# Supplement A

“Rules” Governing the Structures of Human Milk Glycans Based on Knowledge of Published Structural Analyses and Enzyme Specificities.

A. *“Rules” for Core Structures in Human Milk Glycans* – Human milk oligosaccharides are all based on lactose, which provides the reducing end of all HMGs. Thus, all HMGs have 1 hexose that is glucose, and all other hexoses are galactose. The structures below are drawn using the following designations:



The structures are drawn to indicate linkage positions:

 

1. All milk glycans are based on Lactose:

 

2. Terminal Gal of lactose can be extended in milk glycans by addition of a GlcNAc in a β3 linkage. Here the terminal Gal of lactose is extended to the trisacchairde, but this trisaccharide does not occur in human milk. The β3-GlcNAc transferase that catalyzes this reaction transfers only to Galβ1-4Glc or GlcNAc, but not to Galβ1-3GlcNAc[1](#_ENREF_1) [2](#_ENREF_2), [3](#_ENREF_3).



3. There are no terminal GlcNAc residues in milk glycans. Terminal GlcNAc residues are immediately extended by active Gal transferases to add either Gal-β4 or -β3 to make either Lacto-N-neoTetraose (LnNT) or Lacto-N-Tetraose (LNT), respectively. Therefore, there are only 2 tetrasaccharides in human milk glycans.

 

 LNnT LNT

4. The core structures of human milk glycans, therefore, “grow” as if a disaccharide were added to the non-reducing end, but a Galβ1-3GlcNAc (Type 1 disaccharide) is a terminal structure since the β3-GlcNAc transferase that catalyzes this reaction transfers only to Galβ1-4GlcNAc; not to Galβ1-3GlcNAc (Rule 2), making the type 1 terminal disaccharide a prominent determinant in the human milk soluble glycan glycome[4](#_ENREF_4). Thus, there are only 2 possible linear hexaoses in human milk that result from the extension of Lacto-N-neoTetraose.

 

*para*-Lacto-N-neoHexaose *para*-Lacto-N-Hexaose

5. The Core Branching Rule - A branch can occur only on a Gal residue that is in a liner

 sequence. The branch point is always a GlcNAcβ1-6 to an internal Gal; never to a terminal Gal and the enzyme responsible for this biosynthesis is presumably the developmently regulated I-branching enzyme[3](#_ENREF_3), [5](#_ENREF_5). Examples below show branching of LNnT and LNT. But according to rule 3, there are no terminal GlcNAc residues so the branched LNnT and LNT must be extended, but the Galβ1-3transferase apparently cannot add a Gal to GlcNAcβ1-6Gal. Thus, the 6-arm of branched structures is always a type 2 N-acetyllactosamine (Galβ1-4GlcNAc) and there are only 2 branched hexasaccharides in human milk soluble glycans.

  

 Lacto-N-neoHexaose (LnNH) Lacto-N-Hexaose (LNH)

The branching can also occur on an internal LacNAc of a longer linear glycan based on the recent discovery of the following core structure [6](#_ENREF_6).:



B. *“Rules” for addition of Fucose to Core Structures in Human Milk Glycans* – The core structures are “decorated” with fucose residues based on the specificity of at least 3 fucosyltransferases.

1. An α(1,2)-Fucosyltransferase - The α(1,2)fucosyltransferase expressed in human milk is genetically regulated by the “Secretor” gene that is associated with the ability of an individual to produce H antigen in secretions such as saliva, mucins and milk[7](#_ENREF_7). The Secretor (*Se*) locus is on Chromosome 19 at 19q13.3. "Secretors" have a *Se*/*Se* or *Se*/*se* genotype producing at least one copy of the *Se* gene that encodes a functional α(1,2)fucosyltransferase, which is the Fuc-T2 enzyme [8](#_ENREF_8). Non-secretors (se/se) are unable to produce a soluble form of H antigen. Fuc-T2 transfers α-fucose to the 2 position of the terminal Galactose only on type 1 glycans (Galβ1-3GlcNAc-R) to generate H type 1 glycans, but not H type 2 glycans[9](#_ENREF_9). Thus, human milk contains only traces of H-type 2 (Fucα1-2Galβ1-4GlcNAcβ1-3Galβ1-4Glc), but significant amounts of 2’Fucosyllactose (Fucα1-2Galβ1-4Glc), which is only present in milk of mothers that are Secretors[10](#_ENREF_10) indicating that this enzyme will use lactose at high concentrations as a substrate or is able to utilize the lactose, which is a reducing disaccharide where the glucose is in an equilibrium of α and β conformations[11](#_ENREF_11).

  

 H-type 1 glycan H-type 2 glycan 2’-Fucosyllactose

2. The Lewis fucosyltransferase - The α(1,3/4)fucosyltransferase expressed in human milk is genetically regulated by the “Lewis” gene that is associated with the ability of an individual to produce Lewis a antigen in secretions such as saliva, mucins and milk and the Lewis b antigen in individuals that have an active Secretor gene (*Se/Se* or *Se/se*) and an active Lewis gene (*Le/Le or Le/le*)[12](#_ENREF_12).

  

Lewis a antigen Lewis b antigen

The Lewis (*Le*) locus is on Chromosome 19 at 19q13.3, and the gene product encodes a functional α(1,3/4)fucosyltransferase, which is the Fuc-T3 enzyme[13](#_ENREF_13). This enzyme is unique in its ability to transfer fucose from GDP-Fucose to GlcNAc in both α1-3 and α1-4 linkages. This enzyme can, therefore, add fucose to either a type 1 glycan to form the Lewis a and b antigens or to a type 2 glycan to form the Lewis x and Lewis y antigens in concert with the H blood group α(1,2)fucosyltransferase[14](#_ENREF_14" \o "Larsen, 1990 #1235) (Fuc-T1). Fuc-T3 is also responsible for transferring fucose to the 3 position of glucose making 3-fucosyllactose and difucosyllactose dependent on the expression of this this enzyme.

   

Lewis x antigen Lewis y antigen 3-fucosyllactose difucosyllactose

Since the H blood group α(1,2)fucosyltransferase (Fuc-T1) is not expressed in human milk, there is little if any Lewis y antigen present; however, relatively large amounts of the Lewis x antigen are found in human milk. This structure can obviously be synthesized by the action of the Lewis gene, which codes for the Fuc-T3 enzyme that can transfer a fucose to the GlcNAc of both type 1 and type 2 glycans. The 3-fucosyl lactose can be extended as linear glycans, but based on the available known structures in human milk glycans, the branching enzyme will not act on Lacto-N-neoTetraose (LnNT) or Lacto-N-Tetraose (LNT) that possess the fucose on the 3-position of the reducing glucose. The first of the linear series of glycans with a fucose in this position are Lacto-N-Fucopentaose V and Lacto-N-*neo*Fucopentaose V. We assume that linear glycans can have this fucose, but that branched glycans will not. This rule may be altered if branched structures with this fucose are discovered.

 

 Lacto-N-Fucopentaose V Lacto-N-*neo*Fucopentaose V

3. α(1,3)-Fucosyltransferase – Since individuals with the genotype *le/le* also present significant amounts of Lewis x antigen in milk, there must be a ubiquitous α(1,3)-fucosyltransferase expressed in milk that can synthesize this structure. Analysis of the fucosyltransferase activities in human milk indicated that the α(1,3/4)fucosyltransferase activity (Fuc-T3) could be physically separated from an α(1,3)-fucosyltransferase activity present in virtually all individuals[15](#_ENREF_15), and that the α(1,3)-fucosyltransferase cannot transfer fucose to the 3-position of glucose[16](#_ENREF_16) making the presence of the 3-fucosylated glucose dependent on the expression of the Lewis fucosyltransferase (α(1,3/4)fucosyltransferase). Thus, the α(1,3)-fucosyltransferase is distinct from the Lewis fucosyltransferase. The human gene that expresses the ubiquitous α(1,3)-fucosyltransferase (Fuc-T6) was cloned[17](#_ENREF_17), [18](#_ENREF_18). The recombinant Fuc-T6 gene product was expressed in insect cells and displayed acceptor activity identical to the ubiquitous human milk α(1,3)-fucosyltransferase[16](#_ENREF_16).

C. *“Rules” for addition of Sialic Acid to Core Structures in Human Milk Glycans* – The core structures are “decorated” with sialic acid residues based on the specificity of at least 3 sialyltransferases. Although the amounts and distribution of sialylated glycans in milk of many species has been intensely investigated, studies on expression of the sialyltransferase genes in human milk are limited [19](#_ENREF_19). Thus, the rules of sialic acid addition to the soluble human milk glycans are based on the sialic acid linkages that have been described.

1. Neu5Acα2-6Galβ1-4GlcNAc – This structure is presumably the product of the sialyltransferase that is the product of the ST6Gal I gene, that is known to be expressed in human lactating mammary gland [19](#_ENREF_19). This enzyme is apparently responsible for synthesis of the two major glycans in milk possessing this linkage, 6’-Sialyllactose and LSTc.

  

 6’-Sialylactose LSTc

 The α2,6 sialic acid linkage does not occur on type 1 free glycans in human milk, which is consistent with the specificity of the ST6Gal I sialyltransferase specificity for type 2 glycans[20](#_ENREF_20). The sialyltransferase that adds sialic acid in α2,3 linkage to type 1 glycans is apparently not expressed in human lactating mammary gland.

2. Neu5Acα2-3Galβ1-3GlcNAc – The Neu5Acα2-3Gal linkage occurs in human milk in several major glycan structures including 3’-Sialyllactose and LSTa, but this sialic acid linkage does not occur on glycans with a terminating type 2 LacNAc (Galβ1-4GlcNAc).

  

 3’-Sialylactose LSTa

 The gene responsible for the expression of the active α2,3-sialyltransferase in human milk has not been reported, but its specificity is very restricted for the type 1 glycan (Galβ1-3GlcNAc); however, 3’-Sialylactose, which a major component of human milk soluble glycans, has a type 2-related acceptor (Galβ1-4Glc) acceptor. If this enzyme is responsible for the biosynthesis of 3’-Sialylactose, its ability to use the type 2 lactose as an acceptor may be due to the high concentration of lactose available in milk or the “flexibility” of the reducing disaccharide as proposed for the addition of fucose to lactose by Fuc-T2 Fuc[11](#_ENREF_11).

3. Galβ1-3(Neu5Acα2-6)GlcNAc – The Neu5Acα2-6GlcNAc is an internal structure found only on Type 1 glycans and never on Type 2 glycans or on the 6-position of Glucose in lactose. Two major free milk glycans possessing this structure are LSTb and disialy-Lacto-N-Tetraose (DSL).

  

 LSTb Disialy-Lacto-N-Tetraose (DSL)

This unique α2,6-linkage has never been found in a Lacto-N-neoTetraose (Galβ1-4GlcNAcβ1-3Galβ1-4Glc), and the gene responsible for the expression of this sialyltransferase has not been identified.

1. Sasaki, K. et al. Expression cloning of cDNA encoding a human beta-1,3-N-acetylglucosaminyltransferase that is essential for poly-N-acetyllactosamine synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 14294-14299 (1997).

2. Renkonen, O. Enzymatic in vitro synthesis of I-branches of mammalian polylactosamines: generation of scaffolds for multiple selectin-binding saccharide determinants. *Cellular and molecular life sciences : CMLS* **57**, 1423-1439 (2000).

3. Mattila, P. et al. The centrally acting beta1,6N-acetylglucosaminyltransferase (GlcNAc to gal). Functional expression, purification, and acceptor specificity of a human enzyme involved in midchain branching of linear poly-N-acetyllactosamines. *The Journal of biological chemistry* **273**, 27633-27639 (1998).

4. Urashima, T. et al. The predominance of type I oligosaccharides is a feature specific to human breast milk. *Advances in nutrition* **3**, 473S-482S (2012).

5. Bierhuizen, M.F., Mattei, M.G. & Fukuda, M. Expression of the developmental I antigen by a cloned human cDNA encoding a member of a beta-1,6-N-acetylglucosaminyltransferase gene family. *Genes & development* **7**, 468-478 (1993).

6. Yu, Y. et al. Human milk contains novel glycans that are potential decoy receptors for neonatal rotaviruses. *Mol Cell Proteomics* **13**, 2944-2960 (2014).

7. Lowe, J.B. The blood group-specific human glycosyltransferases. *Bailliere's clinical haematology* **6**, 465-492 (1993).

8. Kelly, R.J., Rouquier, S., Giorgi, D., Lennon, G.G. & Lowe, J.B. Sequence and expression of a candidate for the human Secretor blood group alpha(1,2)fucosyltransferase gene (FUT2). Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the non-secretor phenotype. *The Journal of biological chemistry* **270**, 4640-4649 (1995).

9. Sarnesto, A., Kohlin, T., Hindsgaul, O., Thurin, J. & Blaszczyk-Thurin, M. Purification of the secretor-type beta-galactoside alpha 1----2-fucosyltransferase from human serum. *The Journal of biological chemistry* **267**, 2737-2744 (1992).

10. Grollman, E.F. & Ginsburg, V. Correlation between secretor status and the occurrence of 2'-fucosyllactose in human milk. *Biochemical and biophysical research communications* **28**, 50-53 (1967).

11. Castanys-Munoz, E., Martin, M.J. & Prieto, P.A. 2'-fucosyllactose: an abundant, genetically determined soluble glycan present in human milk. *Nutr Rev* **71**, 773-789 (2013).

12. Grollman, E.F., Kobata, A. & Ginsburg, V. An enzymatic basis for Lewis blood types in man. *The Journal of clinical investigation* **48**, 1489-1494 (1969).

13. Kukowska-Latallo, J.F., Larsen, R.D., Nair, R.P. & Lowe, J.B. A cloned human cDNA determines expression of a mouse stage-specific embryonic antigen and the Lewis blood group alpha(1,3/1,4)fucosyltransferase. *Genes & development* **4**, 1288-1303 (1990).

14. Larsen, R.D., Ernst, L.K., Nair, R.P. & Lowe, J.B. Molecular cloning, sequence, and expression of a human GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase cDNA that can form the H blood group antigen. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 6674-6678 (1990).

15. Johnson, P.H. & Watkins, W.M. Purification of the Lewis blood-group gene associated alpha-3/4-fucosyltransferase from human milk: an enzyme transferring fucose primarily to type 1 and lactose-based oligosaccharide chains. *Glycoconjugate journal* **9**, 241-249 (1992).

16. De Vries, T., Palcic, M.P., Schoenmakers, P.S., Van Den Eijnden, D.H. & Joziasse, D.H. Acceptor specificity of GDP-Fuc:Gal beta 1-->4GlcNAc-R alpha 3-fucosyltransferase VI (FucT VI) expressed in insect cells as soluble, secreted enzyme. *Glycobiology* **7**, 921-927 (1997).

17. Koszdin, K.L. & Bowen, B.R. The cloning and expression of a human alpha-1,3 fucosyltransferase capable of forming the E-selectin ligand. *Biochemical and biophysical research communications* **187**, 152-157 (1992).

18. Weston, B.W., Smith, P.L., Kelly, R.J. & Lowe, J.B. Molecular cloning of a fourth member of a human alpha (1,3)fucosyltransferase gene family. Multiple homologous sequences that determine expression of the Lewis x, sialyl Lewis x, and difucosyl sialyl Lewis x epitopes. *The Journal of biological chemistry* **267**, 24575-24584 (1992).

19. Maksimovic, J., Sharp, J.A., Nicholas, K.R., Cocks, B.G. & Savin, K. Conservation of the ST6Gal I gene and its expression in the mammary gland. *Glycobiology* **21**, 467-481 (2011).

20. Paulson, J.C. & Colley, K.J. Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation. *The Journal of biological chemistry* **264**, 17615-17618 (1989).