

ISOLATION AND CHARACTERIZATION OF A PARTIAL cDNA FOR A HUMAN  
SIALYLTRANSFERASE<sup>1</sup>Peter Lance\*, Karen M. Lau\*, and Joseph T.Y. Lau<sup>†</sup>Department of Molecular and Cellular Biology  
Roswell Park Memorial Institute  
Buffalo, New York 14263

Received August 21, 1989

**SUMMARY:** A probe generated from the coding sequence of the rat hepatic  $\beta$ -galactoside  $\alpha$ 2,6-sialyltransferase was used to screen a human cDNA library constructed of human submaxillary gland mRNA lambda gt-11. We report the isolation and characterization of a human cDNA, HSM-ST1, that is putatively the human homolog of the  $\beta$ -galactoside  $\alpha$ 2,6-sialyltransferase. The largest human clone contains a 1.3 kb cDNA insert and is predicted to encompass 75% of the coding sequence as well as a small portion of the 3' untranslated region. Comparative analysis of this insert with the rat hepatic  $\alpha$ 2,6-sialyltransferase sequence indicates 79% nucleotide similarity between the two sequences in the predicted coding region. On the amino acid level, the degree of conservation is 86%. Substantial sequence similarity is observed in the 3'-untranslated region between the rat and human sequences as well. S<sub>1</sub> nuclease analysis was performed to demonstrate the expression of HSM-ST1 transcripts in the human hepatoma cell line, HepG2, and in the human colonic adenocarcinoma cell lines, IS174T.

© 1989 Academic Press, Inc.

It is well known that sialylated oligosaccharides are principal epitopes for a variety of cellular recognition processes including homing of lymphocytes to peripheral lymph nodes (1), lymphocyte activation (2), viability of serum proteins in circulation (3), and the invasiveness of a number of pathogenic organisms and toxins (4). Indeed, cellular transformation is frequently accompanied by alterations to sialylated carbohydrate epitopes (5). The attachment of sialic acids to the termini of oligosaccharides is catalyzed by the sialyltransferases, a family of enzymes about which very little is known. This is due primarily to the lack of sequence information from which specific molecular probes can be constructed.

Recent and independent efforts from this laboratory (6) and from Dr. James Paulson's laboratory (7) have resulted in the sequence elucidation of

<sup>1</sup>This work was supported by a Veterans Administration RAG grant (P.L.) and by grant GM38193 from the National Institutes of Health (J.T.Y.L.).

\* Current address: GI Unit (111G), VA Medical Center, 3495 Bailey Ave, Buffalo, New York 14215.

<sup>†</sup> To whom correspondence should be addressed.

one of the sialyltransferases from rat liver. This enzyme, the  $\beta$ -galactoside  $\alpha$ 2,6-sialyltransferase, mediates the synthesis of the predominant sialyl linkage found in serum glycoproteins. In liver, the major site of serum glycoprotein biosynthesis, glucocorticoids apparently modulate the expression of this enzyme by altering the rate of transcript synthesis (6). The expression of this enzyme, however, is not limited to the liver. Although tissue-specific regulation of the  $\alpha$ 2,6-sialyltransferase is an issue yet to be fully addressed, a recent report from this laboratory documents that the rat  $\alpha$ 2,6-sialyltransferase gene sequence is differentially utilized to generate, in a tissue specific manner, a family of related mRNAs (8).

In this study, we used the coding region of the rat hepatic  $\beta$ -galactoside  $\alpha$ 2,6-sialyltransferase (RL-ST1) to probe a human cDNA library. We report the isolation of a cDNA sequence that apparently represents the human homolog of RL-ST1 based on nucleotide and predicted amino acid similarity.

#### MATERIALS AND METHODS

##### Materials

[ $\alpha$ - $^{32}$ P] dATP (3000 Ci/mmol), [ $\gamma$ - $^{32}$ P]ATP (3000 Ci/mmol), and [ $\gamma$ - $^{35}$ S]dATP (1000 Ci/mmol) were purchased from Amersham Corp. Restriction enzymes and other DNA modifying enzymes were from either Boehringer Mannheim Biochemicals, Stratagene, or Bethesda Research Laboratories. All tissue culture reagents were from GIBCO Laboratories. All other biochemicals were of the highest quality commercially available, and the chemicals were of reagent grade or higher.

##### Isolation and analysis of HSM-ST1 cDNA

Approximately 500,000 plaque forming units from a human submaxillary gland library constructed in lambda gt11 (Clontech Laboratory, Palo Alto, CA) were screened using a 780 bp BstII/BstII fragment from the coding region of the rat liver  $\beta$ -galactoside  $\alpha$ 2,6-sialyltransferase (RL-ST1) (6). Filter hybridization probes were generated by the random primer method (9). cDNA inserts were cloned into PUC or M13, and the nucleotide sequence was determined by the dideoxy chain terminator method (10). Sequencing reactions were primed with the M13 sequencing primer or the synthetic oligonucleotides HS1-P2 (5' GCACCACAGCCAAC 3') and HS1-P3 (5' GTGCCCATCTTGTTC 3').

##### Cell cultures

A human hepatoma cell line, HepG2, and a colon cancer cell line, LS174T, were obtained from American Type Culture Collection and grown in monolayer culture in DMEM supplemented with 10% fetal bovine serum, 2 mM glutamine, 25 mM glucose, 1 mM pyruvate, 100 units/ml of penicillin, and 100  $\mu$ g/ml of streptomycin. Non-essential amino acids were also added to media for LS174T cultures. LS174T-LD and LS174T-ND are liver and neck colonizing variants of the parent LS174T line, respectively. These variants were recovered from xenogeneic tumor foci in liver (LD) and neck (ND) of NIH.nu/nu mice (Taconic Park, NJ) that had received an ileocolic vein injection of  $10^6$  LS174T cells three to four weeks earlier. All cells were maintained at 37°C in 5% CO<sub>2</sub> atmosphere.

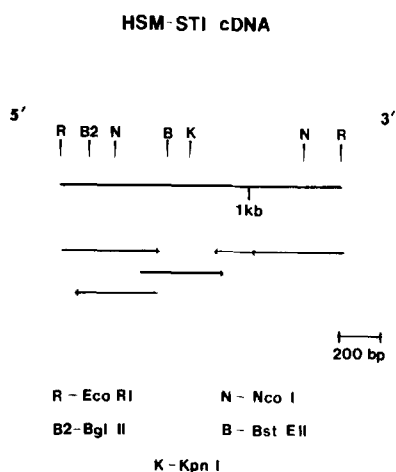
## Analysis of RNA

Total RNA from cultured cells was extracted by guanidinium isothiocyanate (11) and purified by centrifugation through a cushion of 5.7M CsCl (12). Blot analysis (13) was performed after fractionation of the RNA in formaldehyde agarose gels (14). Transcript protection analysis was performed as described previously (15) using 30  $\mu$ g of RNA from HepG2 or 60  $\mu$ g of RNA from the IS174T cells.

## RESULTS AND DISCUSSION

A 780 bp BstEII/BstEII fragment of the rat liver  $\beta$ -galactoside  $\alpha$ 2,6-sialyltransferase cDNA which represents the distal two-thirds of the coding region (6) was used to screen a human submaxillary gland cDNA library in lambda gt11. Previous experiments indicated that this rat coding region probe hybridizes to an approximately 4.7 kb human transcript from HepG2 cells, LS174T cells, and submaxillary gland tissues (Lance and Lau, unpublished observations). 5 positively hybridizing clones were isolated after 6 successive rounds of plaque purification. The cDNA inserts, ranging up to 1.3 kb in size, showed similar restriction digest patterns and were apparently derived from the same transcript. The partial restriction map of the largest of the inserts, HSM-ST1, is shown in Fig 1. The complete HSM-ST1 insert was cloned into the Eco RI site of M13. Single-stranded DNA was prepared and sequenced according to the strategy diagrammed in the lower portion of Fig 1.

The 1.3 kb HSM-ST1 insert is apparently comprised of a 1020 nt protein coding domain followed by a 258 nt segment of the 3'-untranslated region. Fig 2 shows the nucleotide sequence comparison of HSM-ST1 cDNA and the



**Fig. 1.** Partial restriction map of the human cDNA, HSM-ST1. Horizontal lines with arrows in the lower portion of the figure indicate the strategy that was used for sequence elucidation of HSM-ST1.

1	--GCACCCAGGACCCCCACAGGCCCGCAGACCCCTCGGCAGTCTCAGAGGCCATGCCAAGGCCAACACAGAGGCCCTCTTCCAGGTGTGG	HUMAN
372	AGCAAGCAAGACCCCTAAGGAAGACATTCCAATCCTCAGTTACCACAGGGTCACAGCCAAAGGTCAAACACAGCCCTCTTCCAGGTGTGG	RAT
88	AACAAGGACAGCTCTTCCAAAAACCTTATCCCTAGGCTGCAAAAGATCTGGAAGAATTACCTAAGCATGAACAAAGTACAAAGTGCTCTAC	HUMAN
462	GACAAGGACTCCACATACTCAAACCTTAACCCAGGGCTGCTGAAGATCTGGAGAAACTATCTGAACATGAACAAATATAAAGATCTCTAC	RAT
178	AAGGGGCCAGGACAGGCATCAAGTTCAGTGCAGAGGCCCTGCGCTGCCACATCCGGGACCATGTGAATGTATCCATGGTAGAGGTCACA	HUMAN
552	AAGGGACCGGGCCAGGAGTCAAGTTCAGCTGAGAAGCACTGCGTTGCCACCTTCGAGACCATGTCAACGTGGCGATGATAGAGGCCACA	RAT
268	GATTTTCCCTTCAATACCTCTGAATGGGAGGGTTATCTGCCAAGGAGAGCATTAGGACCAAGGCTGGGCTTGGGGCAGGTGTGCTGTT	HUMAN
642	GATTTTCCCTTCAACACCACTGAGTGGGAGGGTTACCTGCCAAGGAGAACTTTAGAAACAAGGTTGGGCTTGGCAAAAGGTGTGCCGTG	RAT
358	GTGTCGTGAGGGGATCTCAGAAGTCTCCCAACTAGGCAGAGAAATCGATGATCATGACGCAGTCTGAGGTTTAATGGGGCACCACA	HUMAN
732	GTCTCTCTGCGAGGATCTCTGAAAACTCCACGCTTGGTCGAGAGATTGATAATCATGATGCAGTTCTGAGGTTAATGGGGCCCTACC	RAT
448	GCCAACTTCAACAAGATGTGGGCACAAAACCTACCAATTCGCCTGATGAACCTCTCAGTTGGTTACCACAGAGAAGCGCTTCTCAAGAGC	HUMAN
822	GACAACCTCAACAGGATGTGGGCTCAAAAACCTACCAATTCGCCTAATGAACCTCTCAGTTAGTCACCACAGAAAAGCGCTTCTCAAGGAC	RAT
538	AGTTTGTACAATGAAGGAATCCTAATTGTATGGGACCCATCTGTATACCACTCAGATATCCCAAAGTGGTACCAGAAATCCGATTATAAT	HUMAN
912	AGTTTGTACACCGAAGGAATCCTAATTGTATGGGACCCATCCGTGTATCATGCAGATATCCCAAAGTGGTATCAGAAACACAGACTACAAT	RAT
628	TTCTTTAACAACACAGACTTATCGTAAGCTGCACCCCAATCAGCCCTTTTACATCCTCAAGCCCAGATGCCTTGGGAGCTATGGGAC	HUMAN
1002	TTCTTCGAAACCTATAAGAGTTACCGAAGGCTGAACCCAGCCAGCCATTTTATATCCTCAAGCCCAGATGCCATGGGAACGTGGGAC	RAT
718	ATTCTTCAAGAAATCTCCCAAGAGGATTCAGCCAAACCCCCATCCTCTGGGATGCTTGGTATCATCATGATGACGCTGTGTGAC	HUMAN
1092	ATCATTCAGGAAATCTCTGAGATGCTGATTAGCCAAATCCCCATCCTCGGGATGCTGGGATATCATCATGATGACTCTGTGTGAC	RAT
808	CAGGTGGATATTTAGGTTCTCCCATCCAAGCGCAAGACTGACGTGTGCTACTACTACCAGAAGTTCTTCGATAGTCCTGCACGATG	HUMAN
1182	CAGGTAGATATTTACGAGTTCCTCCCATCCAAGCGCAAGCGGACGTGTGCTATTATCACCAAAAGTTCTTTGACAGCGCTGCACCATG	RAT
898	GGTGCTACCAACCGCTGCTCTATGAGAAGAAATTTGGTGAAGCATCTCAACAGGGCACAGATGAGGACATCTACCTGCTTGGAAAAAGCC	HUMAN
1272	GGTGCTACGACCCGCTCCTCTTCGAGAAGAAATGTTGAAGCATCTCAATGAGGGAACAGATGAAGACATTTATTGTTTGGGAAAGCC	RAT
988	ACACTGCTGCGCTCCGGACCATTCAGTCTAA 1020	HUMAN
1362	ACCTTTTCGGCTTCCGGAACATTCGTTGTTGA 1394	RAT

[illegible]

corresponding portion of the rat liver  $\alpha 2,6$ -sialyltransferase sequence (6-8). Immediately obvious feature is the close sequence similarity between the rat liver sialyltransferase and the putatively human homolog, HSM-ST1. Similarity is most striking within the coding region which exhibits 79% nucleotide sequence similarity (Fig 2A). Quite unexpectedly, sequence similarity extends into a portion of the 3'-untranslated region immediately adjacent to the translation stop codon (Fig 2B).

86% conservation between HSM-ST1 and the rat liver  $\alpha 2,6$ -sialyltransferase is observed on the amino acid level (Fig 3). This alignment suggests that the human cDNA contains approximately three quarters of the complete coding sequence. The 5' end of the HSM-ST1 coincides with the second codon of a soluble form of rat sialyltransferase protein that is putatively generated by proteolytic removal of a 63aa  $\text{NH}_2$ -terminal membrane domain (7). A summary of the comparative analysis between the human clone, HSM-ST1, and the rat liver sialyltransferase, RL-ST1, is diagrammed in Fig 4. While it is apparent that the 5' information that putatively encodes the  $\text{NH}_2$  membrane anchor is missing in HSM-ST1, it is quite probable that the sequences required for catalytic activity and specificity are present in HSM-ST1. This hypothesis is currently being tested by expression of the human sequence under the control of a heterologous promoter elements in cultured cells.

Northern blot analysis of RNA isolated from human hepatoma and colon cancer cell lines probed with the HSM-ST1 fragment suggests that the human sialyltransferase sequence is carried on a large, approximately 5.0 kb mRNA (data not shown). Not surprisingly, identical results were obtained using the highly homologous rat  $\alpha 2,6$ -sialyltransferase coding region probe. Since the signal on RNA blots can be the result of HSM-ST1 sequences as well as related but distinctly different transcripts (for example, mRNAs of

1	-APRPTGROTLGSLRGLAKAPEASFQVWVKDSSSKNLIPRLQKIWKNYLSMNKYKVS YKGGPGIKFSAEALRCHLRDHVNVSMVEVT	HUMAN
64	SKQDPKEDIPLSYHRVTAKVKPQPSFQVMDKSTYSKLNPRLLKIWRNYLNMNKYKVS YKGGPGVKFSVEALRCHLRDHVNVSMIEAT	RAT
90	DFFPNTSEWEGYLPKESIRTKAGPWGRCAVVSAGSLKSSQLGREIDHDVLRFGNAPTANFQDVGTKTIRLMNSQLVTTEKRFKLD	HUMAN
154	DFFPNTTEWEGYLPKENFRYKVGPRCAVVSAGSLKNSQLGREIDHDVLRFGNAPDNFQDVGSKTIRLMNSQLVTTEKRFKLD	RAT
180	SLYNEGILIVDPSVYHSDIPKYYQNPDYNNFYNNKYTKRLHPNQPFYILKPQMPWELWDILQEISPEEIQPNPPSSGMLGIIMNTLCD	HUMAN
244	SLYTEGILIVDPSVYHADIPKYYQKPDYNNFFETYKSYRRLNPSQPFYILKPQMPWELWDIIQEISADLIQPNPPSSGMLGIIMNTLCD	RAT
270	QVDIYEFPLSKRKTDVCCYYQKFFDSACTNGAYHPLLYEKNLVKHLNQGTDEDIYLLGKATLPGFRTINC	HUMAN
334	QVDIYEFPLSKRKTDVCCYYHQKFFDSACTNGAYDPLLFKNMVKHLNEGTDEDIYLFGKATLSGFRNIRC	RAT

Fig. 3. Comparison of the deduced amino acid sequences of HSM-ST1 and RL-ST1. the symbol ♦ indicates predicted amino acid similarity between the two sequences at that position.

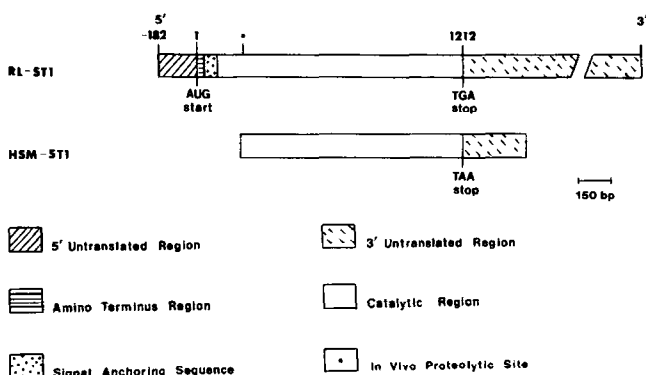
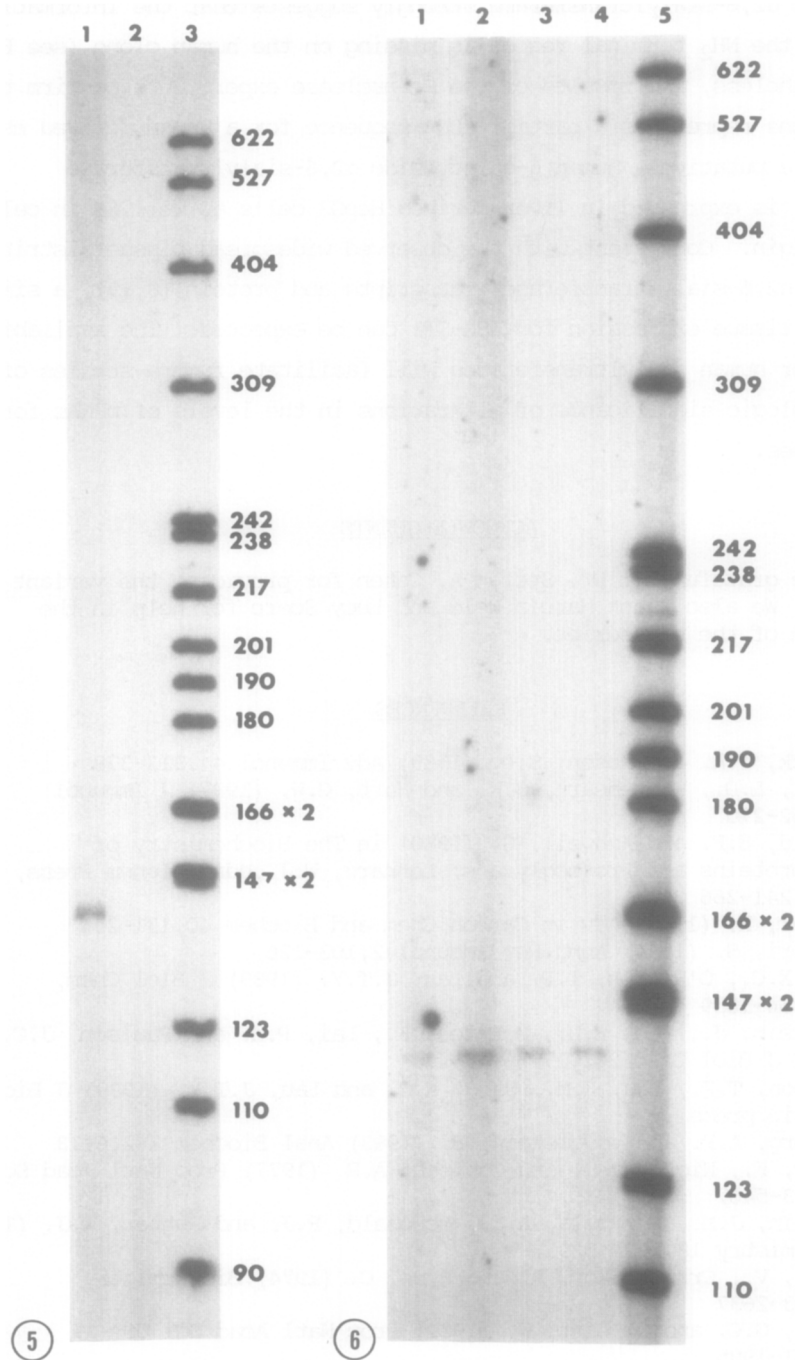


Fig. 4. Schematic comparison of HSM-ST1 and RL-ST1.

sialyltransferases with different linkage specificities),  $S_1$  mapping analysis was used to examine directly the expression of HSM-ST1 mRNA. An end-labeled probe was prepared using the Bgl II site that resides 140 bp downstream of the 5' end of the HSM-ST1 cDNA clone. The specificity of this analysis for HSM-ST1 transcripts is demonstrated in Fig 5. RNA from human Hep G2 cells (lane 1) protected the entire 140 nt HSM-ST1 region present on the probe. In contrast, no signal is observed using RNA from rat liver, a tissue source known to be enriched for the highly similar rat  $\alpha 2,6$ -sialyltransferase mRNA (lane 2).  $S_1$  analysis was also used to examine expression of HSM-ST1 mRNA in other human derived cell lines. Fig 6 illustrates that HSM-ST1 is expressed in LS174T cells (lanes 1-3) as well as in HepG2 cells (lane 4). However, the HSM-ST1 transcript level is a reproducible >2 fold higher in HepG2 cells than in the LS174T lines; for the  $S_1$  analysis shown in Fig 6, only 30  $\mu$ g of HepG2 RNA was used compared to 60  $\mu$ g of RNA for each of the LS174T cell lines. There are also differences in the level of HSM-ST1 expression among the LS174T lines. HSM-ST1 mRNA is lowest in the parental LS174-CD line (Fig 6, lane 1), and relatively higher in the liver and neck colonizing variants, LS174-LD and LS174-ND (Fig 6, lanes 2 and 3, respectively). Overall, the levels of HSM-ST1 expression in the different cell lines are: HepG2 > LS174T-ND and LS174T-LD > LS174T-CD.

Taken together, the data indicate that HSM-ST1 represents a partial clone for the human equivalent of the rat hepatic  $\beta$ -galactoside  $\alpha 2,6$ -sialyltransferase. Since all known mammalian glycosyltransferases including human (16), bovine (17), and mouse (18) galactosyltransferases, and the rat sialyltransferase are encoded on mRNAs with extensive 3'-untranslated regions, it is highly probable that this feature is shared by HSM-ST1 transcript and that this complete region is not represented in our cDNA clones. Likewise, alignment of the HSM-ST1 coding sequence with the



**Fig. 5.** S<sub>1</sub> nuclease analysis of HepG2 (lane 1) and rat liver mRNA (lane 2). HSM-ST1 cDNA insert subcloned into PUC vector was linearized at the unique Bgl II site located at position 140 within the cDNA insert and labeled with <sup>32</sup>P using polynucleotide kinase. This end-labeled fragment was used as probed for the S<sub>1</sub> mapping experiment as described in Materials and Methods. Size markers are PER fragments generated by digestion with Hpa II (lane 3).

**Fig. 6.** S<sub>1</sub> nuclease analysis of LS174T and HepG2 RNAs. Lane 1, LS174T-CD; lane 2, LS174T-ND; lane 3, LS174T-LD; and lane 4, HepG2. Analysis was performed as described in Fig 5 with 30 µg of HepG2 RNA or 60 µg of each of the LS174T RNAs.

rat hepatic  $\alpha 2,6$ -sialyltransferase strongly suggests that the information for 25% of the  $\text{NH}_2$  terminal region is missing on the human clone (see Fig 4). Nevertheless, the results of the  $\text{S}_1$  nuclease experiments confirm that HSM-ST1 clone represents a partial cDNA sequence for a human derived mRNA. This mRNA, a putatively human  $\beta$ -galactoside  $\alpha 2,6$ -sialyltransferase transcript, is expressed in liver derived HepG2 cells as well as in cells of colonic origin. Consistent with the observed widespread tissue distribution of the rat  $\alpha 2,6$ -sialyltransferase transcripts and protein (8,19), a similar pattern of tissue expression for HSM-ST1 can be expected. The availability of cDNAs for human sialyltransferases will facilitate future studies of the pathophysiologic significance of alterations in the levels of mRNAs for these enzymes.

#### ACKNOWLEDGMENTS

We are grateful for Dr. Stefan A. Cohen for providing the variant lines of LS174T. We also thank Marcia Held and Lucy Scere for help in the preparation of the manuscript.

#### REFERENCES

1. Yednock, T.A. and Rosen, S.D. (1989) *Adv Immunol* 44,313-378
2. Powell, L.D., Whiteheart, S.W. and Hart, G.W. (1987) *J Immunol* 139,262-270
3. Neufeld, E.F. and Ashwell, G. (1980) in *The Biochemistry of Glycoproteins and Proteoglycans*; Lennarz, W.J., eds (Plenum Press, New York) 241-266
4. Schauer, R. (1982) *Adv in Carboh Chem and Biochem* 40,131-234
5. Hakomori, S. (1984) *Annu Rev Immunol* 2,103-126
6. Wang, X.C., O'Hanlon, T.P. and Lau, J.T.Y. (1989) *J Biol Chem*, 264,1854-1859
7. Weinstein, J., Lee, E.U., McEntee, K., Lai, P.H. and Paulson, J.C. (1987) *J Biol Chem* 262,17735-17743
8. O'Hanlon, T.P., Lau, K.M., Wang, X.C. and Lau, J.T.Y. (1989) *J Biol Chem*, in press
9. Feinberg, A.P. and Vogelstein, B. (1983) *Anal Biochem* 132,6-13
10. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc Natl Acad Sci USA* 74,5463-5467
11. Chirgwin, J.M., Przybyla, A.E., McDonald, R.J. and Rutter, W.J. (1979) *Biochemistry* 18,5294-5299
12. Glisin, V., Crkvenjakov, R. and Byus, C. (1974) *Biochemistry* 13,2633-2637
13. Church, G.C. and Gilbert, W. (1984) *Proc Natl Acad Sci USA* 81,1991-1995
14. Boedtker, H. (1971) *Biochim Biophys Acta* 240,448-453
15. Lau, J.T.Y., Pittenger, M.F. and Cleveland, D.W. (1985) *Mol Cell Biol* 5,1611-1620
16. Masri, K.A., Appert, H.E. and Fukuda, M.N. (1988) *Biochem Biophys Res Comm* 157,657-663
17. Narimatsu, H., Sinha, S., Brew, K., Okayama, H. and Qasba, P.K. (1986) *Proc Natl Acad Sci USA* 83,4720-4724
18. Shaper, N.L., Hollis, G.F., Douglas, J.G., Kirsch, I.R. and Shaper, J.H. (1988) *J Biol Chem* 263,10420-10428
19. Paulson, J.C., Weinstein, J. and Schauer, A. (1989) *J Biol Chem* 264,10931-10934