Vol. 194, No. 1, 1993
July 15, 1993
BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
Pages 375-382

CLONING AND EXPRESSION OF HUMAN Galβ1,3(4)GlcNAc α2,3-SIALYLTRANSFERASE

Hiroshi Kitagawa and James C. Paulson

Cytel Corporation and the Departments of Chemistry and Molecular Biology, Scripps Research Institute, San Diego, California 92121

Received June 1, 1993

Here we report the cDNA sequence of human Galβ1,3(4)GlcNAc α2,3-sialyltransferase. The cDNA was isolated from a human placental cDNA library by screening with a 150bp probe generated by PCR using degenerate primers based on the sequences found in the sialyl motif. Comparative analylsis of this cDNA with the rat liver α2,3-sialyltransferase sequence indicates 91% nucleotide similarity between the two sequence in the predicted coding region. On the amino acid level, the degree of conservation is 97%. Surprisingly, Northern analysis indicated that the gene is expressed at low levels in human placenta but is abundantly expressed in skeletal muscle and fetal tissues.

Sialyltransferases are a family of 10-12 enzymes which catalyze the transfer of sialic acid to terminal positions on the carbohydrate groups of glycoproteins and glycolipids (1). Sialylated oligosaccharides are key determinants for a variety of cellular recognition process, such as invasiveness of a number of pathogenic organisms (2), clearance of asialo glycoproteins from circulation (3), and adhesion of leukocytes to endothelial cells mediated by selectins (4).

Recent efforts of our laboratory have resulted in the sequence elucidation of three sialyltransferases, a Galβ1,4GlcNAc α2,6-sialyltransferase, ST6N (5), a Galβ1,3GalNAc α2,3-sialyltransferase, ST3O (6) and a Galβ1,3(4)GlcNAc α2,3-sialyltrasferase, ST3N (7). Comparison of the deduced amino acid sequences has revealed a highly

0006-291X/93 $4.00

conserved region, the sialylmotif, in the catalytic domain of these enzymes (8). Degenerate PCR using primers based on the sequences found in the sialylmotif has resulted in the isolation of several new members of the sialyltransferase gene family (9). During these studies, we obtained the human homolog of the previously cloned rat ST3N using human placenta cDNA as a template in the PCR reactions.

Since it was reported earlier that the ST3N enzyme in human placenta prefered Galβ1,4GlcNAc chain acceptors (type 2) while the one in rat liver prefered Galβ1,3GlcNAc chain acceptors (type 1) (10,11), we have expressed the human enzyme and compared its specificity to the rat enzyme.

Materials and Methods

PCR cloning with degenerate oligonucleotides: Based on the sequence information of the sialylmotif (8), two degenerate oligonucleotides were synthesized (Genosys), which were predicted to yield a 150bp amplified fragment. The sequence of the 5' and 3' primers were 5'-GGAAGCTTTGSCRN MGSTGYRYCRTCGT and 5'-CCGGATCCGGTRGTYTTNSNSCCSACRTC (N=A+G+T+C, S=G+C, R=A+G, M=A+C, Y=C+T), respectively. For PCR amplification, first strand cDNA synthesized from human placenta total RNA (Clontech) was combined with 100pmol of each primer. Thirty cycles (95°C for 1 minute, 37°C for 1 minute and 73°C for 2 minutes) were run using Pfu polymerase (Stratagene) and the products were digested with BamHI and HindIII and subcloned into these sites of Bluescript SK (Stratagene). Clones were sequenced using a T7 primer.

Isolation of human ST3N cDNA: Random primed human placenta cDNA was ligated with EcoRI linkers then subsequently ligated into EcoRI digested λZAPII (Stratagene). The resultant library was packaged using a Stratagene GigapackII packaging extract and plated on *E. coli* XL-1 Blue (Stratagene). Approximately 1 million plaques were screened with the cloned PCR fragment described above. Two positive clones were plaque-purified and then excised into Bluescript vectors by *in vivo* excision with R408 helper phage.

Construction of soluble forms of human and rat ST3N: Truncated forms of human ST3N and rat one, which lack the first 61 and 60 amino acids of the open reading frame, respectively, were amplified by PCR using a 5' primer containing an in-frame BamHI site and a 3' primer located 50bp downstream of the stop codon. PCR reactions were carried out with Pfu

polymerase by 30 cycles of 95°C for 45 seconds, 55°C for 45 seconds and 73°C for 90 seconds. The fusion vector pGIR201protA was constructed by inserting a BclI/BamHI fragment, isolated from pRIT5 (Pharmacia), encoding the protein A IgG binding domain into the BamHI site of pGIR201 (a gift from Dr. K. Drickamer). The each PCR fragment was subcloned into BamHI site of pGIR201protA resulting in fusion of human ST3N or rat one to the insulin signal sequence and the protein A present in the vector. The each resulting fusion protein was inserted into the expression vector pSVL to yield the expression plasmid A3NHP or A3NR, respectively.

Expression of the soluble forms of the sialyltransferases and assaying enzyme activity: Each expression plasmid (10μg) was transfected into COS-1 cells on 100mm plates using Lipofectin (BRL). After 48h, the cell culture media was collected and incubated with IgG Sepharose (Pharmacia) for 1h. The beads were assayed for sialyltransferase activity using oligosaccharides and glycoprotein as acceptor substrates. Transfer of sialic acid to the substrate was monitored using ion-exchange or Sephadex G-50 chromatography (11).

Northern analysis: Multiple tissue northern blots of poly(A)+ RNAs were purchased from Clontech Laboratories for the analysis. The cDNA inserts of clones ST3NHP1 and ST3N-1 (7) were gel-purified, radiolabeled (>1X10^{9}cpm/mg), and used as probes, respectively.

Results and Discussion

Using human placenta cDNA as a template, PCR experiments with degenerate primers to a conserved region in the sialylmotif resulted in the amplification of the expected 150bp band which was subcloned for analysis of individual clones (**Materials and Methods**). Of fifty clones sequenced three proved to be the human homolog of the rat ST3N. In order to obtain the entire coding sequence, the human ST3N 150bp fragment was used to screen a human placenta cDNA library. Two positively hybridizing clones were isolated. Characterization of the positive clones revealed that clone ST3NHP-1 contained a 1.3Kb insert, whereas clone ST3NHP-2 was 1.1Kb in length. Sequence analysis revealed that clone ST3NHP-1 contained the complete open reading frame of the sialyltransferase. It consists of a 155-base pair 5'-untranslated region, an open reading frame 1125 base pairs in length, and a 13-base pair 3'-untranslated region. Figure 1 shows the nucleotide sequence comparison between the open

```
   1 ATGGGACTCTTGGTATTTGTGCGCAATCTGCTGCTAGCCCTCTGCCTCTTTCTGGTACTG  Human
     |||||||||||||||||||| ||||| ||||||||||||||||||||||||||||| |||
   1 ATGGGACTCTTGGTATTTGTACGCAACCTGCTGCTAGCCCTCTGCCTCTTTCTGGTCCTG  Rat

  61 GGATTTTTGTATTATTCTGCGTGGAAGCTACACTTACTCCAGTGGGAGGAGGACTCCAAT  Human
     |||||||||||||||||||| |||||||||||||||||||| |||||    |||||||||
  61 GGATTTTTGTATTATTCTGCCTGGAAGCTACACTTACTCCAATGGGAA---GACTCCAAT  Rat

 121 TCAGTGGTTCTTTCCTTTGACTCCGCTGGACAAACACTAGGCTCAGAGTATGATCGGTTG  Human
     ||| || |||||||| |||||||||||||||||||| |||||| |||||||||||| || ||
 118 TCACTGATTCTTTCCCTTGACTCCGCTGGACAAACCCTAGGCACAGAGTATGATAGGCTG  Rat

 181 GGCTTCCTCCTGAATCTGGACTCTAAACTGCCTGCTGAATTAGCCACCAAGTACGCAAAC  Human
     || |||||||||||| |||||||||||||||||||| ||  | ||||||||||||| |||
 178 GGTTTCCTCCTGAAGCTGGACTCTAAACTGCCTGCAGAGCTGGCCACCAAGTACGCTAAC  Rat

 241 TTTTCAGAGGGAGCTTGCAAGCCTGGCTATGCTTCAGCCTTGATGACGGCCATCTTCCCC  Human
     ||||| |||||||| ||||| || ||||| ||||||||| ||||||| ||||||||||||
 238 TTTTCCGAGGGAGCCTGCAAACCCGGCTACGCTTCAGCCATGATGACTGCCATCTTCCCC  Rat

 301 CGGTTCTCCAAGCCAGCACCCATGTTCCTGGATGACTCCTTTCGCAAGTGGGCTAGAATC  Human
     ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||| ||
 298 AGGTTCTCCAAGCCAGCACCCATGTTCCTGGATGACTCCTTTCGCAAATGGGCTAGGATT  Rat

 361 CGGGAGTTCGTGCCGCCTTTTGGGATCAAAGGTCAAGACAATCTGATCAAAGCCATCTTG  Human
     |||||||| ||||| || ||||||||||||||||||||||||||||||||||||||||||
 358 CGGGAGTTTGTGCCACCCTTTGGGATCAAAGGTCAAGACAATCTGATCAAAGCCATCTTG  Rat

 421 TCAGTCACCAAAGAGTACCGCCTGACCCCTGCCTTGGACAGCCTCCGCTGCCGCCGCTGC  Human
     |||||||||||||| ||||||||||||||||||||||||||||||| ||||||||||||
 418 TCAGTCACCAAAGAATACCGCCTGACCCCTGCCTTGGACAGCCTCCACTGCCGCCGCTGC  Rat

 481 ATCATCGTGGGCAATGGAGGCGTTCTTGCCAACAAGTCTCTGGGGTCACGAATTGACGAC  Human
     |||||||| |||||||||||| || || ||||||||||||||||||||||||||||||||
 478 ATCATCGTAGGCAATGGAGGGGTCCTCGCCAACAAGTCTCTGGGGTCACGAATTGACGAC  Rat

 541 TATGACATTGTGGTGAGACTGAATTCAGCACCAGTGAAAGGCTTTGAGAAGGACGTGGGC  Human
     |||||||||||| | ||| |||| |||||||| ||||| ||||||||||||||||||||
 538 TATGACATTGTGATCAGATTGAACTCAGCACCTGTGAAGGGCTTTGAGAAGGACGTGGGC  Rat

 601 AGCAAAACGACACTGCGCATCACCTACCCCGAGGGCGCCATGCAGCGGCCTGAGCAGTAC  Human
     ||||| || || |||||||||||||||| || || ||||||||||||||||||| ||
 598 AGCAAGACCACCCTGCGCATCACCTACCCTGAAGGTGCCATGCAGCGGCCTGAGCAATAT  Rat

 661 GAGCGCGATTCTCTCTTTGTCCTCGCCGGCTTCAAGTGGCAGGACTTTAAGTGGTTGAAA  Human
     || || || |||||||||||| || || |||||||||||||||||||| |||||| ||||
 658 GAACGAGACTCTCTCTTTGTACTAGCTGGCTTCAAGTGGCAGGACTTCAAGTGGCTGAAG  Rat

 721 TACATCGTCTACAAGGAGAGAGTGAGTGCATCGGATGGCTTCTGGAAATCTGTGGCCACT  Human
     |||||||||||||||||||||||||| ||||| |||||||||||||| || ||||||||
 718 TACATCGTCTACAAGGAGAGAGTGAGCGCATCCGATGGCTTCTGGAAGTCCGTGGCCACC  Rat

 781 CGAGTGCCCAAGGAGCCCCCTGAGATTCGAATCCTCAACCCATATTTCATCCAGGAGGCC  Human
     |||||||||||||||||||||||||| || |||||||||| || ||||||||||||||
 778 CGAGTGCCCAAGGAGCCCCCTGAGATCCGCATCCTCAACCCGTACTTCATCCAGGAGGCT  Rat

 841 GCCTTCACCCTCATTGGCCTGCCCTTCAACAATGGCCTCATGGGCCGGGGGAACATCCCT  Human
     |||||||| ||||| || |||||||||||||||||||||||||| | ||||||||||
 838 GCCTTCACGCTCATCGGACTGCCCTTCAACAATGGCCTCATGGGCAGAGGGAACATCCCA  Rat

 901 ACCCTTGGCAGTGTGGCAGTGACCATGGCACTACACGGCTGTGACGAGGTGGCAGTCGCA  Human
     |||||||||||||||||||||||||||||||| | |||||||| || ||||||||||||
 898 ACCCTTGGCAGTGTGGCAGTGACCATGGCACTCGATGGCTGTGATGAAGTGGCAGTCGCG  Rat

 961 GGATTTGGCTATGACATGAGCACACCCAACGCACCCCTGCACTACTATGAGACCGTTCGC  Human
     || |||||||||||||| |||||||||||| |||||||||||||||| || || |||
 958 GGCTTTGGCTATGACATGAACACACCCAACGCCCCCCTGCACTACTATGAAACTGTGCGC  Rat

1021 ATGGCAGCCATCAAAGAGTCCTGGACGCACAATATCCAGCGAGAGAAAGAGTTTCTGCGG  Human
     |||||||||||||||||||||||| ||||| ||||||||||||||||||||||||||||
1018 ATGGCAGCCATCAAAGAGTCCTGGACACACAACATCCAGCGAGAGAAAGAGTTTCTGGTG  Rat

1081 AAGCTGGTGAAAGCTCGCGTCATCACTGATCTAAGCAGTGGCATCTGA              Human
     ||||| |||||||| |||||||||||||  |||||||||| ||||||
1078 AAGCTAGTGAAAGCACGCGTCATCACTGACTTAAGCAGTGGTATCTGA              Rat
```

```
  1  MGLLVFVRNLLLALCLFLVLGFLYYSAWKLHLLQWEEDSNSVVLSFDSAG  Human
     ||||||||||||||||||||||||||||||||||||||||||| |||| || ||||
  1  MGLLVFVRNLLLALCLFLVLGFLYYSAWKLHLLQWE-DSNSLILSLDSAG  Rat

 51  QTLGSEYDRLGFLLNLDSKLPAELATKYANFSEGACKPGYASALMTAIFP  Human
     |||| ||||||||| ||||||||||||||||||||||||||||||||||| ||||||
 50  QTLGTEYDRLGFLLKLDSKLPAELATKYANFSEGACKPGYASAMMTAIFP  Rat

101  RFSKPAPMFLDDSFRKWARIREFVPPFGIKGQDNLIKAILSVTKEYRLTP  Human
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
100  RFSKPAPMFLDDSFRKWARIREFVPPFGIKGQDNLIKAILSVTKEYRLTP  Rat

151  ALDSLRCRRCIIVGNGGVLANKSLGSRIDDYDIVVRLNSAPVKGFEKDVG  Human
     ||||| |||||||||||||||||||||||||||| |||||||||||||||
150  ALDSLHCRRCIIVGNGGVLANKSLGSRIDDYDIVIRLNSAPVKGFEKDVG  Rat

201  SKTTLRITYPEGAMQRPEQYERDSLFVLAGFKWQDFKWLKYIVYKERVSA  Human
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
200  SKTTLRITYPEGAMQRPEQYERDSLFVLAGFKWQDFKWLKYIVYKERVSA  Rat

251  SDGFWKSVATRVPKEPPEIRILNPYFIQEAAFTLIGLPFNNGLMGRGNIP  Human
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
250  SDGFWKSVATRVPKEPPEIRILNPYFIQEAAFTLIGLPFNNGLMGRGNIP  Rat

301  TLGSVAVTMALHGCDEVAVAGFGYDMSTPNAPLHYYETVRMAAIKESWTH  Human
     |||||||||||| |||||||||||||| ||||||||||||||||||||||
300  TLGSVAVTMALDGCDEVAVAGFGYDMNTPNAPLHYYETVRMAAIKESWTH  Rat

351  NIQREKEFLRKLVKARVITDLSSGI                           Human
     |||||||||||||||||||||||||
350  NIQREKEFLRKLVKARVITDLSSGI                           Rat
```

Figure 2. Comparison of the deduced amino acid sequences of the human and rat ST3N. The symbol | indicates that the predicted amino acid in the alignment is identical between the two sequences.

reading frame of ST3NHP-1 cDNA with the corresponding portion of the rat ST3N (7). There is 91% homology at the nucleotide sequence level, and 97% conservation between ST3NHP-1 and the rat ST3N is observed at the amino acid sequencelevel (Fig. 2). The differences include a single amino acid insert (Glu) in the stem region of the human protein. This insertion parallels a similar finding for the human α2,6-sialyltransferase which encoded three additional residues E-K-K in the stem region compared to the rat one (12).

In order to facilitate the functional analysis of the clone, we sought to express the gene in COS-1 cells. As done earlier for previously cloned sialyltransferases (6,7), it was desirable to produce a soluble form of the enzyme which when expressed would be secreted from the cell.

Figure 1. Comparison between the open reading frame of the cDNA sequence encoding the human ST3N with the corresponding portion of the rat ST3N.

Accordingly, rat and human ST3N fusion proteins were constructed replacing the signal-anchor sequence at the amino terminus of the sialyltransferase with the cleavable insulin signal sequence and in addition the protein A IgG binding domain (13) to allow purification on IgG Sepharose (**Materials and Methods**). When the expression plasmids, A3NHP (human) and A3NR (rat), were expressed in COS-1 cells, an approximately 80Kd protein was secreted from each transformant (data not shown) which exhibited sialyltransferase activity. Next, to characterize the substrate specificity of these fusion proteins, they were purified on IgG Sepharose and assayed for sialyltransferase activity against a panel of acceptor substrates. As shown in Table 1, there was no significant difference between them. These results indicate that human ST3N enzyme is quite similar to rat one which found to preferentially act on type 1 chain (Galβ1,3GlcNAc), but which can also catalyze the

TABLE 1. Comparison of acceptor specificity of the human and rat ST3N

Acceptor (0.2mM)*	Relative activity (%)	
	Human	Rat
Galβ1,3GlcNAcβ1,3Galβ1,4Glc (LNT)	100	100
Galβ1,3GlcNAc	48	49
Galβ1,4GlcNAcβ1,3Galβ1,4Glc (LNnT)	5	4
Galβ1,4GlcNAc	22	19
Galβ1,4Glc	9	7
Asialo-α_1-acid glycoprotein	14	9

*The activities are relative to that obtained with LNT, respectively. The rat ST3N kinetic constants for both type 1 and type 2 chains (14) differ primarily in these Km values (0.1-0.6 and 2-4 mM, respectively) and have similar relative Vmax (1.0-1.2 and 0.8-1.0, respectively). For asialo-α_1 acid glycoprotein the concentration was set at 0.2mM relative to the galactose content.

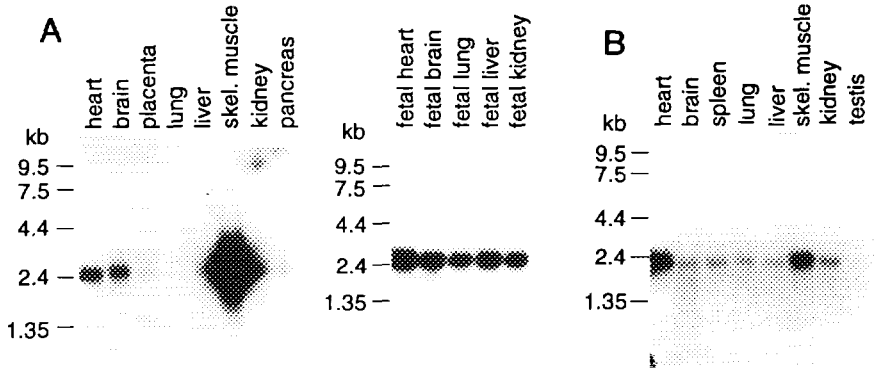

Figure 3. Differential expression of the human ST3N (A) and rat ST3N (B) in various human and rat tissues, respectively. Northen blots with RNA from various human and rat tissues were hybridized with probes for human (A) and rat (B) ST3N, respectively.

sialylation of type 2 chain (Galβ1,4GlcNAc), albeit with lower catalytic efficiency (7,11).

Expression of the human and rat ST3N mRNA in various tissues were compared as shown in Fig. 3. There appears to be a single mRNA band of approximately 2.7Kb for human ST3N and of approximately 2.4Kb for the rat enzyme. Remarkably, both genes, especially human one, are abundantly expressed in skeletal muscle whereas very low levels of mRNA were found in placenta, lung, and liver. The human ST3N was also examined for expressing in fetal human tissues and was highly expressed in all fetal tissues examined. These differential expression patterns are distinct from those observed for the other two cloned sialyltransferases, ST3O and ST6N (5,6). Presently, it is not clear why the ST3N gene is very highly expressed in skeletal muscle.

In summary the human ST3N exhibits virtually identical acceptor specificity to the previously cloned rat ST3N. This is in contrast to the previous reports for α2,3-sialyltransferase in human placenta homogenates which preferentially use the type 2 (Galβ1,4GlcNAc) sequences instead of the type 1 (Galβ1,3GlcNAc) sequences. These observations coupled with the low expression of ST3N in human placenta (Fig.3) suggested the possibility that another as yet unidentified α2,3-sialyltransferase gene may be expressed in this tissue.

Acknowledgments: This work was supported in part by United States Public Health Service Grant GM27904. We thank Dr. Brian Livingston for advice and suggestion throughout this work.

References

1. Beyer, T.A., Sadler, J.E., Rearick, J.I., Paulson, J.C., and Hill, R.L. (1981) Adv. Enzymol. 52, 23-175.
2. Paulson, J.C. (1985) in The Receptors (Conn, P.M., ed.) Vol.2, pp.131-219, Academic Press, New York.
3. Ashwell, G., and Morell, H.G. (1974) Adv. Enzymol. 41, 99-128.
4. Brandley, B.K., Swiedler, S.J., and Robbins, P.W. (1990) Cell 63, 861-863.
5. Weinstein, J., Lee, E.U., McEntee, K., Lai, P., and Paulson, J.C. (1987) J. Biol. Chem. 262, 17735-17743.
6. Gillespie, W., Kelm, S., and Paulson, J.C. (1992) J. Biol. Chem. 267, 21004-21010.
7. Wen, D.X., Livingston, B.D., Medzihradszky, K.F., Kelm, S., Burlingame, A.L., and Paulson, J.C. (1992) J. Biol. Chem. 267, 21011-21019.
8. Drickamer, K. (1993) Glycobiology 3, 2-3.
9. Livingston, B.D., Kitagawa, H., Wen, D.X., Medzihradszky, K.F., Burlingame, A.L., and Paulson,J.C. (1992) Glycobiology 2, 489.
10. Van den Eijnden, D.H., and Schiphorst, W.E.C.M. (1981) J. Biol. Chem. 256, 3159-3162.
11. Weinstein, J., de Souza-e-Silva, U., and Paulson, J.C. (1982) J. Biol. Chem. 257, 13835-13844.
12. Grundmann, U., Nerlich, C., Rein, T., and Zettlmeissl, G. (1990) Nucleic Acids Res. 18, 667.
13. Larsen, R.D., Ernst, L.K., Nair, R.P., and Lowe, J.B. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 6674-6678.
14. Weinstein, J., de Souza-e-Silva, U., and Paulson, J.C. (1982) J. Biol. Chem. 257, 13845-13853.