of the primary structure and expression of recombinant enzyme. Eur. J. Biochem. 268, 6473–6486.
[19] Ghaderi, D., Strauss, H.M., Reinke, S., Cirak, S., Reutter, W., Lucka, L. and Hinderlich, S. (2007) Evidence for dynamic interplay of different oligomeric states of UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase by biophysical methods. J. Mol. Biol. 369, 746–758.