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.

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
doi:10.1016/j.febslet.2007.06.026

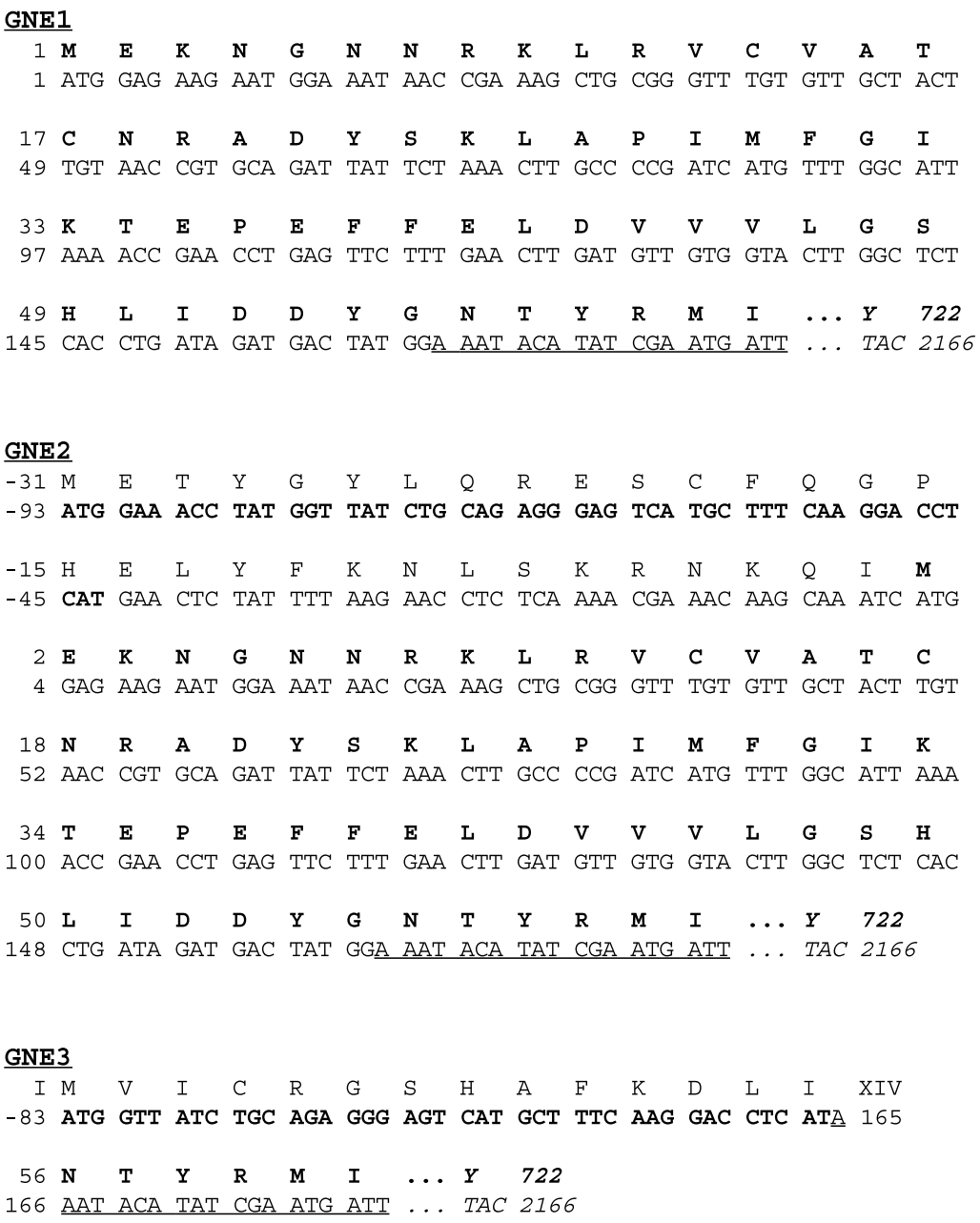

Fig. 1. N-terminal sequences of human GNE isoforms. Upper panel, GNE1; middle panel, GNE2; lower panel, GNE3. In each panel upper rows show amino acid sequences of the N-termini of proteins, lower rows show nucleotide sequences of the open reading frame 5′-ends of corresponding cDNAs. Bold amino acids 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).