Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 2645141 C 2018/07/10

(11)(21) 2 645 141

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2007/03/07
(87) Date publication PCT/PCT Publication Date: 2007/09/13
(45) Date de délivrance/Issue Date: 2018/07/10
(85) Entrée phase nationale/National Entry: 2008/09/08
(86) N° demande PCT/PCT Application No.: EP 2007/052114
(87) N° publication PCT/PCT Publication No.: 2007/101862
(30) Priorité/Priority: 2006/03/09 (US60/780,350)

(51) Cl.Int./Int.Cl. *C12P 19/18* (2006.01), *C12N 1/21* (2006.01), *C12N 9/04* (2006.01), *C12N 9/10* (2006.01), *C12N 9/12* (2006.01), *C12P 19/26* (2006.01)

(72) Inventeur/Inventor:
SAMAIN, ERIC, FR

(73) Propriétaire/Owner:
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS), FR

(74) Agent: NORTON ROSE FULBRIGHT CANADA LLP/S.E.N.C.R.L., S.R.L.

(54) Titre : PROCEDE DE PRODUCTION D'OLIGOSACCHARIDES SIALYLES
(54) Title: METHOD OF PRODUCING SIALYLATED OLIGOSACCHARIDES

(57) Abrégé/Abstract:
The present invention relates to a method for the large scale in vivo synthesis of sialylated oligosaccharides, culturing a microorganism in a culture medium, optionally comprising an exogenous precursor such as lactose, wherein said microorganism comprises heterologous genes encoding a CMP-Neu5Ac synthetase, a sialic acid synthase, a GlcNAc-6-phosphate 2 epimerase and a sialyltransferase, and wherein the endogenous genes coding for sialic acid aldolase (NanA) and for ManNac kinase (NanK) have been deleted or inactivated. The invention also relates to these micoorganisms which are capable of producing internally activated sialic acid.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 September 2007 (13.09.2007)

PCT

(10) International Publication Number
WO 2007/101862 A1

(51) International Patent Classification:
C12P 19/18 (2006.01) *C12N 9/04* (2006.01)
C12P 19/26 (2006.01) *C12N 9/10* (2006.01)
C12N 1/21 (2006.01) *C12N 9/12* (2006.01)

(21) International Application Number:
PCT/EP2007/052114

(22) International Filing Date: 7 March 2007 (07.03.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/780,350 9 March 2006 (09.03.2006) US

(71) Applicant *(for all designated States except US)*: **CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS)** [FR/FR]; 3, rue Michel Ange, F-75016 Paris (FR).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **SAMAIN, Eric** [FR/FR]; 5, allée du Charmant Som, F-38610 Gieres (FR).

(74) Agent: **SCHRIMPF, WARCOIN, AHNER, TEXIER, LE FORESTIER, CALLON DE LAMARCK, COLLIN, TETAZ-Cabinet Regimbeau**; 20, rue de Chazelles, F-75847 Paris Cedex 17 (FR).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD OF PRODUCING SIALYLATED OLIGOSACCHARIDES

(57) **Abstract:** The present invention relates to a method for the large scale in vivo synthesis of sialylated oligosaccharides, culturing a microorganism in a culture medium, optionally comprising an exogenous precursor such as lactose, wherein said microorganism comprises heterologous genes encoding a CMP-Neu5Ac synthetase, a sialic acid synthase, a GlcNAc-6-phosphate 2 epimerase and a sialyltransferase, and wherein the endogenous genes coding for sialic acid aldolase (NanA) and for ManNac kinase (NanK) have been deleted or inactivated. The invention also relates to these micoorganisms which are capable of producing internally activated sialic acid.

WO 2007/101862 A1

METHOD OF PRODUCING SIALYLATED OLIGOSACCHARIDES

FIELD OF THE INVENTION

The present invention relates to a method for the large scale in vivo synthesis of sialylated oligosaccharides, culturing a microorganism in a culture medium, optionally comprising an exogenous precursor such as lactose, wherein said microorganism comprises heterologous genes encoding a CMP-Neu5Ac synthetase, a sialic acid synthase, a GlcNAc-6-phosphate 2 epimerase and a sialyltransferase, and wherein the endogenous genes coding for sialic acid aldolase (*NanA*) and for ManNac kinase (*NanK*) have been deleted or inactivated. The invention also relates to this micoorganism which is capable of producing internally activated sialic acid.

BACKGROUND OF THE INVENTION

N-acetylneuraminic acid (Neu5Ac) is the most common member of the sialic acid family of aminosugars. Neu5Ac is frequently found as a terminal sugar in cell surface complex carbohydrates and plays a major role in many biological processes such as cellular adhesion and binding of toxins and virus (Varki, 1993). Neu5Ac is also a major component of the carbohydrate portion of gangliosides which are notably abundant in brain tissue and are involved in several pathologies (Zhang & Kiechle, 2004).

In reason of their important biological functions, sialic acid containing oligosaccharides has attracted considerable interest and many methods have been developed to synthesize these structures for fundamental research and potential therapeutic applications. However, large scale production of sialylated oligosaccharides has not been reached as of today.

Chemical syntheses are not practical in reason of the multiple protection and deprotection steps and a lot of efforts has been put on enzymatic and biotechnological methods. Sialyltransferases use CMP-Neu5Ac as the activated sugar-nucleotide and the development of efficient processes for enzymatic syntheses of sialyllooligosaccharides has been possible through the identification of bacterial sialyltransferase genes which

are well expressed in *E. coli* and the design of multiple enzymatic systems that mimic the natural pathway of sugar nucleotide biosynthesis (Gilbert et al., 1998).

A significant improvement later came from the use of living bacterial cells to produce sialylooligosaccharides (Priem *et al.*, 2002). In this approach, sialyllactose was directly produced by growing cells of metabolically engineered *Escherichia coli* strains overexpressing the *Neisseria meningitidis* genes for α-2,3-Sialyltransferase and for CMP-Neu5Ac synthase. The bacteria were grown at high cell density with glycerol as the carbon and energy source, while exogenous lactose and Neu5Ac were supplied as precursors for sialyllactose synthesis. During the growth, lactose and Neu5Ac were actively internalized by *E. coli* β-galactoside and Neu5Ac permeases. To prevent catabolism of lactose and Neu5Ac, mutant strains devoid of β-galactosidase and Neu5Ac aldolase activities were used. Lactose and Neu5Ac accumulated in the cytoplasm where Neu5Ac was then converted into CMP-Neu5Ac to be further transferred on lactose to form sialyllactose (our European patent EP 1194584). This system was applied to the production of the carbohydrate portion of gangliosides GM2 and GM1 by additionally expressing the appropriate glycosyltransferase genes (Antoine *et al.*, 2003). Polysialylated oligosaccharides (GD3 and GT3 sugars) were also produced by this method and with the *Campylobacter cstII* gene that encodes a bifunctional α-2,3- and α-2,8- sialyltransferase (our application US 60/690,837 and Antoine *et al.*, 2005).

Large scale production of sialylooligosaccharides by this microbiological method requires important amount of sialic acid as a precursor. Sialic acid can be purified from natural sources such as milk and egg yolk, but the yields are low and the procedure is not suitable for large scale production. Sialic acid is generally prepared by enzymatic synthesis by the sialic acid aldolase using N-acetylmannosamine (ManNAc) and pyruvate as substrate. To reduce the cost, ManNAc is usually prepared by chemical or enzymatic epimerization of N-acetylglucosamine which is a cheaper substrate than ManNAc (Lee et al., 2004; Maru et al., 1998). In spite of these improvements the sialic acid cost is still relatively high and this cost hampers the development of a economical system for the production of sialylooligosaccharides.

Also, strains like *E. coli* K1 and *N. meningitidis* are able to produce CMP-Neu5Ac but they are pathogenic and cannot be used in biotechnological processes for safety reasons. Most of other bacteria, including *E. coli* K12, do not have the enzymatic machinery for the biosynthesis of CMP-Neu5Ac, and it is a goal of the invention to genetically engineer non pathogenic strains which would be able to produce CMP-Neu5Ac from endogenous UDP-GlcNAc.

In connection with the present invention, we have designed a new microbial system for cost-effective large scale production of sialylooligosaccharides without the need of an exogenous supply of sialic acid. The metabolically engineered microorganisms of the invention are viable, non- pathogenic and can be used in large scale and industrial culture processes. They have optimized modified pathways and deletion of futile metabolic cycles and they lead to biosynthesis of activated CMP-Neu5Ac which serves as in situ sialic acid donnor to form sialylated oligosaccharides.

BRIEF SUMMARY OF THE INVENTION

The present invention provides a method of producing sialylated oligosaccharides by fermentative growth of microorganisms. In particular, the invention relates to a method of synthesis of oligosaccharides bearing one or several sialic acid residu(s), without any exogenous sialic acid addition to the culture medium, including but not limited to:

- oligosaccharide moieties of the gangliosides selected from

GM3 (3'sialyllactose, Neu5Acα–3Galβ–4Glc) and oligosaccharides comprising the GM3 motif,

GD3 Neu5Acα–8Neu5Acα–3Galβ–4Glc

GT3 (Neu5Acα–8Neu5Acα–8Neu5Acα–3Galβ–4Glc);

GM2 GalNAcβ-4(Neu5Acα-3)Galβ-4Glc,

GM1 Galβ-3GalNAcβ-4(Neu5Acα-3)Galβ-4Glc,

GD1a Neu5Acα-3Galβ-3GalNAcβ-4(Neu5Acα-3)Galβ-4Glc

GT1a Neu5Acα-8Neu5Acα-3Galβ-3GalNAcβ-4(Neu5Acα-3)Galβ-4Glc

GD2 GalNAcβ-4(Neu5Acα-8Neu5Acα3)Galβ-4Glc

GT2 GalNAcβ-4(Neu5Acα-8Neu5Acα-8Neu5Acα3)Galβ-4Glc

GD1b, Galβ-3GalNAcβ-4(Neu5Acα-8Neu5Acα3)Galβ-4Glc

GT1b Neu5Acα-3Galβ-3GalNAcβ-4(Neu5Acα-8Neu5Acα3)Galβ-4Glc

GQ1b Neu5Acα-8Neu5Acα-3Galβ-3GalNAcβ-4(Neu5Acα-8Neu5Acα3)Galβ-4Glc

GT1c Galβ-3GalNAcβ-4(Neu5Acα-8Neu5Acα-8Neu5Acα3)Galβ-4Glc

GQ1c, Neu5Acα-3Galβ-3GalNAcβ-4(Neu5Acα-8Neu5Acα-8Neu5Acα3)Galβ-4Glc

GP1c Neu5Acα-8Neu5Acα-3Galβ-3GalNAcβ-4(Neu5Acα-8Neu5Acα-8Neu5Acα3)Galβ-4Glc

GD1α Neu5Acα-3Galβ-3(Neu5Acα-6)GalNAcβ-4Galβ-4Glc

Fucosyl-GM1 Fucα-2Galβ-3GalNAcβ-4(Neu5Acα-3)Galβ-4Glc ;

all of which may be extended to the production of the corresponding gangliosides by reacting the above oligosaccharide moities with ceramide.

- other sialylated sugars including :

- **6'sialyllactose** (Neu5Acα-6Galβ-4Glc) and oligosaccharides comprising **6'sialyllactose**
- **SGG hexasaccharide** (Neu5Acα-3Galβ-3GalNacβ-3Galα-4Galβ-4Gal)
- **Sialylated tetrasaccharide** (Neu5Acα-3Galβ-4GlcNacβ-4GlcNAc)
- **pentasaccharide LST_D** (Neu5Acα-3Galβ-4GlcNacβ-3Galβ-4Glc)

In one particular aspect, the process of the invention is based on the active uptake of an exogenous precursor, such as for example a mono, di or tri-saccharide, more particularly an exogenous precursor selected from lactose, galactose, β-galactoside, and α-galactoside such as globotriose (Galα–4Galβ–4Glc), while the cells are growing on an alternative carbon substrate, such as glycerol or glucose. The expression "exogenous precursor" is intended to denote a compound involved in the biosynthetic pathway of the oligosaccharide according to the invention that is internalized by the cells.

It also provides metabolically engineered microorganisms that can specifically produce the above sialylated oligosaccharides without side products such as GA1, GA2, GA3, GA4, and GA5 and to the use of the *cstIII* gene isolated from *C. jejuni* strains expressing lipooligosaccharide structures that mimic the GM1 ganglioside, such as the *C. jejuni* strain NCTC Accession No 111168, for the specific production of GM1.

DESCPRIPTION OF THE DRAWINGS

Figure 1 shows the production of sialyllactose by metabolically engineered E. coli strain. **Figure 2** shows the formation of side-products during production of GM1 sugar by metabolically engineered *E. coli* strains

Figure 3 shows the formation of side-products during biosynthesis of GM2 sugar

Figure 4 shows the UDP-Gal biosynthetic pathway

Figure 5 shows the formation of side-products during biosynthesis of GD2 sugar

Figure 6 shows the strategy for GT1a production using sialyltransferase

Figure 7 shows the metabolically engineered pathway for the production of the LST_D sugar

(Neu5Acα-3Galβ-4GlcNacβ-3β-4Gal) from exogenous lactose

Figure 8 shows the metabolically engineered pathway for the production of the SGG hexasaccharide (Neu5Acα-3Galβ-3GalNacβ-3Galα-4Galβ-4Gal) from exogenous globotriose.

Figure 9 shows the metabolically engineered pathway for the production of the sialylated tetrasaccharide (Neu5Acα-3Galβ-4GlcNacβ-4GlcNAc).

Figure 10 is a TLC analysis of intracellular and extracellular fraction of high cell density culture of strain DC7 with a continous feeding of lactose. Lanes 1: standard solution (2 $mg.ml^{-1}$ each) of lactose, lacto-*N*-neotetraose (LNnT), lacto-*N*-neohexaose (LNnH). Lanes 2, 3, 4, 5, 6 and 7 : intracellular fractions withdrawn 0, 7, 23, 31, 47 and 54 hours after lactose addition. Lanes 8, 9, 10, 11, 12 and 13 : extracellular fractions withdrawn 0, 7, 23, 31, 47 and 54 hours after lactose addition. Sialyllactose (**2**) has been previously shown to migrate as the tetrasaccharide LNnT.

Figure 11 shows the production of sialyllactose in high cell density culture of strain DC7 with a continous feeding of lactose. (▲) cumulated amount of added lactose, (□) intracellular sialyllactose, (■) extracellular sailayllactose, (—) bacterial growth.

Figure 12 is a TLC analysis of oligosaccharides produced by high cell density culture of strain NF03 containing the plasmids pUC18-cstII and pBBR3-SS. The initial lactose concentration was 3 $g.l^{-1}$. Lanes 1: standard solution (2 $mg.ml^{-1}$ each) of lactose lacto-*N*-neotetraose (LNnT), lacto-*N*-neohexaose (LNnH). Lanes 2 and 3:

intracellular fractions withdrawn 7 and 24 hours after lactose addition. Lanes 4 and 5: extracellular fractions withdrawn 7 and 24 hours after lactose addition.

Figure 13 is a TLC analysis of oligosaccharides produced by high cell density culture of strain DC15 (pBS-cgtAII-nst, pBBR3-SS-wbpP, pSU-cgtB). The initial lactose concentration was 5 $g.l^{-1}$. Lanes 1: standard solution (2 $mg.ml^{-1}$ each) of lactose lacto-*N*-neotetraose (LNnT), lacto-*N*-neohexaose (LNnH). Lanes 2,3,4,5 : intracellular fractions withdrawn 7, 20,30 and 44 hours after lactose addition. Lanes 6,7,8,9 : extracellular fractions withdrawn 7, 20,30 and 44 hours after lactose addition.. Sialyllactose (**1**) and the GM1 sugar (**3**) migrate as LNnT and LNnH respectively.

Figure 14 is a TLC analysis of oligosaccharides produced by high cell density culture of strain DC21 (pBS-cgtAII-nst, pBBR3-SS-gne, pSU-cgtB). The initial lactose concentration was 5 $g.l^{-1}$. Lanes 1 an 10: standard solution (2 $mg.ml^{-1}$ each) of lactose lacto-*N*-neotetraose (LNnT), lacto-*N*-neohexaose (LNnH). Lanes 2,3,4,5: intracellular fractions withdrawn 5, 20,28 and 44 hours after lactose addition. Lanes 6,7,8,9: extracellular fractions withdrawn 5, 20,28 and 44 hours after lactose addition.. Sialyllactose (**1**) and the GM1 sugar (**3**) migrate as LNnT and LNnH respectively.

Figure 15 is a TLC analysis of oligosaccharides produced by high cell density culture of strain DC22. (pBS-cstIII-cgtAII, pBBR3-SS-gne, pSU-cgtB) The initial lactose concentration was 5 $g.l^{-1}$ .Lanes 1 and 10: standard solution (2 $mg.ml^{-1}$ each) of lactose lacto-*N*-neotetraose (LNnT), lacto-*N*-neohexaose (LNnH). Lanes 2,3,4,5: intracellular fractions withdrawn 7, 20,30 and 44 hours after lactose addition. Lanes 6,7,8,9: extracellular fractions withdrawn 7, 20,30 and 44 hours after lactose addition.. Sialyllactose (**1**) and the GM1 sugar (**3**) migrate as LNnT and LNnH respectively.

Figure 16 is a TLC analysis of oligosaccharides produced by high cell density culture of strain ZWT (ZLKA, pBS-nst, pBBR3-SS-gne, pWKS-cgtAII). The initial lactose concentration was 5 $g.l^{-1}$. Lanes 1: standard solution (2 $mg.ml^{-1}$ each) of lactose, lacto-*N*-neotetraose (LNnT) lacto-*N*-neohexaose (LNnH). Lanes 2,3,4, : intracellular fractions withdrawn 0, 7 and 22 hours after lactose addition. Lanes 5, 6,7,: extracellular fractions withdrawn 0, 7, and 22 hours after lactose addition.

Figure 17 is a TLC analysis of oligosaccharides produced by high cell density culture of strain ZWU2 (ZWU, pBS-nst, pBBR3-SS-gne, pWKS-cgtAII). The initial lactose concentration was 5 $g.l^{-1}$. Lanes 1: standard solution (2 $mg.ml^{-1}$ each) of lactose, lacto-*N*-

neotetraose (LNnT) lacto-*N*-neohexaose (LNnH). Lanes 2,3,4, and 5 : intracellular fractions withdrawn 0,7, 22 and 30 hours after lactose addition. Lanes 6, 7, 8 and 9 : extracellular fractions withdrawn 0, 7, 22 and 30 hours after lactose addition.

Figure 18 shows a ESI⁻ mass spectrum of fraction compound (**2**) purified from the intracellular fraction of control strain ZWT (**A**) and *galU* mutant strain ZWU2 (**B**) Peak at m/z 835 corresponds to GM2 sugar and peak at m/z 794 corresponds to the galactosylated analog (GA2 sugar).

Figure 19 is a TLC analysis of fractions obtained after separation of the intracellular fraction from a one liter culture of strain ZWU1. The separation was carried out on Dowex 1 ($HCO3^-$ form) using a 0-1M $NaHCO_3$ gradient. The volume of each tube was 10 ml. The yield of faction A, B, C and D were 0.5 g, 0.85, 0.6 and 1.2 g respectively.

Figure 20 is a TLC analysis of oligosaccharides produced by high cell density culture of strain NF17 (ZLKA, pBS-cgtA-cstII, pSU-cgtB, pBBR-SS-gne). The initial lactose concentration was 5 $g.l^{-1}$. Lanes 1 and 10: standard solution (2 $mg.ml^{-1}$ each) of lactose, lacto-*N*-neotetraose (LNnT) lacto-*N*-neohexaose (LNnH). Lanes 2, 3, 4, and 5: intracellular fractions withdrawn 7, 22, 30 and 46 hours after lactose addition. Lanes 6, 7, 8, and 9: extracellular fractions withdrawn 7, 22, 30 and 46 hours after lactose addition.

DISCLOSURE OF ORIGIN OF GENETIC MATERIAL

Table 1. Genes, plasmids and *Escherichia coli* strains used in present invention

	Description	Reference or source
Genes		
nst	α-2,3-Sialyltransferase from *N. meningitidis* L3 strain MC58	U60660
neuA	CMP-Neu5Ac synthetase from *C. jejuni* strain ATCC 43438	AF400048
neuB	Sialic acid synthase from *C. jejuni* strain ATCC 43438	AF400048
neuC	GlcNAc-6-phosphate 2 epimerase from *C. jejuni* strain ATCC 43438	AF400048
cgtA	β-4 GalNAc transferase from *C. jejuni* O:19 strain OH4384	AF130984
wbpP	UDP-GlcNAc C4 epimerase from *P. aeruginosa*	AF035937
gne (Cj1131c)	UDP-GlcNAc C4 epimerase from *C. jejuni* O:2 strain NCTC 11168	AL139077
cstIII (Cj1140)	α-2,3 Sialyltransferase from *C. jejuni* O:2 strain NCTC 11168	AL139077
cgtAII	β-4 GalNAc transferase from *C. jejuni* O:36 strain ATCC 43456	AF401528
cgtB (Cj1139c)	β-3 Gal transferase from *C. jejuni* O:2 strain NCTC 11168	AL139077
cstII	α-2,3 α-2,8-Sialyltransferase from *C. jejuni* strain ATCC43438	AF400048
nodC	*N*-acetylglucosaminyltransferase from *Azorhizobium. caulinodans*	AAB51164
lgtB	β1,4-galactosyltransferase from *Neisseria meningitidis*	AAC44085
chiA	chitinase from *Bacillus circulans*	AAA81528
Plasmids		
pWKS130	Cloning vector, Km^r, Plac promoter, low copy number, pSC101 replicon	(Wang & Kushner, 1991)
pSU27-18	pACYC184-derived cloning vectors Cm^r P_{lac} promoter,	(Martinez *et al.*, 1988)
pBAD33	pACYC184-derived cloning vectors Cm^r P_{ara} promoter,	Guzman et al 1995
pBS-nst	pBluescript II SK derivative carrying *nst* (previously called NST-01)	(Priem *et al.*, 2002)
pBBR1MCS-3	Cloning vector, Tc^r, P_{lac} promoter, low copy number,	(Kovach *et al.*, 1995)
pBBR3-SS	pBBR1MCS-3 derivative carrying *neuABC*	present invention
pBBR3-SS-wbpP	pBBR1MCS-3 derivative carrying *neuABC* and wbpP	present invention
pBBR3-SS-gne	pBBR1MCS-3 derivative carrying *neuABC* and *gne*	present invention
pUC-cstII	pUC18 derivative carrying *cstII*	present invention
pBS-cgtAII-nst	pBluescript II SK derivative carrying *cgtAII* and *nst*	present invention
pBS-cgtA-cstII	pBluescript II KS derivative carrying, *cgtA*, *cstII*	present invention

pSU18-cgtB	pSU27 18 derivative carrying *cgtB*	present invention
pBS-cstIII-cgtAII	pBluescript II KS derivative carrying, *cstIII and cgtAII*	present invention
pWKS-cgtAII	pWKS130 derivative carrying *cgtAII*,	present invention
pBAD33-cgtAII	pBAD33 derivative carrying *cgtAII*,	present invention
pWKS-lgtB-chiA	pWKS130 derivative carrying *nst* and the chitinase *chiA*	(Dumon *et al.*, 2005)
pBS-nst-nodC	pBluescript II SK derivative carrying *A. caulinodans nodC* and *nst*	present invention
pBS-cstII-cgtB	pBluescript II KS derivative carrying *cstII* and *cgtB*	present invention
Strains		
DC	DH1 *lacZ lacA*	(Dumon *et al.*, 2005)
GLK	DH1 *lacZ lacA galK*	(Dumon *et al.*, 2005)
ZLKA	DC *ΔnanKETA*	present invention
AZL	DC *nanA*	present invention
AZK	DC *ΔnanK nanA*	present invention
ZWU	ZLKA *galU*	present invention
GLKA	GLK *ΔnanKETA*	present invention
ZW	ZLKA *melA wcaJ*	present invention
DC6	DC (pBS-nst, pBBR3-SS)	present invention
AW1	AZL (pBS-nst, pBBR3-SS)	present invention
DC7	ZLKA (pBS-nst, pBBR3-SS)	present invention
DC7	ZLKA (pBS-nst, pBBR3-SS)	present invention
DC0	ZLKA (pBS-nst)	present invention
AZK1	AZK (pBS-nst, pBBR3-SS)	present invention
NF3	ZLKA (pUC-cstII, pBBR3-SS)	present invention
DC15	ZLKA (pBS-cgtAII-nst, pBBR3-SS-wbpP, pSU-cgtB)	present invention
DC21	ZLKA (pBS-cgtAII-nst, pBBR3-SS-gne, pSU-cgtB)	present invention

DC22	ZLKA (pBS-cstIII-cgtAII, pBBR3-SS-gne, pSU-cgtB)	present invention
ZWT	ZLKA (pBS-nst, pBBR-SS-gne, pWKS-cgtAII)	present invention
ZWU2	ZWU (pBS-nst, pBBR-SS-gne, pWKS-cgtAII)	present invention
NF08	ZLKA (pUC18-cstII, pBBR3-SS-gne, pWKS-cgtAII)	present invention
NF09	ZLKA (pUC18-cstII, pBBR3-SS-gne, pBAD33-cgtAII)	present invention
ZWU1	ZWU (pUC18-cstII, pBBR3-SS-gne, pBAD33-cgtAII)	present invention
NF17	ZLKA (pBS-cgtA-cstII, pSU-cgtB, pBBR-SS-gne)	present invention
GLK7	GLKA (pBS-nst, pBBR3-SS)	present invention
SN4	ZLKA (pBS-nst-nodC, pBBR3-SS, pWKS-lgtB-chiA)	present invention
NF21	ZLKA (pBS-cstII-cgtB, pBBR3-SS-gne, pBAD33-cgtAII)	present invention

DETAILED DESCRIPTION OF THE INVENTION

In a first embodiment, the invention relates to a method for producing oligosaccharides comprising at least one sialic acid residu, herein referred to sialylated oligosaccharides, the method comprising the step consisting of culturing a microorganism in a culture medium, optionally comprising an exogenous precursor, wherein said microorganism is capable of producing internally activated sialic acid and comprises heterologous genes encoding a CMP-Neu5Ac synthetase, a sialic acid synthase, a GlcNAc-6-phosphate 2 epimerase and a sialyltransferase, and wherein the endogenous genes coding for sialic acid aldolase (*NanA*) and for ManNac kinase (*NanK*) have been deleted or inactivated.

In the above method, and depending on the end-point, the heterologous sialyltransferase gene may be selected from α-2,3-Sialyltransferase, for example encoded by *nst*, α-2,3 α-2,8-Sialyltransferase (*cstII*), and α-2,3-Sialyltransferase (*cstIII*) or α-2,6-Sialyltransferase. The heterologous CMP-Neu5Ac synthetase may be *neuA*, the heterologous sialic acid synthase may be *neuB*, and the heterologous GlcNAc-6-phosphate 2 epimerase may be *neuC*. The *neuA*, *neuB*, and the *neuC genes* can be

isolated from bacterial strains that contain sialylated structure in their cells envelope, such as *C. jejuni* strain ATCC Accession No. 43438.

The *nanT, nanA, nanK* and *nanE* genes are part of the same operon, which is regulated by the DNA binding protein NanR and induced by Neu5Ac (Kalivoda *et al.*, 2003). Thus, the microorganisms of the invention can also be *nanKEAT*-. The production of Neu5Ac as intermediate during the synthesis of CMP-Neu5Ac by genetically engineered strain ovexpressing the *neuBCA* genes can thus induce the pathway of sialic acid catabolism and create two futile cycles that will reduce the capacity of CMP-Neu5Ac biosynthesis of the bacteria. A first futile cycle can result from the combined activity of the sialic acid synthase NeuB with the sialic acid aldolase NanA. A second futile cycle can result from the combined action of the UDP-GlcNAc 2 epimerase NeuC with the four enzymes NanK NanE NagA GlmM and GlmU that catalyse the formation of UDP-GlcNAc from ManNAc. According to the method proposed herein, degradation of Neu5Ac and ManNAc is prevented. This can be advantageously done by disrupting the *nanA* and *nanK* genes in the strains which will be used for sialylooligosaccharides production. In one specific embodiment, the *nanT, nanA, nanK* and *nanE* genes are deleted or inactivated. This can be practiced by removing the all operon for example.

In a preferred embodiment, the above microorganism encodes a protein that facilitates uptake of lactose and lacks enzymes that metabolize lactose. For example, in E. coli, the cell is preferrably *LacY*+ (β-galactoside permease), *LacZ*- (β galactosidase), and optionally *MelA*- (α-galactosidase).

In another preferred embodiment, the medium comprises an exogenous precursor which is selected for example from lactose, galactose, β-galactoside, and α–galactoside, such as globotriose (Galα–4Galβ–4Glc).

The invention also relates to the above microorganism and to a cell culture medium comprising the above microorganism and an exogenous precursor selected from lactose, galactose, β-galactoside, and α–galactoside, such as globotriose (Galα–4Galβ–4Glc).

DEFINITIONS

The term "sialic acid" refers to any member of a family of nine-carbon carboxylated sugars. The most common member of the sialic acid family is N-acetyl-neuraminic acid (2-keto-5-acetamido-3,5-dideoxy-D-glycero-D-galactononulopyranos-1-onic acid (often abbreviated as Neu5Ac, Neu5Ac, or NANA). A second member of the family is N-glycolyl-neuraminic acid (Neu5Gc or NeuGc), in which the N-acetyl group of Neu5Ac is hydroxylated. A third sialic acid family member is 2-keto-3-deoxy-nonulosonic acid (KDN) (Nadano *et al.* (1986) *J. Biol. Chem.* **261**: 11550-11557; Kanamori *et al.*, *J. Biol. Chem.* **265**: 21811-21819 (1990)). Also included are 9-substituted sialic acids such as a 9-O-C_1-C_6 acyl-Neu5Ac like 9-O-lactyl-Neu5Ac or 9-O-acetyl-Neu5Ac, 9-deoxy-9-fluoro-Neu5Ac and 9-azido-9-deoxy-Neu5Ac. For review of the sialic acid family, *see*, *e.g.*, Varki, *Glycobiology* **2**: 25-40 (1992); *Sialic Acids: Chemistry, Metabolism and Function*, R. Schauer, Ed. (Springer-Verlag, New York (1992)). The synthesis and use of sialic acid compounds in a sialylation procedure is disclosed in international application WO 92/16640, published October 1, 1992.

The term "bifunctional *Campylobacter jejuni* CstII sialyltransferase" refers to a sialyltransferase which exhibits both α-2,3 and α-2,8-Sialyltransferase activities. In some embodiments, the CstII sialyltransferase from ATCC Accession No. 43438 is used.

An "acceptor substrate" or an "acceptor saccharide" for a glycosyltransferase is an oligosaccharide moiety that can act as an acceptor for a particular glycosyltransferase. When the acceptor substrate is contacted with the corresponding glycosyltransferase and sugar donor substrate, and other necessary reaction mixture components, and the reaction mixture is incubated for a sufficient period of time, the glycosyltransferase transfers sugar residus from the sugar donor substrate to the acceptor substrate. For example, an acceptor substrate for the sialyltransferases used in the methods of the invention is lactose Galβ1,4-Glc.

A "donor substrate" for glycosyltransferases is an activated nucleotide sugar. Such activated sugars generally consist of uridine, guanosine, and cytidine monophosphate derivatives of the sugars (UMP, GMP and CMP, respectively) or diphosphate derivatives of the sugars (UDP, GDP and CDP, respectively) in which the nucleoside

monophosphate or diphosphate serves as a leaving group. For example, a donor substrate for sialyltransferases used in the methods of the invention is CMP-Neu5Ac.

A "culture medium" refers to any liquid, semi-solid or solid media that can be used to support the growth of a microorganism used in the methods of the invention. In some embodiments, the microorganism is a bacteria, *e.g.*, *E. coli*. Media for growing microorganisms are well known, see, *e.g.*, Sambrook *et al.* and Current Protocols in Molecular Biology, F.M. Ausubel *et al.*, eds., *Current Protocols*, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1998 Supplement) (Ausubel). Media can be rich media, *e.g.*, Luria broth or terrific broth, or synthetic or semi-synthetic medium, *e.g.*, M9 medium. In some preferred embodiments the growth medium comprises lactose and sialic acid.

"Commercial scale" refers to gram scale production of a sialylated product saccharide in a single reaction. In preferred embodiments, commercial scale refers to production of greater than about 50, 75, 80, 90 or 100, 125, 150, 175, or 200 grams.

The term "operably linked" refers to functional linkage between a nucleic acid expression control sequence (such as a promoter, signal sequence, or array of transcription factor binding sites) and a second nucleic acid sequence, wherein the expression control sequence affects transcription and/or translation of the nucleic acid corresponding to the second sequence.

A "heterologous polynucleotide" or a "heterologous gene", as used herein, is one that originates from a source foreign to the particular host cell, or, if from the same source, is modified from its original form. Thus, a heterologous sialyltransferase gene in a cell includes a gene that is endogenous to the particular host cell but has been modified. Modification of the heterologous sequence may occur, *e.g.*, by treating the DNA with a restriction enzyme to generate a DNA fragment that is capable of being operably linked to a promoter. Techniques such as site-directed mutagenesis are also useful for modifying a heterologous sequence.

A "recombinant expression cassette" or simply an "expression cassette" is a nucleic acid construct, generated recombinantly or synthetically, with nucleic acid elements that are

capable of affecting expression of a structural gene in hosts compatible with such sequences. Expression cassettes include at least promoters and optionally, transcription termination signals. Typically, the recombinant expression cassette includes a nucleic acid to be transcribed (*e.g.*, a nucleic acid encoding a desired polypeptide), and a promoter. Additional factors necessary or helpful in effecting expression may also be used. Transcription termination signals, enhancers, and other nucleic acid sequences that influence gene expression, can also be included in an expression cassette. When more than one heterologous protein is expressed in a microorganism, the genes encoding the proteins can be expressed on a single expression cassette or on multiple expression cassettes that are compatible and can be maintained in the same cell. As used herein, expression cassette also encompasses nucleic acid constructs that are inserted into the chromosome of the host microorganism. Those of skill are aware that insertion of a nucleic acid into a chromosome can occur, *e.g.*, by homologous recombination. An expression cassette can be constructed for production of more than one protein. The proteins can be regulated by a single promoter sequence, as for example, an operon. Or multiple proteins can be encoded by nucleic acids with individual promoters and ribosome binding sites.

The term "isolated" refers to material that is substantially or essentially free from components which interfere with the activity biological molecule. For cells, saccharides, nucleic acids, and polypeptides of the invention, the term "isolated" refers to material that is substantially or essentially free from components which normally accompany the material as found in its native state. Typically, isolated saccharides, oligosaccharides, proteins or nucleic acids of the invention are at least about 50%, 55%, 60%, 65%, 70%, 75%, 80% or 85% pure, usually at least about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% pure as measured by band intensity on a silver stained gel or other method for determining purity. Purity or homogeneity can be indicated by a number of means well known in the art, such as polyacrylamide gel electrophoresis of a protein or nucleic acid sample, followed by visualization upon staining. For certain purposes high resolution will be needed and HPLC or a similar means for purification utilized. For oligosaccharides, *e.g.*, sialylated products, purity can be determined using, *e.g.*, thin layer chromatography, HPLC, or mass spectroscopy.

The terms "identical" or percent "identity," in the context of two or more nucleic acid or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residus or nucleotides that are the same, when compared and aligned for maximum correspondence, as measured using one of the following sequence comparison algorithms or by visual inspection.

The phrase "substantially identical," in the context of two nucleic acids or polypeptides, refers to two or more sequences or subsequences that have at least 60%, preferably 80% or 85%, most preferably at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% nucleotide or amino acid residu identity, when compared and aligned for maximum correspondence, as measured using one of the following sequence comparison algorithms or by visual inspection. Preferably, the substantial identity exists over a region of the sequences that is at least about 50 residus in length, more preferably over a region of at least about 100 residus, and most preferably the sequences are substantially identical over at least about 150 residus. In a most preferred embodiment, the sequences are substantially identical over the entire length of the coding regions.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

Optimal alignment of sequences for comparison can be conducted, *e.g.*, by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by visual inspection (*see generally*, *Current Protocols in Molecular Biology*, F.M. Ausubel *et al.*, eds., Current

Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1995 Supplement) (Ausubel)).

Examples of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.* (1990) *J. Mol. Biol.* 215: 403-410 and Altschuel *et al.* (1977) *Nucleic Acids Res.* 25: 3389-3402, respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residus; always > 0) and N (penalty score for mismatching residus; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residu alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)).

In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (*see, e.g.*, Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)),

which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

"Conservatively modified variations" of a particular polynucleotide sequence refers to those polynucleotides that encode identical or essentially identical amino acid sequences, or where the polynucleotide does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given polypeptide. For instance, the codons CGU, CGC, CGA, CGG, AGA, and AGG all encode the amino acid arginine. Thus, at every position where an arginine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent substitutions" or "silent variations," which are one species of "conservatively modified variations." Every polynucleotide sequence described herein which encodes a polypeptide also describes every possible silent variation, except where otherwise noted. Thus, silent substitutions are an implied feature of every nucleic acid sequence which encodes an amino acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine) can be modified to yield a functionally identical molecule by standard techniques. In some embodiments, the nucleotide sequences that encode the enzymes are preferably optimized for expression in a particular

Similarly, "conservative amino acid substitutions," in one or a few amino acids in an amino acid sequence are substituted with different amino acids with highly similar properties are also readily identified as being highly similar to a particular amino acid sequence, or to a particular nucleic acid sequence which encodes an amino acid. Such conservatively substituted variations of any particular sequence are a feature of the present invention. Individual substitutions, deletions or additions which alter, add or delete a single amino acid or a small percentage of amino acids (typically less than 5%, more typically less than 1%) in an encoded sequence are "conservatively modified variations" where the alterations result in the substitution of an amino acid with a

chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. *See, e.g.*, Creighton (1984) *Proteins*, W.H. Freeman and Company.

BIFUNCTIONAL SIALYLTRANSFERASES

As noted above, bifunctional sialyltransferases are used in the methods of the invention. Nucelic acids encoding such enzymes have been isolated from *C. jejuni* and are disclosed in U.S. Patent Nos. 6,699,705 and 6503744 and WO/02074942. Exemplary *C. jejuni* strains which can be used as sources of bifunctional sialyltransferases include OH4384 (nucleic acid sequences are found in GenBank accessions AR271700 and AX934425), OH4382, O:10 (nucleic acid sequences are found in GenBank accessions AR271701 (SEQ ID No 1), AX934427 (SEQ ID No 2), O:23, and O:41 (nucleic acid sequences are found in GenBank accessions AR271702 (SEQ ID No 3) and AX934429 (SEQ ID No 4)). It shall be understood that conservatively modified variations as defined above of SEQ ID No 1, 2, 3 and 4 may be applied herein.

HOST CELLS

The recombinant cells of the invention are generally made by creating or otherwise obtaining a polynucleotide that encodes the particular enzyme(s) of interest, placing the polynucleotide in an expression cassette under the control of a promoter and other appropriate control signals, and introducing the expression cassette into a cell. More than one of the enzymes can be expressed in the same host cells using a variety of methods. For example, a single extrachromosomal vector can include multiple expression cassettes or more that one compatible extrachromosomal vector can be used maintain an expression cassette in a host cell. Expression cassettes can also be inserted into a host cell chromosome, using methods known to those of skill in the art. Those of skill will recognize that combinations of expression cassettes in extrachromosomal vectors and expression cassettes inserted into a host cell chromosome can also be used. Other modification of the host cell, described in detail below, can be performed to enhance production of the desired oligosaccharide. For example, the microorganism may be LacY+ allowing active transport of lactose. Host cells don't need to be NanT+

since activated sialylic acid is produced internally with the method according to the invention.

The recombinant cells of the invention are generally microorganisms, such as, for example, yeast cells, bacterial cells, or fungal cells. Examples of suitable cells include, for example, *Azotobacter sp.* (*e.g.*, *A. vinelandii*), *Pseudomonas sp.*, *Rhizobium sp.*, *Erwinia sp.*, *Bacillus sp.*, *Streptomyces sp.*, *Escherichia sp.* (*e.g.*, *E. coli*), and *Klebsiella sp.*, among many others. The cells can be of any of several genera, including Saccharomyces (*e.g.*, *S. cerevisiae*), Candida (*e.g.*, *C. utilis, C. parapsilosis, C. krusei, C. versatilis, C. lipolytica, C. zeylanoides, C. guilliermondii, C. albicans, and C. humicola*), Pichia (*e.g.*, *P. farinosa and P. ohmeri*), Torulopsis (*e.g.*, *T. candida, T. sphaerica, T. xylinus, T. famata, and T. versatilis*), Debaryomyces (*e.g.*, *D. subglobosus, D. cantarellii, D. globosus, D. hansenii, and D. japonicus*), Zygosaccharomyces (*e.g.*, *Z. rouxii and Z. bailii*), Kluyveromyces (*e.g.*, *K. marxianus*), Hansenula (*e.g.*, *H. anomala and H. jadinii*), and Brettanomyces (*e.g.*, *B. lambicus and B. anomalus*).

Promoters for use in *E. coli* include the T7, trp, or lambda promoters. A ribosome binding site and preferably a transcription termination signal are also provided. For expression of heterologous proteins in prokaryotic cells other than *E. coli*, a promoter that functions in the particular prokaryotic species is required. Such promoters can be obtained from genes that have been cloned from the species, or heterologous promoters can be used. For example, the hybrid trp-lac promoter functions in *Bacillus* in addition to *E. coli*. Methods of transforming prokaryotes other than E. coli are well known. For example, methods of transforming Bacillus species and promoters that can be used to express proteins are taught in U.S. Patent No. 6,255,076 and U.S. Patent No. 6,770,475.

In yeast, convenient promoters include GAL1-10 (Johnson and Davies (1984) *Mol. Cell. Biol.* 4:1440-1448) ADH2 (Russell *et al.* (1983) *J. Biol. Chem.* 258:2674-2682), PHO5 (*EMBO J.* (1982) 6:675-680), and MFα (Herskowitz and Oshima (1982) in *The Molecular Biology of the Yeast Saccharomyces* (eds. Strathern, Jones, and Broach) Cold Spring Harbor Lab., Cold Spring Harbor, N.Y., pp. 181-209). Another suitable promoter for use in yeast is the ADH2/GAPDH hybrid promoter as described in Cousens *et al.*, *Gene* 61:265-275 (1987). For filamentous fungi such as, for example, strains of the fungi Aspergillus (McKnight *et al.*, U.S. Patent No. 4,935,349), examples

of useful promoters include those derived from *Aspergillus nidulans* glycolytic genes, such as the ADH3 promoter (McKnight *et al.*, *EMBO J.* 4: 2093 2099 (1985)) and the *tpi*A promoter. An example of a suitable terminator is the ADH3 terminator (McKnight *et al.*).

In some embodiments, the polynucleotides are placed under the control of an inducible promoter, which is a promoter that directs expression of a gene where the level of expression is alterable by environmental or developmental factors such as, for example, temperature, pH, anaerobic or aerobic conditions, light, transcription factors and chemicals. Such promoters are referred to herein as "inducible" promoters, which allow one to control the timing of expression of the glycosyltransferase or enzyme involved in nucleotide sugar synthesis. For *E. coli* and other bacterial host cells, inducible promoters are known to those of skill in the art. These include, for example, the *lac* promoter. A particularly preferred inducible promoter for expression in prokaryotes is a dual promoter that includes a *tac* promoter component linked to a promoter component obtained from a gene or genes that encode enzymes involved in galactose metabolism (*e.g.*, a promoter from a UDPgalactose 4-epimerase gene (*gal*E)).

Inducible promoters for other organisms are also well known to those of skill in the art. These include, for example, the arabinose promoter, the lacZ promoter, the metallothionein promoter, and the heat shock promoter, as well as many others.

The construction of polynucleotide constructs generally requires the use of vectors able to replicate in bacteria. A plethora of kits are commercially available for the purification of plasmids from bacteria. For their proper use, follow the manufacturer's instructions (*see*, for example, EasyPrepJ, FlexiPrepJ, both from Pharmacia Biotech; StrataCleanJ, from Stratagene; and, QIAexpress Expression System, Qiagen). The isolated and purified plasmids can then be further manipulated to produce other plasmids, and used to transfect cells. Cloning in *Streptomyces* or *Bacillus* is also possible.

Selectable markers are often incorporated into the expression vectors used to construct the cells of the invention. These genes can encode a gene product, such as a protein, necessary for the survival or growth of transformed host cells grown in a selective culture medium. Host cells not transformed with the vector containing the selection gene will not survive in the culture medium. Typical selection genes encode proteins

that confer resistance to antibiotics or other toxins, such as ampicillin, neomycin, kanamycin, chloramphenicol, or tetracycline. Alternatively, selectable markers may encode proteins that complement auxotrophic deficiencies or supply critical nutrients not available from complex media, *e.g.*, the gene encoding D-alanine racemase for Bacilli. Often, the vector will have one selectable marker that is functional in, *e.g.*, *E. coli*, or other cells in which the vector is replicated prior to being introduced into the target cell. A number of selectable markers are known to those of skill in the art and are described for instance in Sambrook *et al.*, *supra*. A preferred selectable marker for use in bacterial cells is a kanamycin resistance marker (Vieira and Messing, *Gene* 19: 259 (1982)). Use of kanamycin selection is advantageous over, for example, ampicillin selection because ampicillin is quickly degraded by β-lactamase in culture medium, thus removing selective pressure and allowing the culture to become overgrown with cells that do not contain the vector.

Construction of suitable vectors containing one or more of the above listed components employs standard ligation techniques as described in the references cited above. Isolated plasmids or DNA fragments are cleaved, tailored, and re-ligated in the form desired to generate the plasmids required. To confirm correct sequences in plasmids constructed, the plasmids can be analyzed by standard techniques such as by restriction endonuclease digestion, and/or sequencing according to known methods. Molecular cloning techniques to achieve these ends are known in the art. A wide variety of cloning and *in vitro* amplification methods suitable for the construction of recombinant nucleic acids are well-known to persons of skill.

A variety of common vectors suitable for constructing the recombinant cells of the invention are well known in the art. For cloning in bacteria, common vectors include pBR322 derived vectors such as pBLUESCRIPT™, and λ-phage derived vectors. In yeast, vectors include Yeast Integrating plasmids (*e.g.,* YIp5) and Yeast Replicating plasmids (the YRp series plasmids) and pGPD-2.

The methods for introducing the expression vectors into a chosen host cell are not particularly critical, and such methods are known to those of skill in the art. For example, the expression vectors can be introduced into prokaryotic cells, including *E.*

coli, by calcium chloride transformation, and into eukaryotic cells by calcium phosphate treatment or electroporation. Other transformation methods are also suitable.

PREFERED EMBODIMENTS

Production of 3'sialyllactose.

3'sialyllactose production can be carried out in a metabolically engineered strain defined above that expresses A gene coding for an enzyme comprising a α-2,3 sialyltransferase activity, such as a α-2,3 sialyltransferase or a bifunctional α-2,3 and α-2,8 sialyltransferase that use lactose as acceptor. As indicated, this strain is devoid of sialic acid aldolase, ManNAc kinase and β-Galactosidase activity and expresses heterologous genes encoding CMP-Neu5Ac synthetase, a sialic acid synthase, and a GlcNAc-6-phosphate 2 epimerase. For large scale production of sialyllactose, this strain can be cultivated at high cell density on inexpensive substrate such as glucose or glycerol and fed with lactose which will be internalized by the lactose permease and sialylated by the recombinant sialyltransferase using the CMP-Neu5Ac endogenously generated fom UDP-GlcNAc as shown in figure 1.

Production of 6'sialyllactose

Here, the sialyltransferase gene is a gene encoding α-2,6 sialyltransferase such as the gene from *Photobacterium damsela* (Yamamoto *et al.*, 1998) which results in the production of 6'sialyllactose (Neu5Acα-6Galβ-4Glc).

Production of GD3 and GT3 sugar

GD3 (Neu5Acα-8Neu5Acα-3Galβ-4GlcCer) is a minor ganglioside found in most normal tissues in higher vertebrates including humans. The GD3 level has been shown to increase during some pathological situations, such as cancers (glioma, melanoma) and to have an important role in tumor angiogenesis Zeng, *et al. Cancer Res*, 60:6670 (2000). Thus, high scale and cost-effective production of GD3 is of particular interest. To reach this end-point, the heterologous sialyltransferase gene refered above is chosen from a gene encoding a bifunctional α-2,3 and α-2,8 sialyltransferase, for example such as the cstII gene from *Campylobacter jejuni* (Gilbert *et al.*, 2002 and deposited under

ATCC Accession No. 43438) and results in the production of GD3 and GT3 sugar. The bifunctional sialyltransferase polypeptide catalyzes the transfer of a sialyl moiety from an activated sialic acid molecule produced internally to the Neu5Acα-3Galβ-4Glc (GM3) to form Neu5Acα-8Neu5Acα-3Galβ-4Glc.This reaction may be further extended to produce GT3 which is the precursor of C series gangliosides which are the major constituents in adult fish brain and are found abundantly in fetal brains of higher vertebrates (Letinic, *et al. Neuroscience*, *86*, 1 (1998)). They are also found in various neuroectodermal tumors and there is thus potentially great interest in having easy access to the GT3 oligosaccharide. To this end, the method herein further comprises culturing the microorganism such that the bifunctional *Campylobacter jejuni* sialyltransferase polypeptide catalyzes the transfer of a sialyl moiety from an activated sialic acid molecule produced internally to the Neu5Acα-8Neu5Acα-3Galβ-4Glc (GD3) to form Neu5Ac5α-8Neu5Acα-8Neu5Acα-3Galβ-4Glc (GT3).

Production of GM1 sugar

As illustrated in figure 2, the system for sialyllactose production displayed above (see also figure 1) can be extended to the production of carbohydrate portion of the ganglioside GM1 by expressing the additional genes for a β-1,4GalNActransferase and for a β-1,3-Galactosyltransferase. Here, the microorganism of the invention is as displayed above in fig 1 and further comprises heterologous sequences encoding β-1,4GalNActransferase as well as β-1,3-Galactosyltransferase. In this embodiment, the β-1,4GalNActransferase transfers a UDP-GalNac residu to sialyllactose (GM3) to form GM2 and the β-1,3-galactosyltransferase transfers a Galactosyl residu to GM2 to form GM1. Strains may not be able to naturally produce UDP-GalNAc such as E. coli K12 strains. In this case, the strain of the invention can be complemented by a gene that encodes a UDP-GlcNAc 4 epimerase, such as for example the *wbpP* gene from *P. aeruginosa* (Creuzenet *et al.*, 2000) and the *gne* gene from *C. jejuni* (Bernatchez *et al.*, 2005). The GM1 sugar has a terminal non reducing galactose and this structure can be used as acceptor by α-2,3 sialyltransferase to produce the GD1a sugar. The GD1a sugar has the same terminal non reducing disaccharide structure than sialylactose and can be used as acceptor by the β-1,4GalNActransferase to form a heptasaccharide intermediate

which can be galactosylated into the octasaccharide represented as GA1 (Galβ-3GalNAcβ-4(Neu5Acα-3)Galβ-3GalNAcβ-4(Neu5Acα-3)Galβ-4Glc) in figure 2.

The formation of the GD1a and its larger derivatives reduce the production yield of the GM1 sugar and it is one particular embodiment of the invention to reduce or abolish the formation of these side products. To this end, we discovered an α-2,3 sialyltransferase which is not able to use the GM1 sugar as acceptor. The *C. jejuni* strain NCTC 111168 expresses lipooligosaccharide structures that mimic the GM1 ganglioside (Linton *et al.*, 2000). This lipooligosaccharide outer core structure was referred by Linton et al (2000) as the Outer core "VI". *C. jejuni* NCTC 111168 contains a gene called *cstIII* which encodes a protein showing a 53 % sequence identity with sialyltransferase (CstII) from other *C. jejuni* strains that express GD1a mimic (Gilbert *et al.*, 2002). The sialyltransferase activity of the CstIII protein allows advantageously the production of the GM1 sugar as the only oligosaccharide product. Thus, in one preferred embodiment, the invention contemplates the above method for the specific production of GM1, wherein the heterologous sialyl transferase is a α-2,3 sialyltransferase encoded by the *cstIII* gene isolated from the *C. jejuni* strains expressing lipooligosaccharide structures that mimic the GM1 ganglioside, such as the *C. jejuni* strain NCTC Accession No 111168.

Production of GM2 sugar

The system for sialyllactose production as described above can be extended to the production of carbohydrate portion of the ganglioside GM2 by expressing the additional genes for a β-1,4-GalNActransferase and for UDP-GlcNAc 4 epimerase. It has been previously reported that the CgtA β-1,4-GalNActransferase from C. jejuni exibits a β-1,4 Galactosyltransferase side activity which resulted in the production of the GM2 sugar analog designated as GA2 in figure 3 (Antoine *et al.*, 2003). The GA2 sugar contains a terminal non reducing galactose and we have found that this galactose can served as acceptor for sialyltransferase to produce a disialylated tetrasaccharide (GA5 , figure 3) which in turn can be converted into the pentasaccharide GA4 or GA3 (figure 3) by the very active CgtAII β-1,4-GalNActransferase.

The formation of the side products GA2, GA3, GA4 and GA5 reduces the GM2 sugar production yield and makes its purification more difficult. It is thus within the scope of the invention to abolish the formation of these side products. The β-1,4 Galactosyltransferase side activity can be advantageously suppressed by using mutant unable to produce UDP-Gal. As illustrated in figure 4, such mutant can be obtained by disrupting one of the three following genes: the *galE* gene encoding the UDP-glucose epimerase, the *galU* gene which encodes the UDP-Glc pyrophosphorylase, and the *pgm* gene that encodes the phosphoglucomutase. Thus, the invention is directed to a method as defined above for producing sialylated oligosaccharides, which is extended to the production of carbohydrate portion of the ganglioside GalNAcβ-4(Neu5Acα-3)Galβ-4Glc (GM2), wherein the microorganism further comprises heterologous sequences encoding a β-1,4-GalNActransferase, such as the *CgtAII* gene from *C. jejuni* O:36 strain ATCC Accession No 43456, and a UDP-GlcNAc 4 epimerase and wherein the micoorganism has at least one of the three following genes deleted or inactivated to disrupt the endogenous production of UDP-Gal: the *galE* gene encoding the UDP-glucose epimerase, the *galU* gene which encodes the UDP-Glc pyrophosphorylase, and the *pgm* gene which encodes the phosphoglucomutase, said disruption avoiding the production of side products such as GA2, GA3, GA4 and GA5.

Production of GD2 sugar

The system for the production of GD3 sugar can be extended to the production of carbohydrate portion of the ganglioside GD2 by expressing the additional heterologous genes coding for a UDP-GlcNAc 4 epimerase and a β-1,4-GalNActransferase that use the GD3 sugar as acceptor, such as the CgtAII protein from *C. jejuni* O:36 strain ATCC Accession No 43456.

In this system, both the CstII sialyltransferase and the CgtAII GalNAc transferase compete for utilization of sialyllactose as acceptor (figure 5). To favour the production of GD2 sugar, we reduced the expression of *cgtAII* in the first phase of the culture (to allow sialyllactose to be mainly converted into GD3 sugar) and then increase the *cgtAII* expression in a second phase to convert GD3 into GD2. This can be done by placing the *cgtAII* gene under the control of a promotor which is regulated independently from the promotors that control the other genes. The β-1,4-

Galactosyltransferase side activity of the CgtAII protein can result in the production of galactosylated analogs such as compounds GA2 GA5 and GA6 represented in figure 5. While these oligosaccharides may be of interest, it is also within the scope of the invention to prevent the formation of these side products by using mutant strain unable to produce UDP-Gal as depicted above for the GM2 production to lead to specific production of GD2.

Production of GD1b, GT1c, GT1b, GQ1b, GQ1c and GP1c sugars

The production of GD1b sugar (Galβ-3GalNAcβ-4(Neu5Acα-8Neu5Acα-3)Galβ-4Glc) and GT1c sugar (Galβ-3GalNAcβ-4(Neu5Acα-8Neu5Acα-8Neu5Acα3)Galβ-4Glc) can be achieved by using the same combination of gene as in the GD2 sugar production system depicted above and by additionally expressing a β-3 Gal transferase gene. Here, the microorganism further comprises heterologous sequences encoding UDP-GlcNAc 4 epimerase, a β-1,4-GalNActransferase that use the GD3 and GT3 sugar as acceptor, such as the CgtAII protein from *C. jejuni* O:36 strain ATCC Accession No 43456. and a β-1,3-Galactosyltransferase, such as the *cgtB gene* from *C. jejuni* O:2 strain NCTC Accession No 11168.

The system for the production of of GD1b and GT1c can be extended to the production of GT1b, GQ1c, GQ1b and GP1c by expressing a gene encoding a sialyltransferase that is able to use GD1b and GT1c as acceptor.

Production of GD1a and GT1a sugars

The strategy for the production of GT1a sugar is illustrated in figure 6 and relies on the coexpression of the *cstII* gene for the bifunctional sialyltransferase with the *cgtA* gene that encode a β-1,4GalNAc transferase which does not use GD3 sugar as acceptor, the end products being GD3, GD1a and GT1a sugars. Here, the invention relates to a method according as depicted above for producing sialylated oligosaccharides which is applied for producing specifically Neu5Acα-3Galβ-3GalNAcβ-4(Neu5Acα-3)Galβ-4Glc (GD1a) and Neu5Acα-8Neu5Acα-3Galβ-3GalNAcβ-4(Neu5Acα-3)Galβ-4Glc (GT1a), wherein the microorganism comprises heterologous sequences coding for a

bifunctional α-2,3 α-2,8-Sialyltransferase, such as the *cstII* gene from *C. jejuni* strain ATCC Acccession No 43438, a β-1,4-GalNAc transferase, such as the *cgtAII* gene from *C. jejuni* O:36 strain ATCC Accession No 43456, which does not use GD3 sugar as acceptor, and a β-3 Gal transferase, such as the *cgtB gene* from *C. jejuni* O:2 strain NCTC Accession No 11168.

Production of LST_D (sialyl-LNnT) and sialyl-lewis X oligosaccharides

It has already been described that the tetrasaccharide LNnT (Galβ-4GlcNacβ-3Galβ-4Glc) can be produced from exogenous lactose by metabolically engineered strain expressing the *lgtA* and *lgtB* gene encoding β-1,3 GlcNAc transferase and β-1,4Galactosyltransferase respectively (Priem *et al.*, 2002). The above system for producing sialylated oligosaccharides can be extended to the production of the pentasaccharide LST_D (Neu5Acα-3Galβ-4GlcNacβ-3Galβ-4Glc) by expressing the additional gene *nst* and *neuBCA* in a nanK- , nanA- strain as illustrated in figure 7. The system for the synthesis of LST_D can be combined with the fucosylation system that we have described for the production of Lewis X oligosaccharide (Dumon et al., 2004) to produce oligosaccharides carrying the sialyl-lewis X motif (Neu5Acα-3Galβ-4(Fucα-3)GlcNacβ-) at their non reducing end. In this regard, the micro organism further comprises a heterologous sequence coding for a α-1,3-fucosyltransferase such as the *Helicobacter pylori futA* (for example SEQ ID No 18 - from Helicobacter pylori ATCC Accession No 26695) and *futB* (for example SEQ ID No 19 - from Helicobacter pylori ATCC Accession No 26695).

Production of sialosyl galactosyl globoside (SGG) hexasaccharide

The sialosyl galactosyl globoside (Neu5Acα-3Galβ-3GalNAcβ-3Galα-4Galβ-4Gal) has been found to be the preferred binding receptor for uropathogenic *Escherichia coli* (Stapleton *et al.*, 1998) and could potentially be used as an anti-infective agent. The production of the SGG hexasaccharide from globotriose has recently been described using exogenously added sialic acid (Efficient production of globosides sugar using metabolically engineered microorganisms in our US Patent application US 60/711,406). The production of SGG hexasaccharide was carried out by a *nanA- melA-* strain

expressing (i) the *Haemophilus influenzae* (strain rd) lgtD gene that encoded both a GalNAc transferase and a Galactosyltransferase activities, (ii) the *nst* sialyltransferase (α-2,3 sialyltransferase, such as for example from *N. meningitidis*, such the MC58 strain : GenBanK accession number U60660 - SEQ ID No 5, protein_id=AAC44541.1 - SEQ ID No 6) which catalyzes the transfer of a sialyl moiety from an activated sialic acid molecule to globopentaose to form sialosyl galactosyl globoside (SGG) hexasaccharide, (iii) the *neuA* gene for CMP Neu5Ac synthase and (iv) the *wbpP* gene for UDP-GlcNAc epimerase.

This system has been modified to work without sialic acid addition by using a *nanK- nanA- melA-* strain and by additionally expressing the *neuABC* genes as illustrated in figure 8.

In some embodiments, the microorganisms are manipulated to enhance transport of an acceptor saccharide into the cell. Here, where lactose or globotriose is the acceptor saccharide, *E. coli* cells that express or overexpress the LacY permease can be used.

Thus, the invention embraces the above method wherein the microorganism is cultured in a medium with globotriose and is *LacY+*, *MelA-*, *nanT+*, *nanA-*, *nanK-* and comprises heterologous *lgtD*, genes for α-2,3-Sialyltransferase and UGP-GlcNAc C4 epimerase as well as the *neuABC* genes.

The invention also provides a coupling method in which a first microorganism is used to prepare globosides. As mentioned above, the culture medium may include lactose or globotriose but there is no need to supply sialic acid in the configuration herein since it is produced internally. When globotriose is used, the invention also contemplates a set of two separate micoorganisms, the said first microorganism being cultured in a medium with lactose and being *LacY+*, *LacZ-*, *MelA-* and comprising a heterologous *lgtC* gene to produce globotriose (α-1,4-Gal transferase enzyme can be encoded for example by *LgtC* genes of *N. meningitidis*, *N. gonorrhoeae* or *Haemophilus influenzae*, more particularly by the *LgtC* gene of *Neisseria meningititis* L1 (126E) GenBank accession number U65788 - SEQ ID No 7, protein_id AAB48385 - SEQ ID No 8) ; the second microorganism being cultured in a medium with globotriose and being *LacY+*, *MelA-*, *nanT+*, *nanA-*, *nanK-* and comprising heterologous *lgtD*, *wbpP* and *nst* genes as

well as the *neuABC* genes. The gene *lgtD* encodes a β-3 GalNAc transferase to catalyze the transfer of a galactose moiety from UDP-Gal to globotetraose to form globopentaose (*β*-3 Gal transferase activity). For example, the *lgtD* gene from *Haemophilus influenzae* HI1578, GenBanK accession number U32832 - SEQ ID No 9, protein_id=AAC23227 - SEQ ID No 10 can be used.

Production of sialylgalactose

It has recently been shown that *galK* mutant lacking galactokinase ativity can use exogenous galactose as acceptor for the synthesis of olihosaccharides with a terminal reducing galactose (Dumon *et al.*, 2005). The method for producing sialylated oligosaccharides as described above can be advantageously used to produce the disaccharide sialylgalactose (Neu5Acα-3Gal) by a microorganism *galK-, nanA- and nanK-* (or *nanKEAT-*) expressing the gene for sialyltransferase and the *neuBCA* genes cultured in a medium with galactose.

Production of sialylated oligosaccharides with a terminal reducing galactose

The method for the synthesis of sialygalactose can be adapted to the production of analogs of all the sialylated structure mentioned above. The use of galactose as acceptor in place of lactose result in the formation of analogs lacking the terminal glucose residu.

Production of sialylated oligosaccharides with a terminal lactose or galactose carrying latent chemical functions

The broad specificity of the sugar permease can be used to internalize lactose or galactose derivatives carrying latent chemical functions to produce conjugatable oligosaccharides. This strategy has been successfully applied to the synthesis of the oligosaccharide portions of GM2 and GM3 gangliosides with an allyl or a propargyl aglycon (Fort *et al.*, 2005). The alkyne function makes possible an azido addition under aqueous conditions and the alkene function can either be converted into an aldehyde to be linked to proteins by reductive amination, or be transformed into a versatile amino group by the addition of cysteamine. Other chemical function such as azide or amine group can also be used. All these lactose or galactose derivatives can be advantageously used to produce conjugatable analogs of all the sialylated structure mentioned above.

Production of sialylated oligosaccharides with chitooligosaccharide structure at their reducing end

E. coli strains overexpressing the *Azorhizobium caulinodans nodC* gene for chitin-oligosaccharide synthase have been shown to produce more than 2 $g.l^{-1}$ of chitinpentaose when they were cultivated at high cell density (Samain *et al.*, 1997). Once produced in the cytoplasm, chitinpentaose can serve as acceptor for glycosyltransferases that recognize a terminal non-reducing GlcNAc residu. This strategy was used for the synthesis of the hexasaccharide Galβ-4[GlcNAcβ-4]$_4$GlcNAc by an *E. coli* strain that co-expressed the *Azorhizobium caulinodans nodC* gene and the *Neisseria meningitidis lgtB* gene for β-1,4-Galactosyltransferase (Bettler *et al.*, 1999). The terminal *N*-acetyllactosamine motif of this hexasaccharide is an acceptor for sialyltransferase. The sialylated heptasaccharide Neu5Acα-3Galβ-4[GlcNAcβ-4]$_4$GlcNAc can thus be advantageously produced in *nanK, nanA* mutant strains coexpressing *nodC, lgtB, nst* and *neuBCA.* A recently developed strategy to reduce this size is to enzymatically hydrolyze the chitinpentaose by a chitinase in the living bacteria as soon as it is produced by NodC (Cottaz & Samain, 2005). We have found that it is possible within the method of the invention to avoid formation of large size chitinpentaose primer, which considerably increases the molecular weight of the target structures using the chitinase gene from *Bacillus circulans* for example.

As illustrated in figure 9, the method of the invention enable the formation of the sialylated tetrasaccharide Neu5Acα-3Galβ-4GlcNAcβ-4GlcNAc by *nanK-, nanA-* strains coexpressing *nodC, chiA, lgtB, nst* and *neuBCA.* Thus, the invention relates to the above method; or alternatively a method wherein no exogenous is added to the culture medium; for producing sialylated oligosaccharides with chitooligosaccharide structure at their reducing end, such as the sialylated heptasaccharide Neu5Acα-3Galβ-4[GlcNAcβ-4]$_4$GlcNAc, wherein the sialyl-transferase is a α-2,3- Sialyltransferase, such as the Neisseria *nst* gene, and the microorganism further comprises a chitin oligosaccharide synthase such as the *Azorhizobium caulinodans nodC* gene and a β-1,4-Galactosyltransferase gene such as the *Neisseria meningitidis lgtB* gene. It can be extended to the production of sialylated tetrasaccharide Neu5Acα-3Galβ-4GlcNAcβ-

4GlcNAc, wherein the microorganism further comprises a heterologous sequence encoding a chitinase, such as the *chiA* gene.

In still another aspect, the invention is directed to a micoorganism as defined avove as well as to a cell culture medium comprising an exogenous precursor selected from lactose, galactose, β-galactoside, and α-galactoside such as globotriose (Galα–4Galβ–4Glc) and said microorganism.

Example 1: construction of *nanA, nanKA and nanKETA* mutants

All mutants were constructed from strain DC (Dumon *et al.*, 2005) which was a strain DH1 derivative carrying the *lacZ* and *lacA* mutations. Since all derivatives of strain DH1 are *recA* mutant, they were transformed with the low copy plasmid pEXT22 (Dykxhoorn *et al.*, 1996) carrying a functional *recA* gene and a kanamycin resistance to recover a transient $RecA^+$ phenotype for the gene inactivation procedure that involved DNA recombination. Once the gene has been disrupted, the plasmid was cured by growing the cell without kanamycin and screening for $RecA^-$ phenotype.

The strain AZL was constructed from strain DC by inactivating *nanA* using the suicide plasmid pMAK705 (Hamilton *et al.*, 1989) as previously described (Priem *et al.*, 2002).

To construct the strain ZLKA from strain DC, the *nanKETA* genes were disrupted by removing a 3.339 kb segment in the chromosomal DNA using the previously described one-step procedure that employs PCR primers to provide the homology to the targeted sequence (Datsenko & Wanner, 2000). The sequence of the upstream primer was 5'GCAATTATTGATTCGGCGGATGGTTTGCCGATGGTGGTGTAGGCTGGAGCTGCTTC (SEQ ID No 11) and the sequence of the downstream primer was 5'CTCGTCACCCTGCCCGGCGCGCGTGAAAATAGTTTTCGCATATGAATATCCTCCTTAG.((SEQ ID No 12).

The same procedure was used to inactivate the *nanK* gene in strain AZL to obtain the strain AZK except that the size of the deleted fragment was 0.537 kb and that the sequence of the upstream primer was

5'CACTGGCGATTGATATCGGCGGTACTAAACTTGCCGCCGTGTAGGCTGGAGCTGC TTC
(SEQ ID No 13).

Example 2: Cloning of *neuBCA genes*

A 2.995 DNA fragment containing the sequence of the genes *neuBCA* was amplified by PCR using the genomic DNA of *Campylobacter jejuni* strain ATCC 43438 as a template.

A *KpnI* site was added to the left primer:

5'GGTACCTAAGGAGGAAAATAAATGAAAGAAATAAAAATACAA (SEQ ID No 14)

and a *XhoI* site was added to the right primer

5'CTCGAGTTAAGTCTCTAATCGATTGTTTTCCAATG (SEQ ID No 15).

The amplified fragment was first cloned into pCR4Blunt-TOPO vector (Invitrogen) and then sub-cloned into the *KpnI* and *XhoI* sites of pBBR1-MCS3 vector to form pBBR3-SS.

Example 3: Production of sialyllactose by metabolically engineered *E. coli* strains

Sialyllactose production was investigated with different mutant strains contained the *N. meningitidis nst* gene for α-2,3 sialyltranferase and the pBBR3-SS plasmid that contained the *C. jejuni* genes *neuC, neuB* and *neuA* encoding N-acetylglucosamine-6-phosphate-epimerase, sialic acid synthase and CMP-Neu5Ac synthetase respectively. Production of sialyllactose was estimated by the colorimetric quantification of sialic acid in both the intra and extracellular fractions (Table 2). The results showed that the *nanA* mutant AW1 and the DC6 strain, which contained no mutation in the sialic acid operon, produced low amount of sialyllactose with a similar production yield. The two mutants AZK1 and DC7 that carried the *nanK* and *nanA* mutations both produce a four time higher quantities of sialyllactose. No sialic acid could be detected in the control culture of DC7 incubated without lactose, indicating the high level of total sialic acid corresponded to the formation of sialyllactose.

Improvement of sialyllactose production was also confirmed by TLC analysis which showed
that the band corresponding to sialyllactose was much more intense in DC7 and AZK1 extracts that in DC6 and AW1 extracts.

Table 2. Colorimetric quantification of sialic acid in intracellular and extracellular fractions of high cell density cultures of strains genetically engineered for the production of sialyllactose

Strain	mutation	hetrologous genes expressed	accepteur	Neu5Ac concentration (g.l-1) intracellular	extracellular
DC		none	lactose	0	0
DC6		*neuBCA nst*	lactose	0.94	0.43
AW1	*nanA*	*neuBCA nst*	lactose	1.13	0.27
DC7	*nanKETA*	*neuBCA nst*	lactose	2.32	3.25
DC7	*nanKETA*	*neuBCA nst*	none	0.11	0
DC0	*nanKETA*	*neuBCA*	lactose	0	0
AZK1	*nanK nanA*	*neuBCA nst*	lactose	2.16	2.93

Total sialic acid was quantified by the diphenylamine method (Werner & Odin, 1952). Cultures were incubated 30 hours after addition of Lactose was supplied at a concentration of 7.5 g/l except for the control culture of DC7 without lactose.

Example 4: Large scale production of sialyllactose with continuous feeding of lactose

This production yield was increased by extending the cultivation time to 71 hours. Lactose was added at a concentration of 2 $g.l^{-1}$ at the beginning of the fed-batch phases. It was first added continuously with an input rate of 0.52 $g.l^{-1}.h^{-1}$ for 5 hours in the phase with a high glycerol feeding rate. The lactose input rate was then decrease to 0.3 $g.l^{-1}.h^{-1}$ until the end of the culture in the second phase with a low glycerol feeding rate. TLC analysis showed that sialyllactose (compound **2**, **Figure 10**) was continuously produced until the end of the culture and that sialyllactose production was not limited by the supply of lactose (compound **1**) which could always be detected in small amount throughout the culture in the intracellular fraction.

Colorimetric quantification of sialic acid indicated that sialyllactose accumulated mainly in the intracellular fraction in the first part of the culture. The intracellular sialyllactose concentration then plateaued at around 10 $g.l^{-1}$ and the sialyllactose, which was additionally produced, was then secreted in the extracellular medium where it accumulated at a final concentration of 15.5 $g.l^{-1}$ (**Figure 11**).

Example 5: Purification of sialyllactose

Sialyllactose was purified from one liter of DC7 culture obtained as described in example 4. At the end of the culture, the extracellular fraction was separated from the cells by centrifugation. The pH of the extracellular fraction was lowered to 3.00 by the addition of a strong cation exchanger resin (Amberlite IR120 H^+ form). This resulted in the precipitation of proteins which were removed by centrifugation. The pH of the clear supernatant was then adjusted to 6.0 by the addition of a week anion exchanger (Dowex 66 free base form) and half of the supernatant was then loaded on a Dowex 1 (HCO3 form) column (5 x 20 cm). Sialyllactose was retained by Dowex 1 resin and, after washing with distilled water, was eluted with a 0-500mM continuous NaHCO3 gradient. The same procedure was repeated with the other half of the supernatant. Eluted fractions containg sialyllactose were pooled and the NaHCO3 was removed by a treatment with Amberlite IR120 (H^+ form) until pH 3.0. The pH was the adjusted to 6.0 with NaOH and the sialyllactose was freeze-dried.

For the purification of the intracellular fraction, the cells were permeabilized by heating (100°C, 45 min) and resuspended in the same volume as the initial culture medium. Oligosaccharides freely diffused outside of the cells and were recovered in the supernatant after centrifugation. The purification of sialyllactose was then carried out using the same protocol as for the extracellular fraction.

From a one liter culture of strain DC7, the yield of purified sialyllactose was 9 grams from the extracellular fraction and 6 grams from the intracellular fraction. Identification of the purified product as sialyllactose was confirmed by mass spectrometry analysis.

Example 6: Production of GD3 and GT3 sugars

The production of the GD3 and GT3 sugars from exogenous lactose and Neu5Ac has been previously described using metabolically engineered strain expressing the *cstII*

Campylobacter jejuni gene for the bifunctional α-2,3 and α-2,8 sialyltransferase (Antoine *et al.*, 2005). Here we have investigated the production of these two oligosaccharides without exogenous supply of sialic acid by using the system described in example 3. The GD3 producing strain NF3 was a *nanKEAT* mutant which co-expressed *cstII* and *neuBCA* (Table I). The strain NF3 was cultivated at high cell density in presence of 3 $g.l^{-1}$ of lactose. The TLC analysis (**Figure 12**) showed that lactose was entirely converted into three compounds which were presumed to be GM3 (**1**) GD3 (**2**) and GT3 (**3**) sugars.

The intracellular fraction from strain NF03 culture was purified by ion exchange chromatography on Dowex 1 as described in example 4 and three oligosaccharide fractions containing the GM3, GD3, and GT3 sugars respectively were separated. The yields of the three sugar fractions were 0.16 g, 0.75 g and 1.26 g respectively. Identification was confirmed by mass spectrometry analysis of the purified fractions.

Example 7: Production of GM1 sugar using the *N. meningitidis* sialyltransferase and the *P. aeruginosa* UDP-GlcNAc C4 epimerase.

The production of the GM1 sugar from exogenous lactose and Neu5Ac has been previously described using a metabolically engineered strain expressing : (i) the *N. meningitidis nst* gene for α-2,3-Sialyltranferase; (ii) the *cgtA* gene from *C. jejuni* O:19 strain OH4384, which encodes a β-1,4-GalNAc transferase; (iii) the *cgtB* gene from *Campylobacter jejuni* strain NCTC 11168, which encode β-1,4-Galactosyltransferase, (iv) the *wbpP* gene from *P. aeruginosa* which encodes a UDP-GlcNAc C4 epimerase (Antoine *et al.*, 2003). First attempts to produce GM1 sugar without exogenous addition of sialic acid by coexpressing *cgtA*, *cgtB* and *nst* with *neuBCA* indicated that the limiting step was the conversion of sialyllactose into GM2 sugar by the CgtA GalNAc transferase. The GalNAc transferases have been shown to exist in different version depending on the Campylobacter strains. The *cgtA* version cloned from strain OH4384 was reported to have a specific activity largely lower than those of versions from strain ATCC 4356 or NTCC 11168 (Gilbert *et al.*, 2002; Varki, 1993). The *cgtAII* gene was thus cloned by PCR from the genomic DNA of strain ATCC 4356 and subcloned with the *nst* gene into a pBluescript plasmid, yielding to the pBS-cgtAII-nst plasmid (Table I). The *wbpP* gene was cloned from the pBBRwbpP plasmid (Antoine *et al.*, 2003) downstream the

neuBCA gene in pBBR3-SS, yielding to pBBR-SS-wbpP. The *cgtB* was cloned from the pACT3cgtAB plasmid (Antoine *et al.*, 2003) into the pSU27-18 plasmid, yielding to pSU18-cgtB. The DC15 strain was constructed by transforming the *nanKEAT* mutant strain ZLKA with the three plasmids pBS-cgtAII-nst, pBBR-SS-wbpP and pSU18-cgtB (table 1).

As shown in **Figure 13**, the strain DC15 did not accumulate the GM3 sugar (**1**), indicating that the CgtAII from strain ATCC 4356 was considerably more active than CgtA from strain OH4384 . At the end of the culture, the major products were a compound (**5**) which migrated slower than the GM1 sugar and a compound (**2'**) that migrated as the GM2 sugar. After purification on Dowex1, the mass spectrometry analysis indicated that compound (**2'**) was a GM2 sugar analog which has a Gal residu in place of GalNAc and which was further designated as GA2 sugar (Table 3). Structure was : Galβ-4(Neu5Acα-3)Galβ4Glc The mass spectrum of compound (**5**) suggest it was a disialylated octasaccharide formed by the transfert of two sugar residu (one HexNAc and one Hex) on the GD1a sugar. Since these two sugars being most probably added by CgtAII and CgtB, the structure of compound (**5**), which was further designated as GA1 sugar, is likely to be : Galβ-3GalNAcβ-4Neu5Acα-3Galβ-3GalNAcβ-4(Neu5Acα-3)Galβ-4Glc (Table 3).

Table 3 Structure of ganglioside sugar analogs formed as side-products during the syntheseis of ganglioside sugar in reason of the β-1,4Galactosyltransferase activity of the CgtA β-1,4-GalNActransferase.

Structure	name in the text
Galβ-3GalNAcβ-4(Neu5Acα-3)Galβ-3GalNAcβ-4(Neu5Acα-3)Galβ-4Glc	GA1
Galβ-4(Neu5Acα-3)Galβ4Glc	GA2
GalNAcβ-4(Neu5Acα-3)Galβ-4(Neu5Acα-3)Galβ-4Glc	GA3
Galβ-4(Neu5Acα-3)Galβ-4(Neu5Acα-3)Galβ-4Glc	GA4
Neu5Acα-3Galβ-4(Neu5Acα-3)Galβ-4Glc	GA5
Galβ-4(Neu5Acα-8Neu5Acα-3)Galβ-4Glc	GA6

Example 8: Production of GM1 sugar using the *N. meningitidis* sialyltransferase and the Campylobacter UDP-Gal C4 epimerase (GNE).

The GA2 sugar was produced by culture DC15 because the CgtAII is extremely active and can use UDP-Gal as a sugar donor instead of UDP-GalNAc in case of a shortage of UDP-GalNAc. Since this shortage can be due to an insufficient activity of the UDP-GlcNAc C4 epimerase encoded by *wbpP* , the utilisation of the epimerase encoded by the *C. jejuni gne gene* was investigated. This gene has recently been shown to be more active than WbpP (Bernatchez *et al.*, 2005). The *gne* gene was cloned by PCR from the genomic DNA of *C. jejuni* strain NCTC 111168 and subcloned in plasmid pBBR-SS downstream the *neuBCA* genes. The resulting plasmid pBBR-SS-gne was used to construct the strain DC21 which was similar to strain DC 15 except that it expressed *gne* instead of wbpP. (Table 1).

As shown in **Figure 14**, the additional expression of *gne* has almost entirely abolished the accumulation of the isoGM2 sugar and the two major products were the GM1 sugar (**3**) and the GA1 sugar (**5**). The GM1 accumulated transiently; its intracellular concentration reached a maximum 28 hours after the lactose addition and then decreased due to the formation of compound (**5**).

Example 9: Production of GM1 sugar using the *C. jejuni* sialyltransferase CstIII

The *cstIII* gene was cloned by PCR using genomic DNA from *C. jejuni* O:2 strain NCTC 11168 as a template. The *cstIII* gene was then sub-cloned in a pBluescript plasmid upstream *cgtAII*. The resulting plasmid pBS- cstIII-cgtAII was used to construct the strain DC22 which also expressed *cgtB* and the *neuBCA* genes for CMP-Neu5Ac biosynthesis.

TLC analysis (**Figure 15**) showed that the GM1 sugar (**3**) was almost the only oligosaccharide found in the intracellular fraction of strain DC22 at the end of the culture. Sialyllactose (**1**) and the GM2 sugar (**2**) could be barely detected as intermediate. After purification on Dowex1 as described in example 5, the yield of GM1 sugar was 6 g from a one liter of culture.

Example 10: Production of GM2 sugar

The production of the GM2 sugar was investigated with the ZWT strain that was similar to the GM1 producing strain DC21 except it did not expressed the *cgtB* gene. As shown in **figure 16**, lactose was very rapidly converted by strain ZWT in compounds (**2**) that migrated as GM2. Surprisingly a compound (**3**) that migrated slower that GM2 were also produced.

The compounds (**2**) and (**3**) were purified from the intracellular fraction of strain ZWT culture by chromatography on Dowex1. Mass spectrometry analysis indicated that purified compound (**2**) consisted of a mixture of GM2 sugar and of the GA2 sugar analog.

The ESI^- mass spectrum of fraction B showed two peaks at m/z 1310 and 1269.

The peak at m/z 1310 probably corresponds to the quasi-molecular ions $[(M+Na-H)-H]^-$ derived from the hexasaccharide GalNAcβ-4(Neu5Acα-3)Galβ-4(Neu5Acα-3)Galβ-4Glc which was further designated as GA3 sugar (Table 3). The second peak at m/z 1269 corresponds to the quasi-molecular ions $[(M+Na-H)-H]^-$ derived from the Galβ-4(Neu5Acα-3)Galβ-4(Neu5Acα-3)Galβ-4Glc structure (called GA4 sugar).

The formation of GA3 and GA4 sugar is explained as illustrated in figure 11 by a side activity of the Nst sialyltransferase which appears to be able to sialylated the terminal non reducing Gal of the GA2 sugar. The resulting disialylated pentasaccharide (GA6 sugar) can serve as acceptor for the CgtAII to be converted into either GA3 or GA4 sugars

Example 11: Construction of *galU* mutants

To construct the *galU* mutant, a 1.88 kb DNA fragment containing the *galU* sequence was amplified by PCR using the *E. coli* K12 genomic DNA as template and the two following primers: 5'CAATGCCAAATATGGGGAAC (SEQ ID No 16)

and 5'GCGGCCGCGTCTTTTCTGGCTAA (SEQ ID No 17)

The amplified fragment was cloned into the pCR4blunt-TOPO vector (Invitrogen) and a 0.244 kb fragment located between the two *NdeI* sites present in the galU sequence was excised by *NdeI* digestion. The truncated galU gene was subcloned into the *SalI NotI* sites of the pKO3 vector. The *galU* disruption was carried out in strain ZLKAaccording

to the pKO3 gene replacement protocol (Link *et al.*, 1997), yielding to the ZWU host strain.

Example 12: Production of GM2 sugar by *galU* mutants

The strain ZWU2 was similar to the ZWT strain used for GM2 sugar production in example 10, except that the *galU* strain ZWU was used as host strain instead of the ZLKA strain (Table 1) As show in Figure 17, no more side-products (3) was detected at the end of the ZWU2 culture. In addition mass spectrometry analysis showed that compound (**2**) was only composed of GM2 sugar and that the GA2 sugar analog, which was present in the compound **(2)** purified from the ZWT culture, could be detected **(figure 18)**.

Example 13: Production of GD2 sugar.

First attempt to produce the GD2 sugar was made with the strain NF08 which was constructed by transforming the ZLKA host strain with the three plasmids pUC18-cstII, pBBR3-SS-gne, pWKS-cgtAII (Table 1). Characterization of oligosaccharides produced by strain NF08 showed that the major products were the GM2 sugar and its galactosylated analog GA2 sugar. A small amount of GD2 sugar was also formed but it was purified as a mixture with the galactosylated analogs GA5 and GA6 **(figure 5)**.

The *cgtAII* gene was sub-cloned into the pBAD 33 plasmid under the control of the promotor arabinose (P_{ara}). By this way the *cgtAII* expression could be regulated independently from the other genes which were under the control of the P_{lac} promoter. The strain NF09 containing the pBAD33-cgtAII plasmid instead of pWKS-cgtAII was cultivated without arabinose for first 24 hours that followed the lactose addition ($3 g.l^{-1}$). Then arabinose was added and the strain was cultivated for an additional period of 24 hours. This strategy effectively resulted in diminution in the GM2 sugar production with a concomitant increase in the GD2 sugar yield. However the galactosylated analogs GA5 or GA6 were still produced in significant amount. To prevent the formation of these analogs, which are very difficult to

separate from GD2, the production of GD2 was investigated in the galU mutant strain ZWU1.

The strain ZWU1 was similar to the NF09 strain except that the *galU* strain ZWU was used as host strain instead of the ZLKA strain (Table 1) Arabinose was added after 25 hours of culture to increase the expression of cgtAII after that almost all of the sialyllactose has disappeared. Purification of the intracellular fraction on Dowex1 led to the separation of four fractions (**Figure 19**). Mass spectrometry analysis indicated that fraction A was composed of GM2 sugar and that fraction B consisted of GD2 sugar. Fraction C mainly contained GD3 and Fraction D contained GT2 sugars. No galactosylated analogs could be detected and the GD2 sugar was obtained as the major product.

Example 14: Production of GD1b and GT1

The two genes *cstII* and *cgtB* were cloned in the same pBluescript plasmid (pBS-cstII-cgtB). The strain NF21 was constructed by transforming the ZLKA host strain with plasmid pBS-cstII-cgtB and the two plasmids pBBR-SS-gne and pBAD33-cgtAII. Strain NF21 was cultivated at high cell density as described for strain NF09 in example 13 in conditions that favoured the formation of GD2. Oligosaccharides produced from a one liter culture of strain NF21 were first purified on Dowex 1 (HCO_3) resin and three fractions (A,B,and C) were eluted with a $NaHCO_3$ gradient (0-1M).

Mass spectrometry analysis indicated that fraction A (490 mg) contained GD1 sugar, that fraction B (870 mg) contained a mixture of GD1 sugar with GA5 or GA6 analog, and that Fraction C (960 mg) contained GT1 sugar. GD1 and GT1 sugars were purified from Fraction A and C by size exclusion chromatography on a Toyopearl HW40S column using NaHCO3 100mM as eluant. As shown in Table 4, ^{1}H NMR analysis shows that the proton chemical shifts of the H-1 of the terminal galactose of both GD1 and GT1 were very closed to the value of the H-1 of GM1a sugar in which the terminal galactose is unsubstitued. By contrast the signal of the H-1 of the terminal galactose of GD1a and GT1a shifted to 4.62 ppm in reason of its substirution with a sialic acid. These results clearly indicated that the GD1 and GT1 sugars produced by strain NF21 had the structure of GD1b and GT1c .

Table 4. NMR proton chemical shift of gangliosides sugars at 343°K

Strain	Purified sugar	Terminal Galactose		Internal Galactose		GalNAc	Glucose	
		H-1	H-3	H-1	H-3	H-1	α H-1	β H-1
DC22	GM1a	4.57	3.84	4.56	4.17	4.83	5.25	4.68
NF17	GD1a	4.62	4.11	4.54	4.17	4.83	5.25	4.68
NF17	GT1a	4.62	nd	4.53	nd	4.83	5.25	4.68
NF21	GD1b	4.55	nd	4.53	nd	4.83	5.25	4.68
NF21	GT1c	4.55	3.81	4.52	4.17	4.83	5.25	4.68

Example 15: Production of GT1a sugar

The two genes *cstII* and *cgtA* were cloned in the same pBluescript plasmid to be coexpressed with *cgtB* and the neu*ABC* genes in a *nanKEAT* mutant ZLKA host strain (Table 1) . The *cstII* gene was cloned downstream cgtA to have a maximal expression of cgtA and lower expression of *cstII* in order to minimize the formation of GD3. As shown in **figure 20**, TLC analysis of NF17 culture indicated that compounds migrating like sialyllactose (**1**) and GM1 sugar (**3**) were transiently produced. A small amount of compound (**2**) that migrated as GD3 was recovered at the end of the culture but the main final products were compounds (**4**) and (**5**) that migrated slower than GM1. Mass spectrometry analysis of purified compounds indicated that compounds (**4**) and (**5**) have a molecular weight corresponding to GD1a and GT1a respectively.

Example 16: Production of sialylgalactose

The GLKA strain was constructed by deleting the *nanKEAT* genes in the chromosome of the GalK mutant strain GLK (Dumon *et al.*, 2005). The GLK7 strain was obtained by transforming the GLKA strain with the plasmid pBS-nst and pBBR3-SS (Table 1). Cultivation of strain GLK7 at high cell density in presence of 3 $g.l^{-1}$ of galactose resulted in the formation of a disaccharide which was identified to sialylgalactose by mass spectrometry analysis. The sialylgalactose production yield was estimated to be 6 g.l-1 by colorimetric quantification of total sialic acid.

Example 17: Production of the tetrasaccharide Neu5Acα-3Galβ-4GlcNAcβ-4GlcNAc

The plasmid pBS-nst-nodC was constructed by cloning *A.caulinodans nodC* gene from the pBS-nodC plasmid (Cottaz & Samain, 2005) in the *EcoRV KpnI* sites of pBS-nst. The strain SN4 was obtained by transforming the strain ZLKA with the three plasmids pBS-nst-nodC, pBBR3-SS, pWKS-lgtB-chiA (Table 1). Cultivation of strain SN4 at high cell density resulted in the production of a major oligosaccharide which was identified as Neu5Acα-3Galβ-4GlcNAcβ-4GlcNAc by mass spectrometry analysis.

REFERENCES

Antoine, T., Priem, B., Heyraud, A., Greffe, L., Gilbert, M., Wakarchuk, W. W., Lam, J. S. & Samain, E. (2003). Large-scale in vivo synthesis of the carbohydrate moieties of gangliosides GM1 and GM2 by metabolically engineered Escherichia coli. *Chembiochem* **4**, 406-412.

Antoine, T., Heyraud, A., Bosso, C. & Samain, E. (2005). Highly efficient biosynthesis of the oligosaccharide moiety of the GD3 ganglioside by using metabolically engineered Escherichia coli. *Angew Chem Int Ed Engl* **44**, 1350-1352.

Bernatchez, S., Szymanski, C. M., Ishiyama, N., Li, J., Jarrell, H. C., Lau, P. C., Berghuis, A. M., Young, N. M. & Wakarchuk, W. W. (2005). A single bifunctional UDP-GlcNAc/Glc 4-epimerase supports the synthesis of three cell surface glycoconjugates in Campylobacter jejuni. *J Biol Chem* **280**, 4792-4802.

Bettler, E., Samain, E., Chazalet, V., Bosso, C., Heyraud, A., Joziasse, D. H., Wakarchuk, W. W., Imberty, A. & Geremia, A. R. (1999). The living factory: in vivo production of N-acetyllactosamine containing carbohydrates in E. coli. *Glycoconj J* **16**, 205-212.

Cottaz, S. & Samain, E. (2005). Genetic engineering of Escherichia coli for the production of N(I),N(II)-diacetylchitobiose (chitinbiose) and its utilization as a primer for the synthesis of complex carbohydrates. *Metab Eng* **7**, 311-317.

Creuzenet, C., Belanger, M., Wakarchuk, W. W. & Lam, J. S. (2000). Expression, purification, and biochemical characterization of WbpP, a new UDP-GlcNAc C4 epimerase from Pseudomonas aeruginosa serotype O6. *J Biol Chem* **275**, 19060-19067.

Datsenko, K. A. & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-6645.

Dumon, C., Samain, E. & Priem, B. (2004). Assessment of the two Helicobacter pylori alpha-1,3-fucosyltransferase ortholog genes for the large-scale synthesis of LewisX human milk oligosaccharides by metabolically engineered Escherichia coli. *Biotechnol Prog* **20**, 412-419.

Dumon, C., Bosso, C., Utille, J. P., Heyraud, A. & Samain, E. (2005). Production of Lewis x Tetrasaccharides by Metabolically Engineered Escherichia coli. *Chembiochem* **7**, 359-365.

Dykxhoorn, D. M., St Pierre, R. & Linn, T. (1996). A set of compatible tac promoter expression vectors. *Gene* **177**, 133-136.

Fort, S., Birikaki, L., Dubois, M. P., Antoine, T., Samain, E. & Driguez, H. (2005). Biosynthesis of conjugatable saccharidic moieties of GM2 and GM3 gangliosides by engineered E. coli. *Chem Commun (Camb)*, 2558-2560.

Ganguli, S., Zapata, G., Wallis, T., Reid, C., Boulnois, G., Vann, W. F. & Roberts, I. S. (1994). Molecular cloning and analysis of genes for sialic acid synthesis in Neisseria meningitidis group B and purification of the meningococcal CMP-NeuNAc synthetase enzyme. *J Bacteriol* **176**, 4583-4589.

Gilbert, M., Bayer, R., Cunningham, A. M., DeFrees, S., Gao, Y., Watson, D. C., Young, N. M. & Wakarchuk, W. W. (1998). The synthesis of sialylated oligosaccharides using a CMP-Neu5Ac synthetase/sialyltransferase fusion. *Nat Biotechnol* **16**, 769-772.

Gilbert, M., Karwaski, M. F., Bernatchez, S., Young, N. M., Taboada, E., Michniewicz, J., Cunningham, A. M. & Wakarchuk, W. W. (2002). The genetic bases for the variation in the lipo-oligosaccharide of the mucosal pathogen, Campylobacter jejuni. Biosynthesis of sialylated ganglioside mimics in the core oligosaccharide. *J Biol Chem* **277**, 327-337.

Hamilton, C. M., Aldea, M., Washburn, B. K., Babitzke, P. & Kushner, S. R. (1989). New method for generating deletions and gene replacements in Escherichia coli. *J Bacteriol* **171**, 4617-4622.

Kalivoda, K. A., Steenbergen, S. M., Vimr, E. R. & Plumbridge, J. (2003). Regulation of sialic acid catabolism by the DNA binding protein NanR in Escherichia coli. *J Bacteriol* **185**, 4806-4815.

Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M., 2nd & Peterson, K. M. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**, 175-176.

Lee, J., Yi, J., Lee, S., Takahashi, S. & Kim, B. (2004). Production of N-acetylneuraminic acid from N acetylglucosamineand pyruvate using recombinant human renin binding protein and sialic aldolase in one pot. *Enzyme Microb Technol* **35**, 121-125.

Link, A. J., Phillips, D. & Church, G. M. (1997). Methods for generating precise deletions and insertions in the genome of wild-type Escherichia coli: application to open reading frame characterization. *J Bacteriol* **179**, 6228-6237.

Linton, D., Gilbert, M., Hitchen, P. G., Dell, A., Morris, H. R., Wakarchuk, W. W., Gregson, N. A. & Wren, B. W. (2000). Phase variation of a beta-1,3 galactosyltransferase involved in generation of the ganglioside GM1-like lipo-oligosaccharide of Campylobacter jejuni. *Mol Microbiol* **37**, 501-514.

Martinez, E., Bartolome, B. & de la Cruz, F. (1988). pACYC184-derived cloning vectors containing the multiple cloning site and lacZ alpha reporter gene of pUC8/9 and pUC18/19 plasmids. *Gene* **68**, 159-162.

Maru, I., Ohnishi, J., Ohta, Y. & Tsukada, Y. (1998). Simple and large-scale production of N-acetylneuraminic acid from N-acetyl-D-glucosamine and pyruvate using N-acyl-D-glucosamine 2-epimerase and N-acetylneuraminate lyase. *Carbohydr Res* **306**, 575-578.

Plumbridge, J. & Vimr, E. (1999). Convergent pathways for utilization of the amino sugars N-acetylglucosamine, N-acetylmannosamine, and N-acetylneuraminic acid by Escherichia coli. *J Bacteriol* **181**, 47-54.

Priem, B., Gilbert, M., Wakarchuk, W. W., Heyraud, A. & Samain, E. (2002). A new fermentation process allows large-scale production of human milk oligosaccharides by metabolically engineered bacteria. *Glycobiology* **12**, 235-240.

Samain, E., Drouillard, S., Heyraud, A., Driguez, H. & Geremia, R. A. (1997). Gram-scale synthesis of recombinant chitooligosaccharides in Escherichia coli. *Carbohydr Res* **302**, 35-42.

Stapleton, A. E., Stroud, M. R., Hakomori, S. I. & Stamm, W. E. (1998). The globoseries glycosphingolipid sialosyl galactosyl globoside is found in urinary tract tissues and is a preferred binding receptor In vitro for uropathogenic Escherichia coli expressing pap-encoded adhesins. *Infect Immun* **66**, 3856-3861.

Vann, W. F., Tavarez, J. J., Crowley, J., Vimr, E. & Silver, R. P. (1997). Purification and characterization of the Escherichia coli K1 neuB gene product N-acetylneuraminic acid synthetase. *Glycobiology* **7**, 697-701.

Vann, W. F., Daines, D. A., Murkin, A. S., Tanner, M. E., Chaffin, D. O., Rubens, C. E., Vionnet, J. & Silver, R. P. (2004). The NeuC protein of Escherichia coli K1 is a UDP N-acetylglucosamine 2-epimerase. *J Bacteriol* **186**, 706-712.

Varki, A. (1993). Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3**, 97-130.

Vimr, E. R. & Troy, F. A. (1985). Identification of an inducible catabolic system for sialic acids (nan) in Escherichia coli. *J Bacteriol* **164**, 845-853.

Wang, R. F. & Kushner, S. R. (1991). Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in Escherichia coli. *Gene* **100**, 195-199.

Werner, I. & Odin, L. (1952). On the presence of sialic acid in certain glycoproteins and in gangliosides. *Acta Soc Med Ups* **57**, 230-241.

Yamamoto, T., Nakashizuka, M. & Terada, I. (1998). Cloning and expression of a marine bacterial beta-galactoside alpha2,6-sialyltransferase gene from Photobacterium damsela JT0160. *J Biochem (Tokyo)* **123**, 94-100.

Zhang, X. & Kiechle, F. L. (2004). Review: Glycosphingolipids in health and disease. *Ann Clin Lab Sci* **34**, 3-13.

WHAT IS CLAIMED IS:

1. A method for producing oligosaccharides comprising at least one sialic acid residue, herein referred to sialylated oligosaccharides, comprising the step consisting of culturing a microorganism in a culture medium, optionally comprising an exogenous precursor, wherein said microorganism comprises heterologous genes encoding a CMP-Neu5Ac synthetase, a sialic acid synthase, a GlcNAc-6-phosphate 2 epimerase and a sialyltransferase, and wherein the endogenous genes coding for sialic acid aldolase (*NanA*) and for ManNac kinase (*NanK*) have been deleted or inactivated, said microorganism produces internally activated sialic acid.

2. The method according to claim 1, wherein the heterologous sialyltransferase gene is selected from α-2,3 -sialyltransferase, α-2,3- and α-2,8-sialyltransferase (*cstII*), and α- 2,6 sialyltransferase.

3. The method according to claim 1, wherein the heterologous CMP-Neu5Ac synthetase is *neuA*, the heterologous sialic acid synthase is *neuB*, and the heterologous GlcNAc-6-phosphate 2 epimerase is *neuC*.

4. The method according to claim 3, wherein the *neuA*, *neuB*, and the *neuC genes* are isolated from bacterial strains that contain sialylated structures in their cells envelope.

5. The method according claim 1, wherein *nan T* and *nanE* genes are further deleted or inactivated (*nanKEAT*-).

6. The method according claim 1, wherein said microorganism further encodes a protein that facilitates uptake of lactose and lacks enzymes that metabolize lactose.

7. The method according to claim 1, wherein said microorganism is an *E. coli* strain which comprises *LacY* (β-galactoside permease), and wherein *LacZ* (β galactosidase), and optionally *MelA* (α-galactosidase) are further deleted or inactivated.

8. The method according to claim 1, wherein the exogenous precursor is selected from lactose, galactose, β-galactoside, and α-galactoside.

9. A microorganism as defined in any one of claims 1 to 7.

10. A cell culture medium comprising lactose as precursor and the microorganism of claim 9.

11. The method according to claim 1, for the production of 3'sialyllactose or 6'sialyllactose, the microorganism is cultivated at high cell density on carbon substrate, and fed with lactose which is internalized by the lactose permease and sialylated by said recombinant sialyltransferase using the CMP- Neu5Ac endogenously generated from UDP-GlcNAc.

12. The method according to claim 1 for the production of GD3, wherein the heterologous sialyltransferase gene is a bifunctional α-2,3 and α-2,8 sialyltransferase, which catalyzes the transfer of a sialyl moiety from an activated sialic acid molecule produced internally to a Neu5Acα-3Galβ-4Glc (GM3) to form Neu5Acα-8Neu5Acα-3Galβ-4Glc (GD3).

13. The method according to claim 12, which is extended to the production of GT3, wherein the microorganism is further cultured such that the bifunctional sialyltransferase catalyzes the transfer of a sialyl moiety from an activated sialic acid molecule produced internally to the Neu5Acα-8Neu5Acα-3Galβ-4Glc (GD3) to form Neu5Ac5α-8Neu5Acα-8Neu5Acα-3Galβ-4Glc (GT3).

14. The method according to claim 11, which is extended to the production of the carbohydrate portion of the ganglioside Galβ-3GalNAcβ-4(Neu5Acα-3)Galβ-4Glc (GMl), wherein the microorganism further comprises heterologous sequences encoding β-l,4GalNActransferase and β-l,3-Galactosyltransferase, which β-1,4GalNActransferase transfers a UDP-GalNac residue to sialyllactose (GM3) to form GalNAcβ-4(Neu5Acα-3)Galβ-4Glc (GM2) and the β-l,3-Galactosyltransferase transfers a Galactosyl residue to GM2 to form GMl.

15. The method according to claim 14, wherein the microorganism further comprises a heterologous sequence encoding a UDP-GlcNAc 4 epimerase.

16. The method according to claim 11, which is extended to the production of Neu5Acα-3Galβ-3GalNAcβ-4(Neu5Acα-3)Galβ-4Glc (GD1a) and Galβ-3GalNAcβ- 4(Neu5Acα-3)Galβ-3GalNAcβ-4(Neu5Acα-3)Galβ-4Glc (GAl).

17. The method according to claim 14, for the specific production of GMl, wherein the heterologous sialyl transferase is a α-2,3 sialyltransferase encoded by the *cstIII* gene isolated from *C. jejuni* strains expressing lipooligosaccharide structures that mimic the GMl ganglioside.

18. The method according to claim 11, which is extended to the production of the carbohydrate portion of a ganglioside GalNAcβ-4(Neu5Acα-3)Galβ-4Glc (GM2), wherein the microorganism further comprises heterologous sequences encoding a β-1,4- GalNActransferase, and a UDP-GIcNAc 4 epimerase and wherein the micoorganism has at least one of the three following genes further deleted or inactivated to disrupt the endogenous production of UDP-Gal: *galE* gene encoding the UDP-glucose epimerase, *galU* gene which encodes the UDP-Glc pyrophosphorylase, and *pgm* gene which encodes the phosphoglucomutase.

19. The method according to claim 11, which is extended to the production of the carbohydrate portion of the ganglioside of GD2

GalNAcβ-4Galβ-4Glc
|
Neu5Acα-8Neu5Acα3

wherein the microorganism further comprises heterologous sequences encoding UDP-GIcNAc 4 epimerase and a β-l,4-GalNActransferasc that use a GD3 sugar as acceptor.

20. The method according to claim 19, which is extended to the production of GD1b sugar (Galβ-3GalNAcβ-4(Neu5Acα-8Neu5Acα-3)Galβ-4Glc) and GT1c sugar (Galβ-3GalNAcβ-4(Neu5Acα-8Neu5Acα-8Neu5Acα3)Galβ-4Glc) wherein the microorganism further comprises an heterologous sequence encoding a β-3 Gal transferase gene.

21. The method according to claim 20, which is extended to the production of GT1b (Neu5Acα-3Galβ-3GalNAcβ-4(Neu5Acα-8Neu5Acα3)Galβ-4Glc),GQ1c

(Neu5Acα-3Galβ-3GalNAcβ-4(Neu5Acα-8Neu5Acα-8Neu5Acα3)Galβ-4Glc), GQ1b (Neu5Acα-8Neu5Acα-3Galβ-3GalNAcβ-4(Neu5Acα-8Neu5Acα3)Galβ-4Glc) and GP1c (Neu5Acα-8Neu5Acα-3Galβ-3GalNAcβ-4(Neu5Acα-8Neu5Acα-8Neu5Acα3)Galβ-4Glc) wherein the microorganism further comprises an heterologous sequence encoding a gene encoding a sialyltransferase using GDIb and GTIc as acceptor.

22. The method according to claim 19, wherein the micoorganism has at least one of the three following genes further deleted or inactivated to disrupt the endogenous production of UDP-GaI: *galE* gene encoding the UDP-glucose epimerase, *galU* gene which encodes UDP-Glc pyrophosphorylase, and *pgm* gene.

23. The method according to claim 11, which is extended to the production of a carbohydrate portion of the ganglioside of GT1b:

Neu5Acα3 Galβ-3GalNAcβ-4Galβ-4Glc

Neu5Acα-8Neu5Acα3

wherein the microorganism further comprises heterologous sequences encoding UDP-GlcNAc 4 epimerase, a β-l,4-GalNActransferase that used the GD3 sugar as acceptor and a β-l,3-Galactosyltransferase.

24. The method according to claim 1 for producing Neu5Acα-3Galβ- 3GalNAcβ-4(Neu5Acα-3)Galβ-4Glc (GD1a) and Neu5Acα-8Neu5Acα-3Galβ- 3GalNAcβ-4(Neu5Acα-3)Galβ-4Glc (GT1a), wherein the microorganism comprises heterologous sequences coding for a bifunctional α-2,3 and α-2,8-Sialyltransferase, a β-l,4GalNAc transferase, which does not use GD3 sugar as acceptor, and a β-l,3-Galactosyltransferase.

25. The method according to claim 1 for producing a pentasaccharide LST_D (Neu5Acα-3Galβ-4GlcNacβ-3Galβ-4Glc), wherein the sialyl transferase is a α-2,3- Sialyltransferase, and wherein the microorganism further comprises heterologous *lgtA* and *lgtB* genes encoding β-l,3-GlcNAc transferase and β-l,4-Galactosyltransferase respectively.

26. The method according to claim 23 for producing sialyl-lewis X oligosaccharides, wherein the microorganism further comprises a heterologous sequence coding for a α-1,3-fucosyl-transferase.

27. The method according to claim 1 for producing sialosyl galactosyl globoside (Neu5Acα-3Galβ-3GalNAcβ-3Galα-4Galβ-4Gal), wherein the microorganism is *LacY*+ and cultured in a medium with globotriose, and wherein *MelA*, *nanT* are further deleted or inactivated, and further comprises heterologous *lgtD*, genes for α-2,3- Sialyltransferase and UGP-GlcNAc C4 epimerase as well as *neuABC* genes.

28. The method according to claim 1 for producing sialylgalactose (Neu5Acα-3Gal) and sialylated oligosaccharides with a terminal reducing galactose, wherein the microorganism is further inactivated or deleted in *galK* and expresses a gene for sialyltransferase and *neuBCA* genes, and is cultured in a medium with galactose.

29. The method according to claim 1 for producing sialylated oligosaccharides with chitooligosaccharide structure at their reducing end, wherein the sialyltransferase is a α-2,3- Sialyltransferase and the microorganism further comprises a chitin oligosaccharide synthase and a β-l,4-Galactosyltransferase gene.

30. The method according to claim 29 for producing sialylated tetrasaccharide Neu5Acα-3Galβ-4GlcNAcβ-4GlcNAc, wherein the microorganism further comprises a heterologous sequence encoding a chitinase.

31. The method according to claim 29, wherein no exogenous precursor is added to the culture medium.

32. The method according to any one of claims 11 to 31 wherein the microorganisms are grown on glycerol or glucose as carbon substrate.

33. The method of claim 4, wherein said bacterial strain is *C. jejuni* strain ATCC Accession No. 43438.

34. The method of claim 8, wherein said exogenous precursor is globotriose (Galα-4Galβ-4Glc).

35. The method of claim 11, wherein said carbon substrate is glucose or glycerol.

36. The method of claim 12, wherein said heterologous sequence encoding for a bi-functional α-2,3 and α-2.8 sialyltransferase is cstII gene from *Campylobacter jejuni* deposited under ATCC Accession No. 43438.

37. The method of claim 15, wherein said heterologous sequence is *wbpP* gene from *Pseudomonas aeruginosa* or the *gne* gene from *C. jejuni.*

38. The method of claim 17, wherein *C. jenuni* strain is *C. jejuni* strain NCTC Accession No 111168.

39. The method of claim 18, wherein the heterologous sequence is *CgtAII* gene from *C. jejuni* O:36 strain ATCC Accession No 43456.

40. The method of claim 19, wherein the heterologous sequence is CgtAII protein from *C. jejuni* O:36 strain ATCC Accession No 43456.

41. The method of claim 20, wherein the heterologous sequence is *cgtB gene* from *C. jejuni* O:2 strain NCTC Accession No 11168.

42. The method of claim 23, wherein the heterologous sequence encoding UDP-GlcNAc 4 epimerase, a β-l,4-GalNActransferase that use the GD3 sugar as acceptor is CgtAII protein from *C. jejuni* O: 36 strain ATCC Accession No 43456, and said heterologous sequence encoding a 1,3-Galactosyltransferase is *cgtB gene* from *C. jejuni* O:2 strain NCTC Accession No 11168.

43. The method of claim 24, wherein said heterologous sequences encoding for 1) said bi-functional α-2,3 and α-2,8 sialyltransferase is *cstII* gene from *C. jejuni* strain ATCC Acccession No 43438, 2) said β-1,4GalNAc transferase is *cgtAII* gene from *C. jejuni* O:36 strain ATCC Accession No 43456, and 3) for said β1,3-Galactosyltransferase is *cgtB gene* from *C. jejuni* O:2 strain NCTC Accession No 111168.

44. The method of claim 25, wherein said α-2,3 sialyltransferase is encoded by the *nst* gene.

45. The method of claim 26, wherein said heterologous sequence is SEQ ID No 18 (*H. pylori futA*) or SEQ ID No 19 (*fut*B).

46. The method of claim 27, wherein said heterologous *lgtD* gene is the *lgtD* gene from *Haemophilus influenzae* HIl 578, GenBanK accession number U32832.

47. The method of claim 29, wherein said sialylated oligosaccharides with chitooligosaccharide structure at their reducing end is sialylated heptasaccharide Neu5Acα-3Galβ-4[GlcNAcβ-4]$_4$GlcNAc, said α-2,3- Sialyltransferase is encoded by the Neisseria *nst* gene, said chitin oligosaccharide synthase is *Azorhizobium caulinodans nodC* gene, and wherein said β-l,4-Galactosyltransferase gene is *Neisseria meningitidis lgtB* gene.

48. The method of claim 30, wherein said heterologous sequence encoding a chitinase is *chiA* gene.

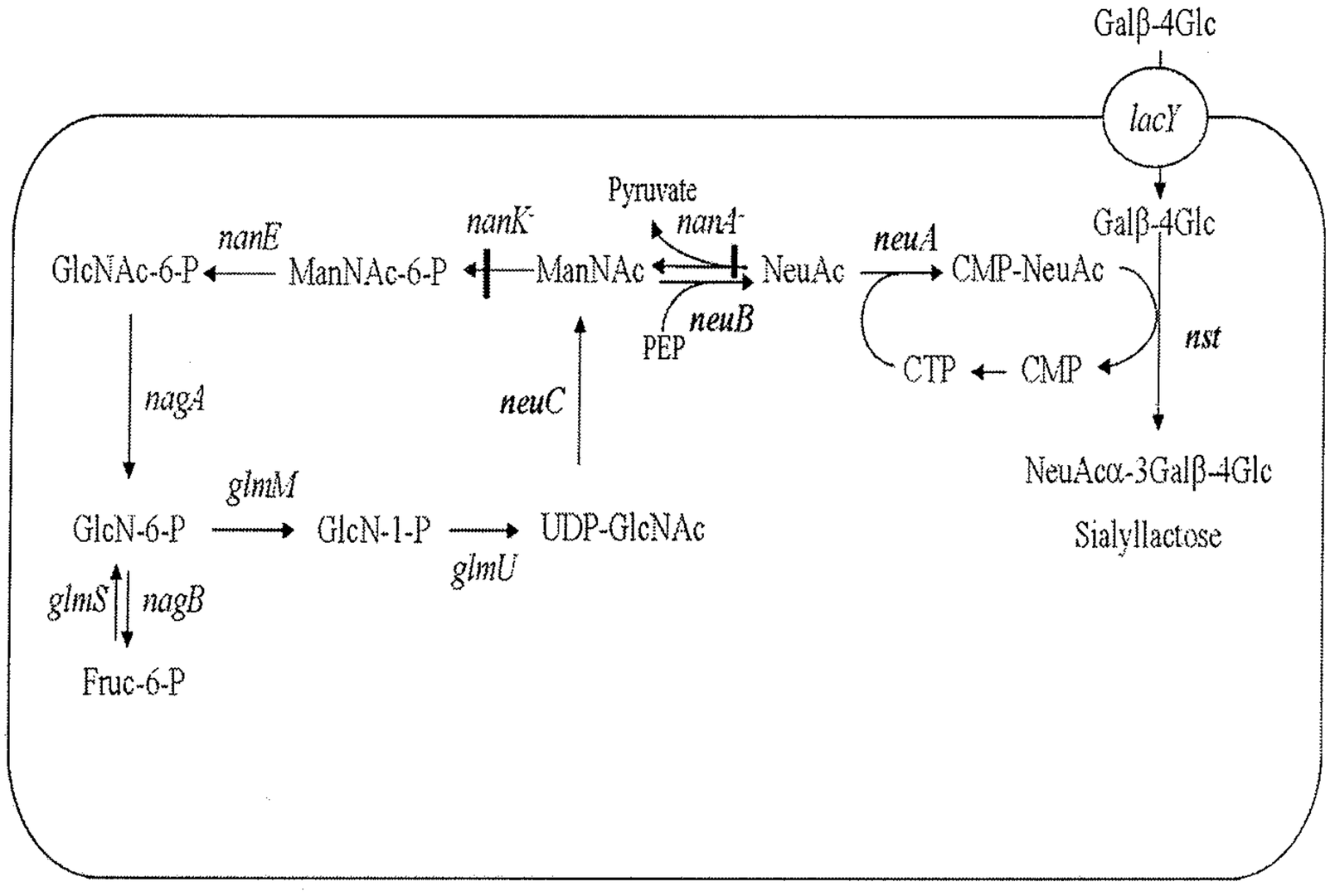


FIGURE 1

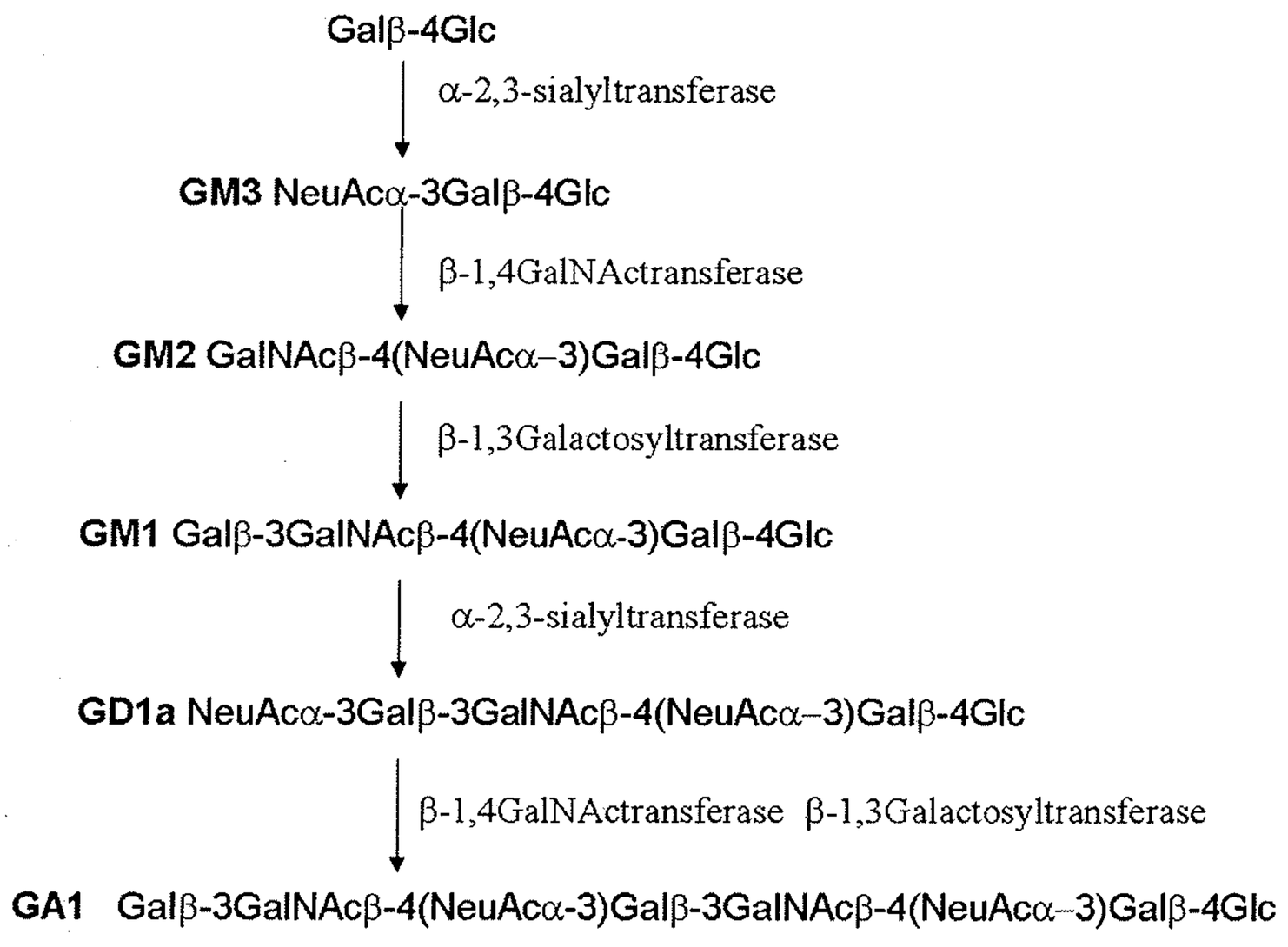


FIGURE 2

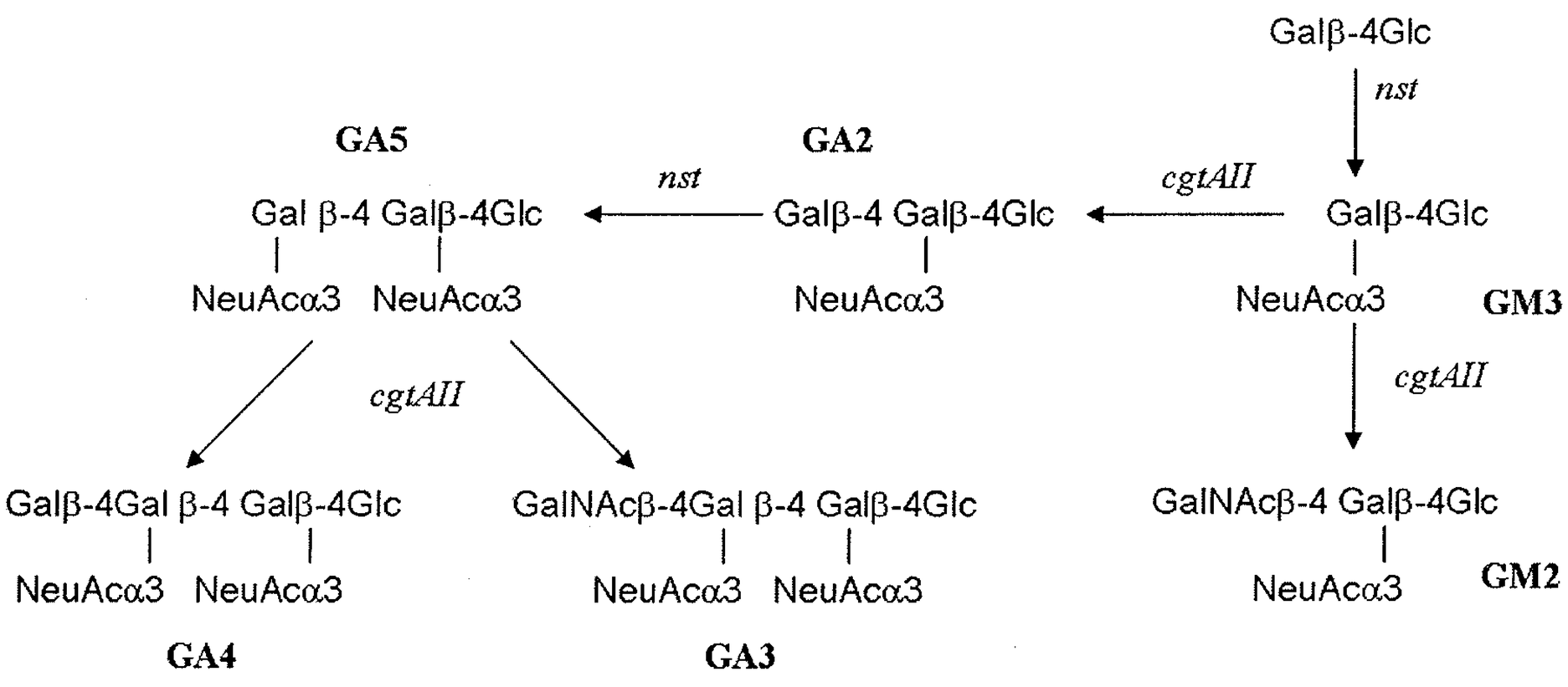


FIGURE 3

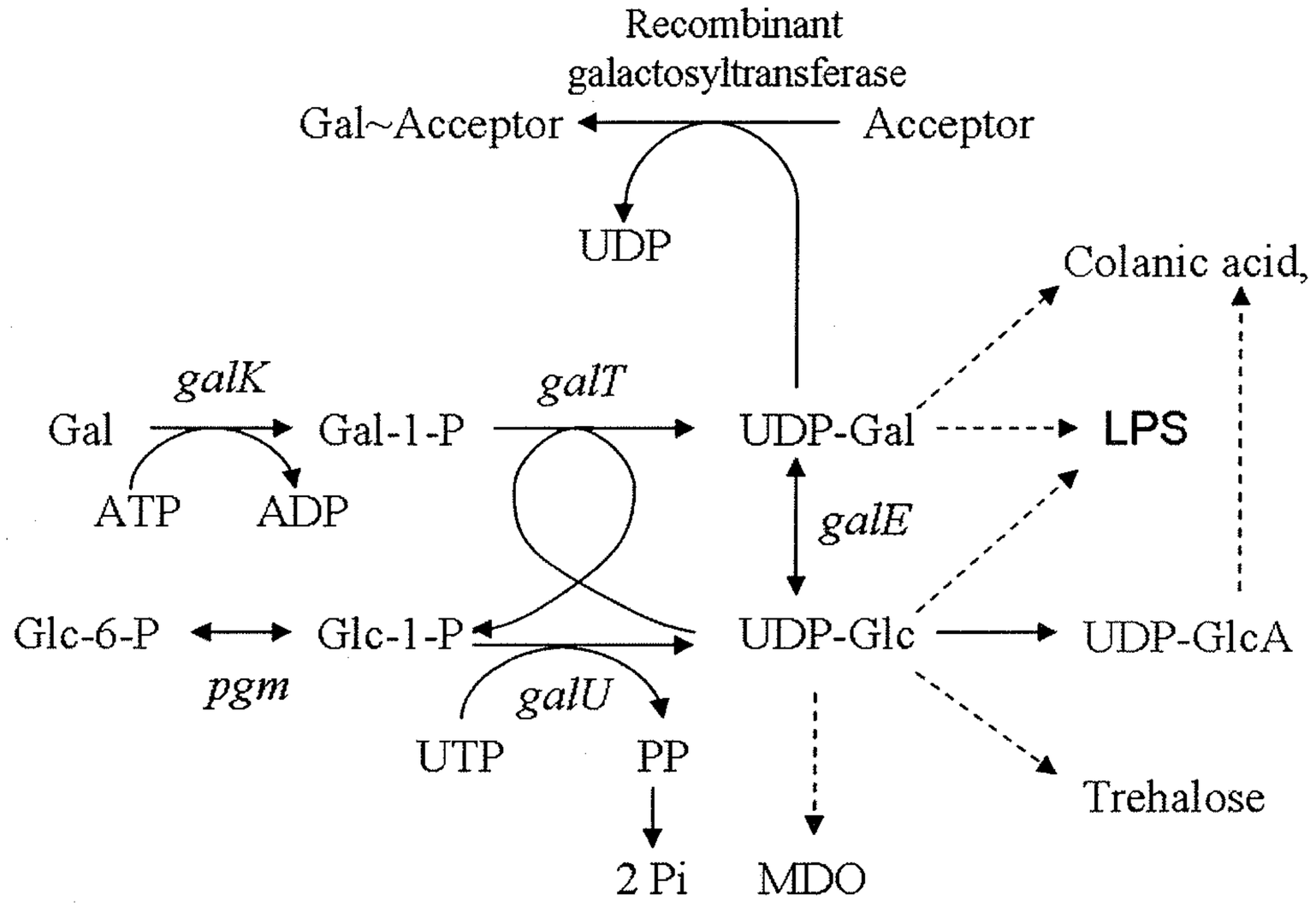


FIGURE 4

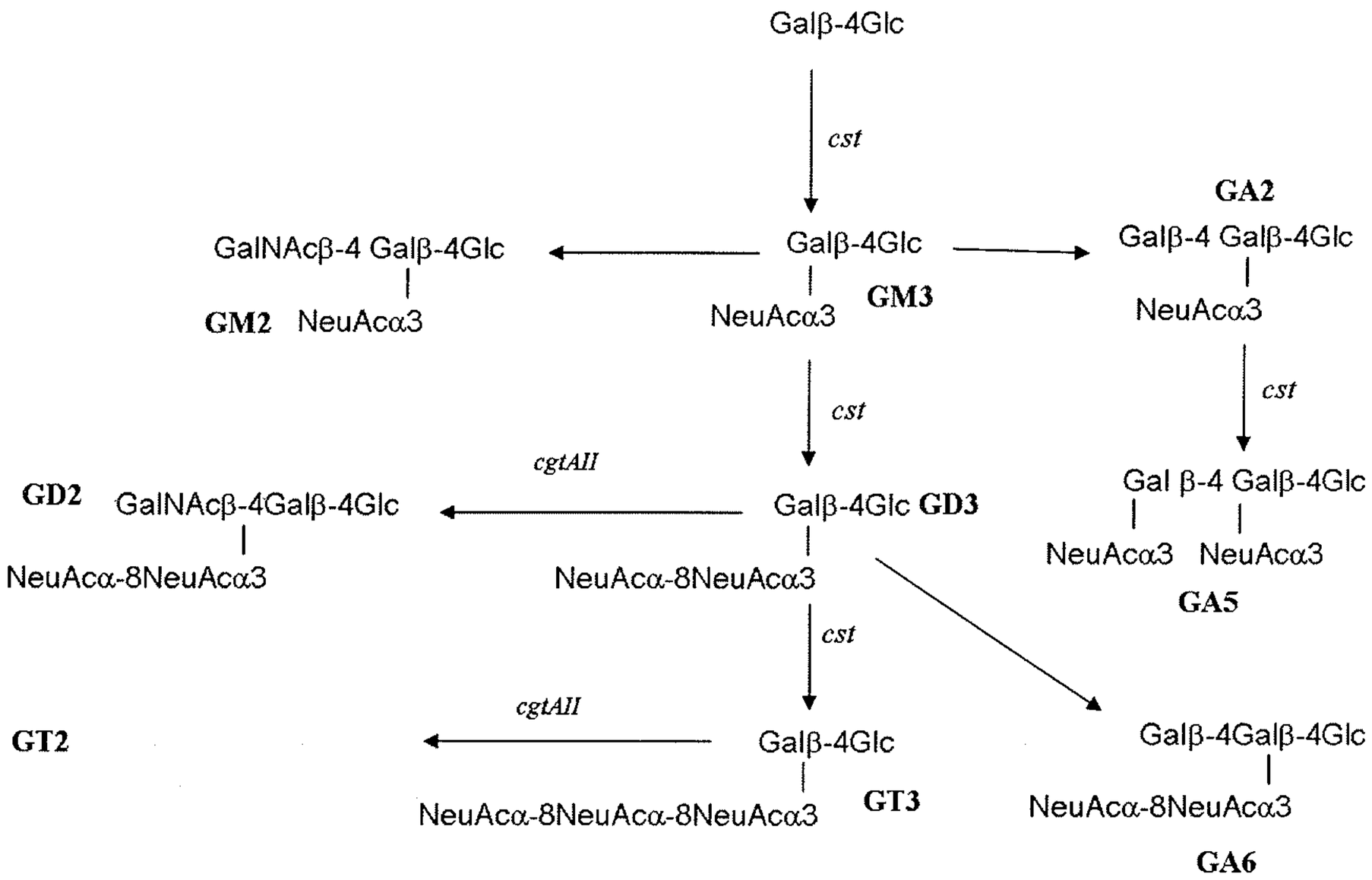


FIGURE 5

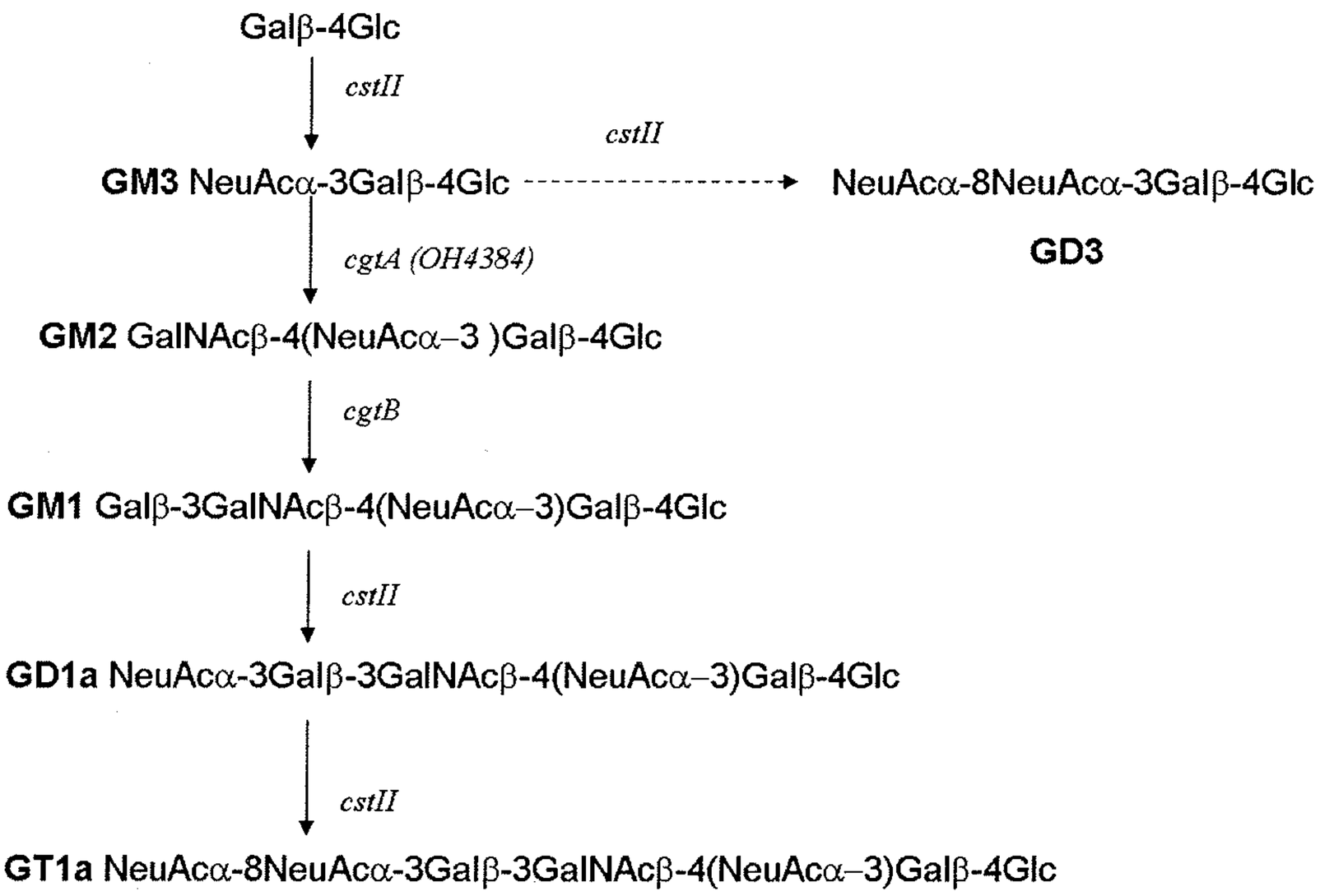


FIGURE 6

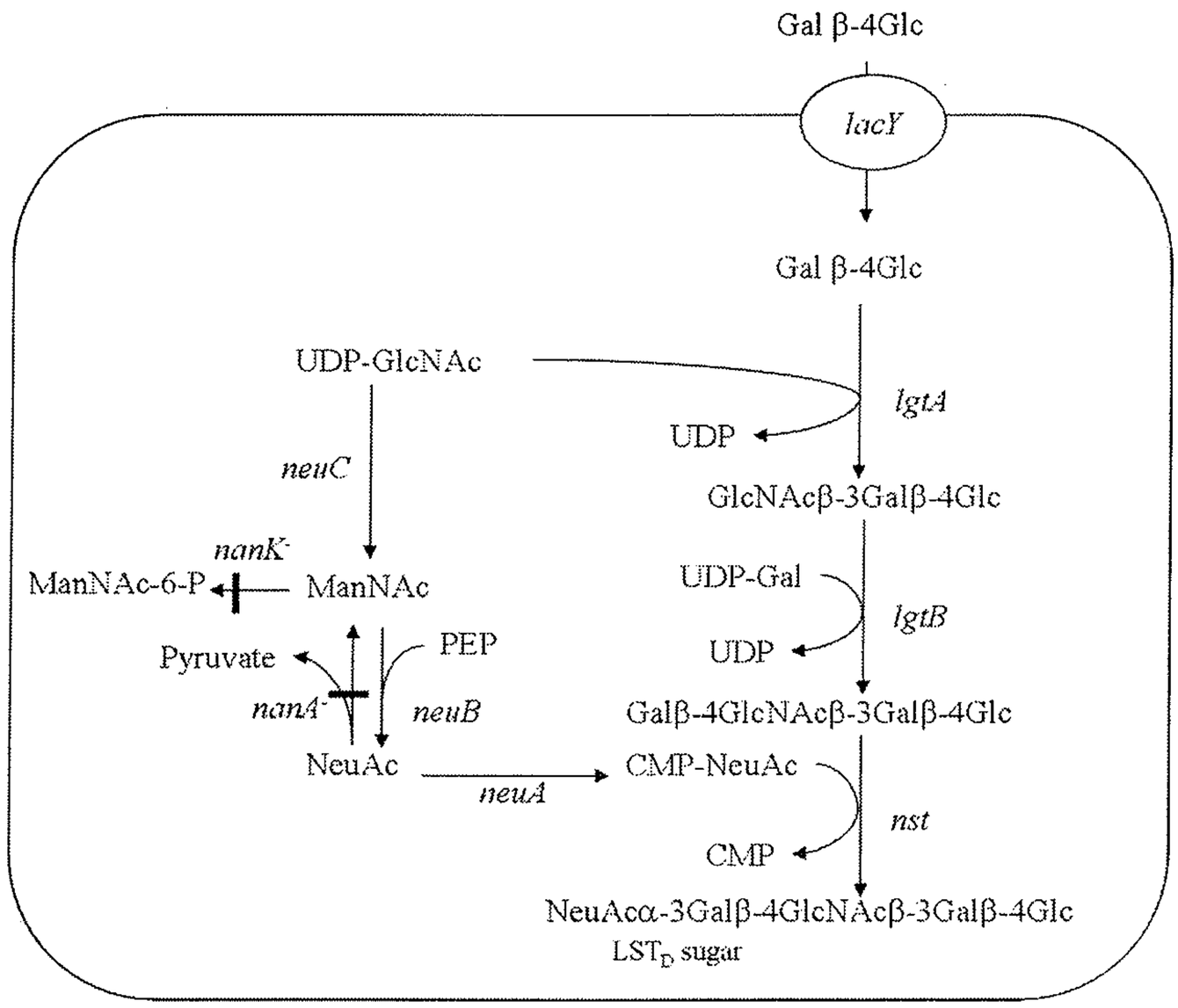


FIGURE 7

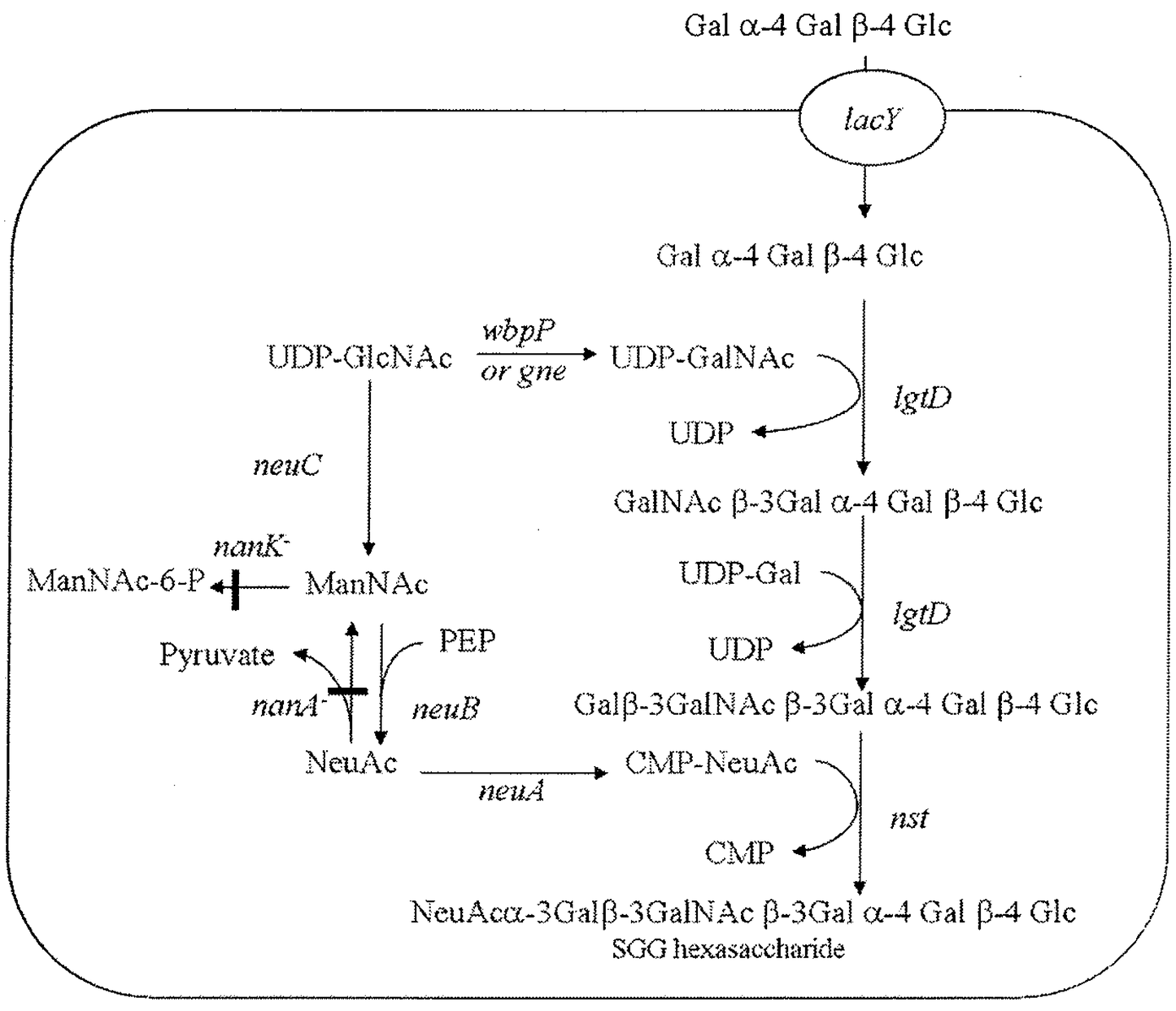


FIGURE 8

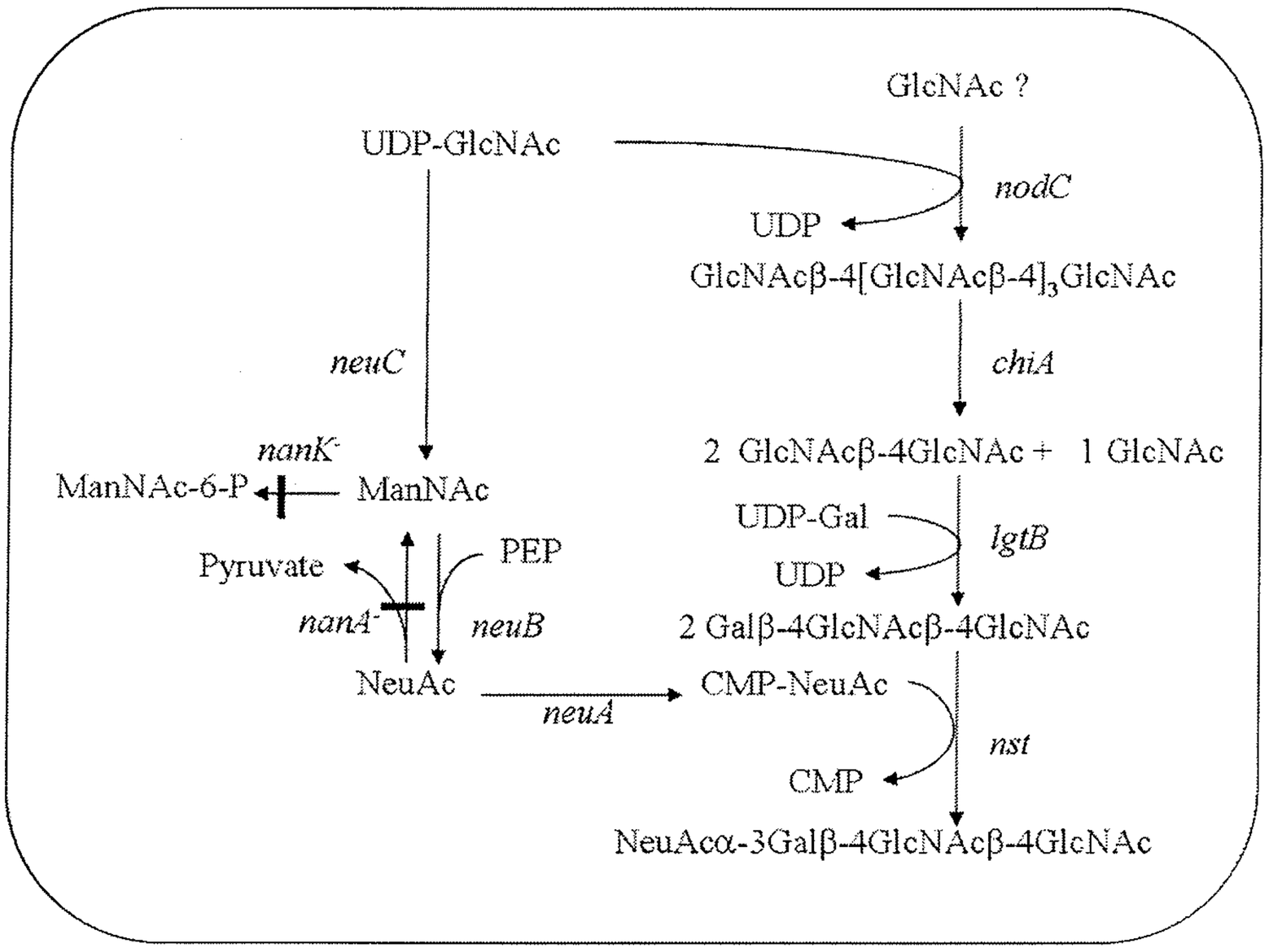


FIGURE 9

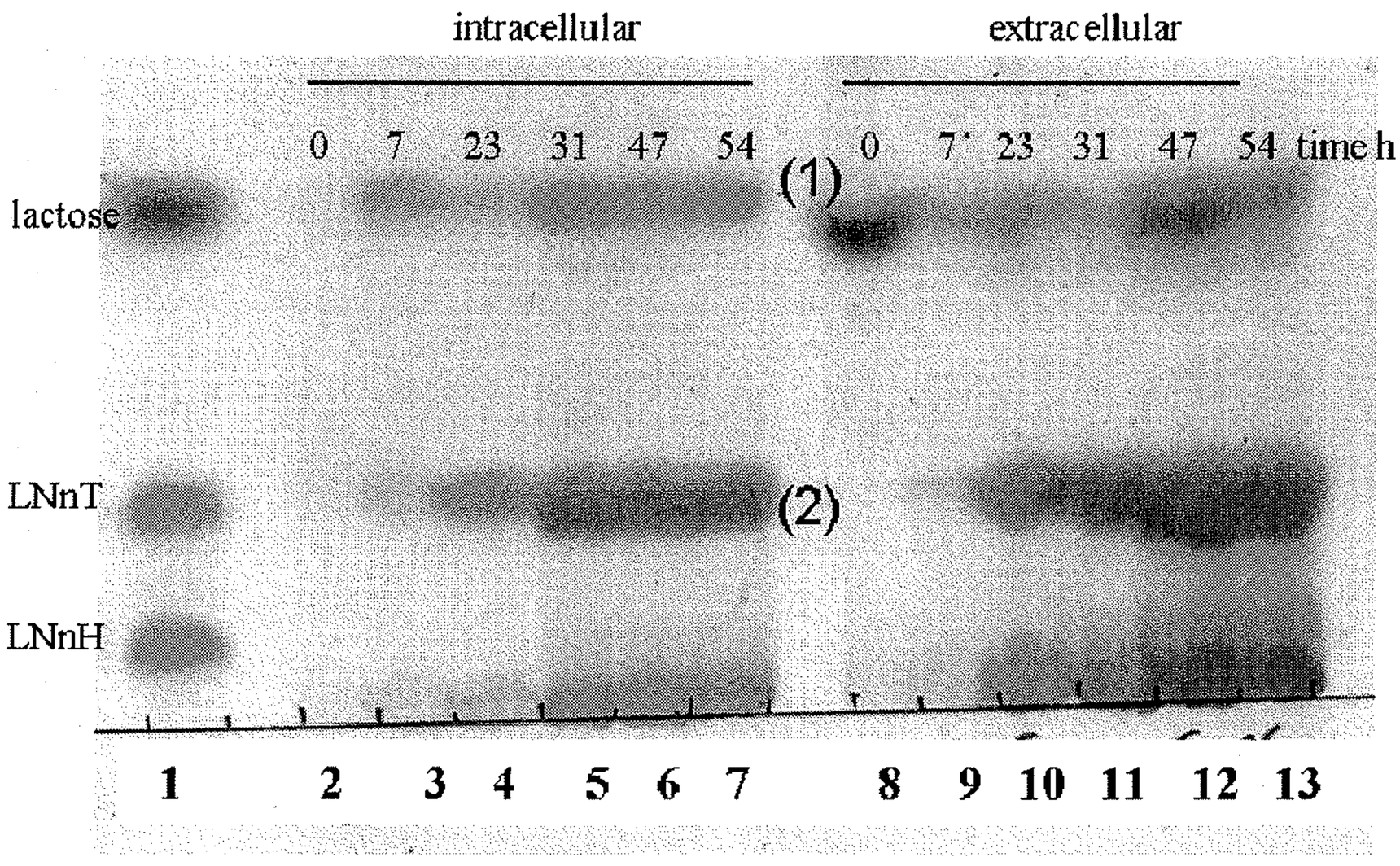


FIGURE 10

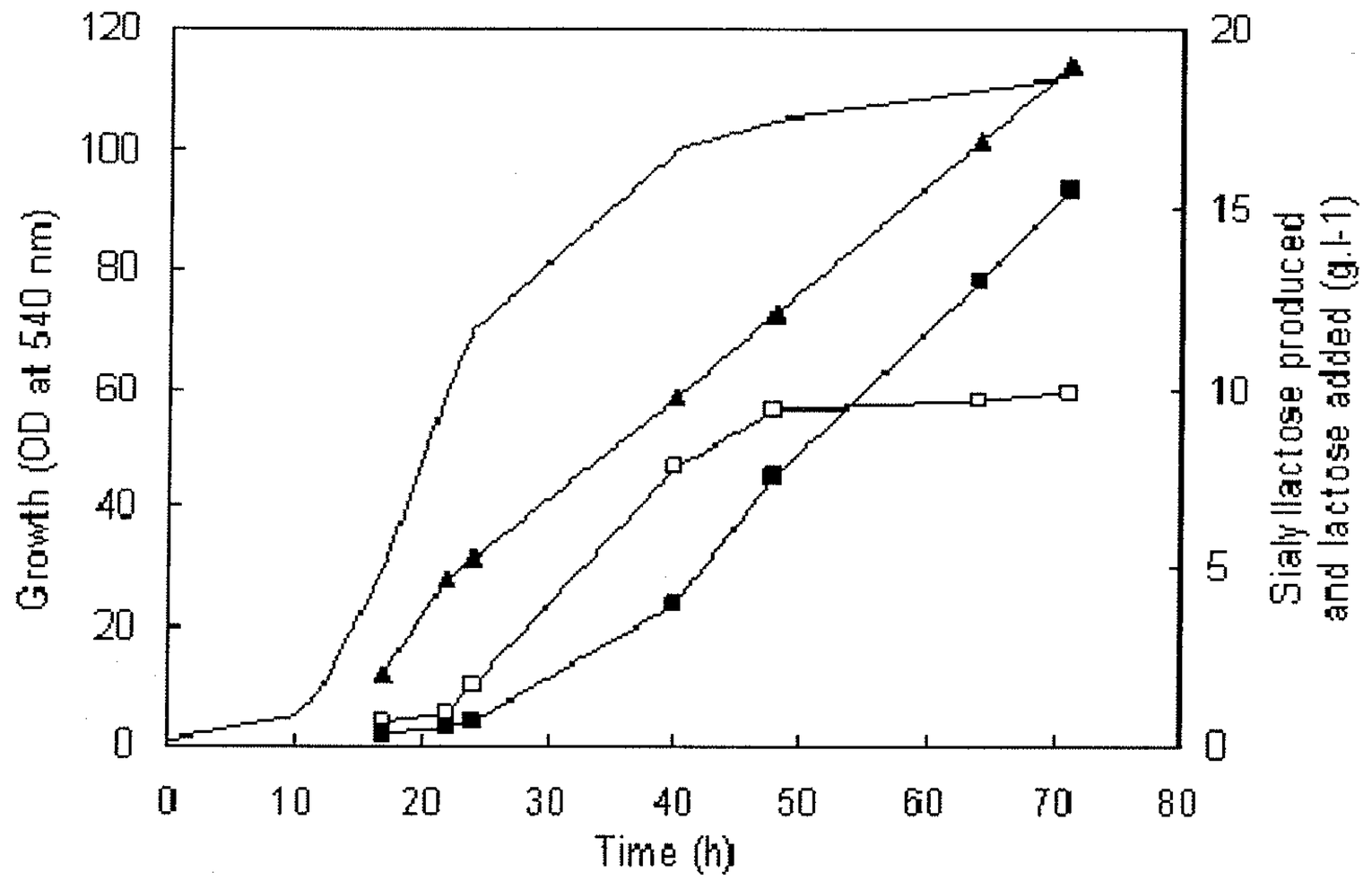


FIGURE 11

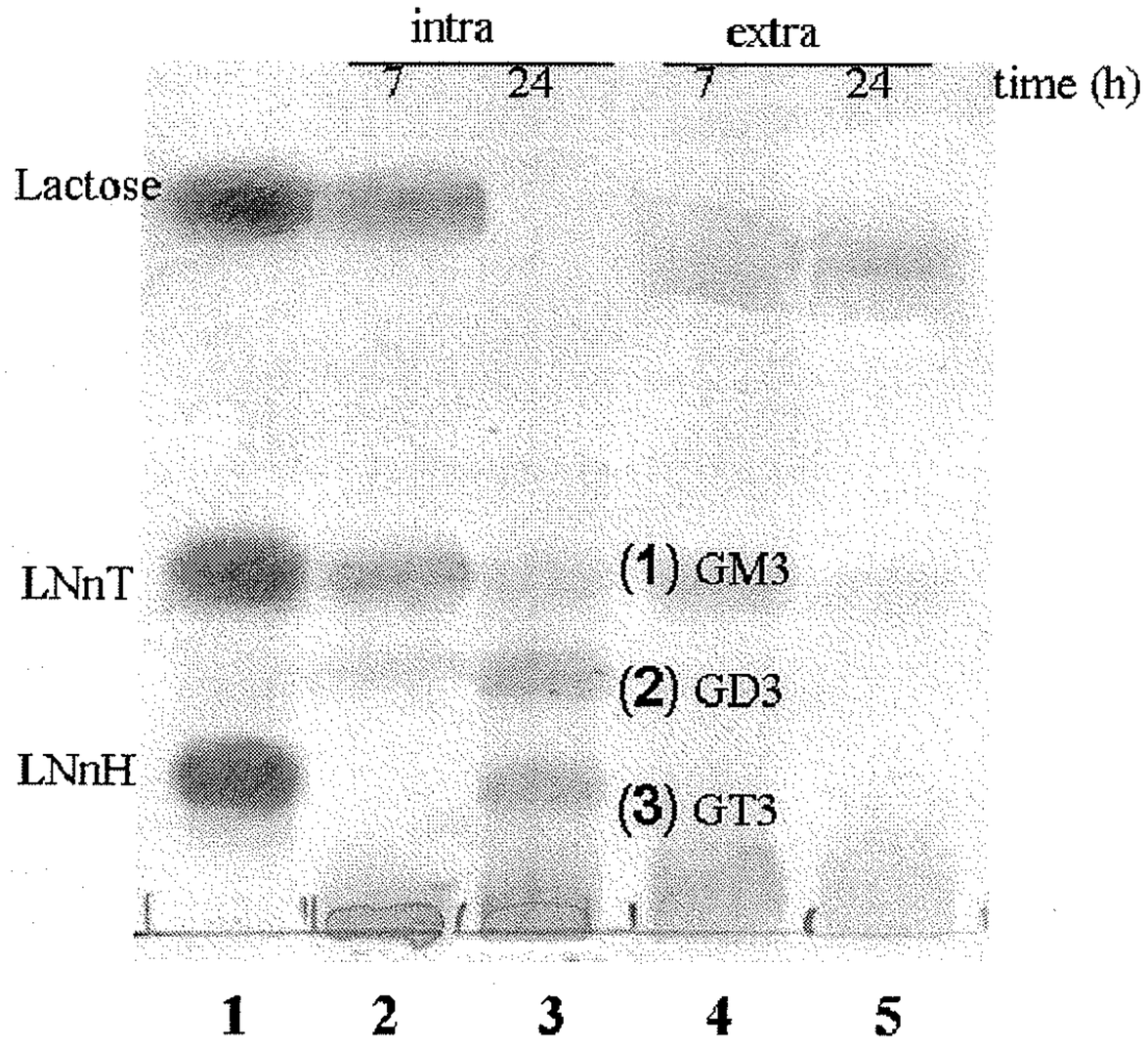


FIGURE 12

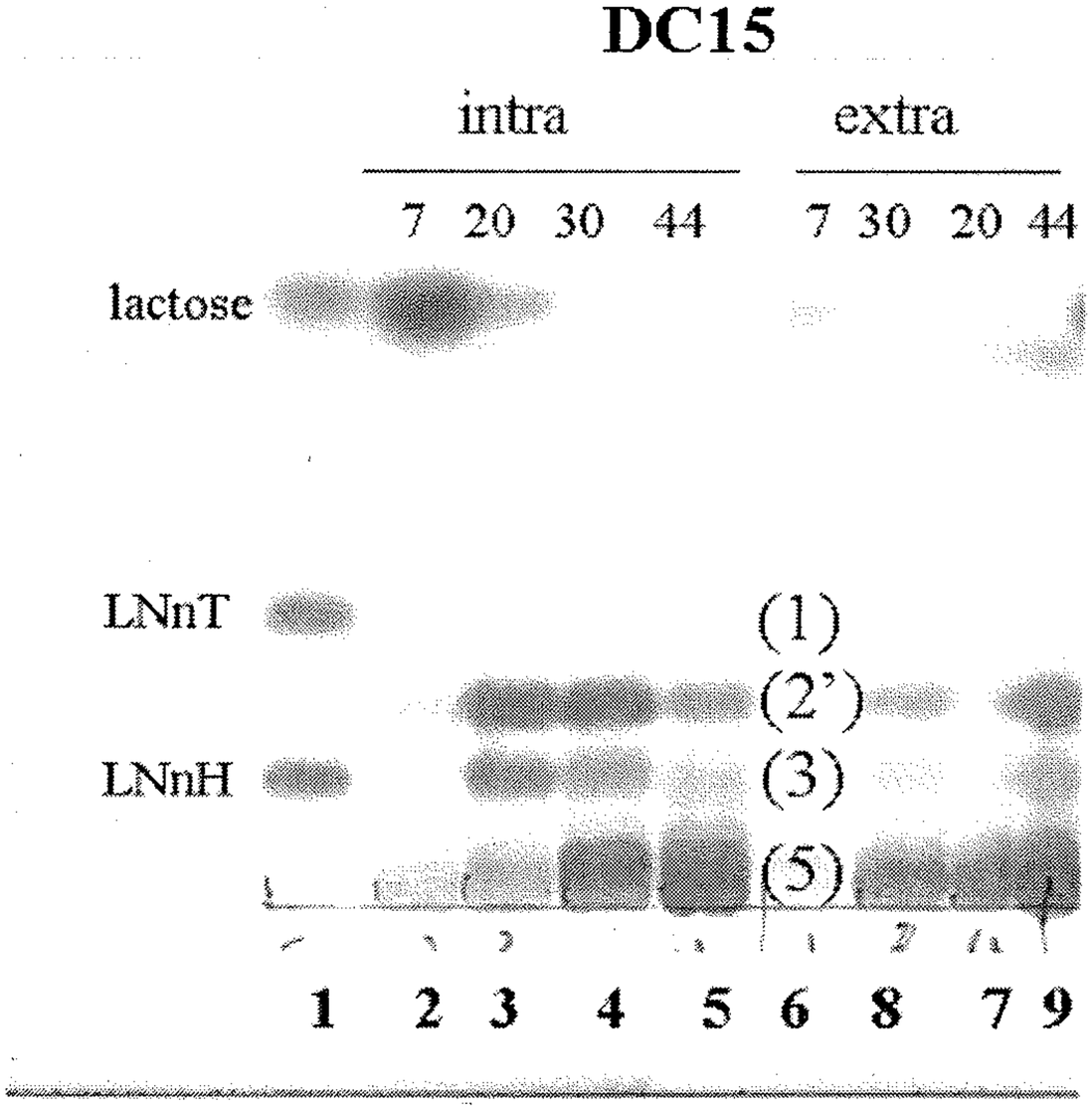


FIGURE 13

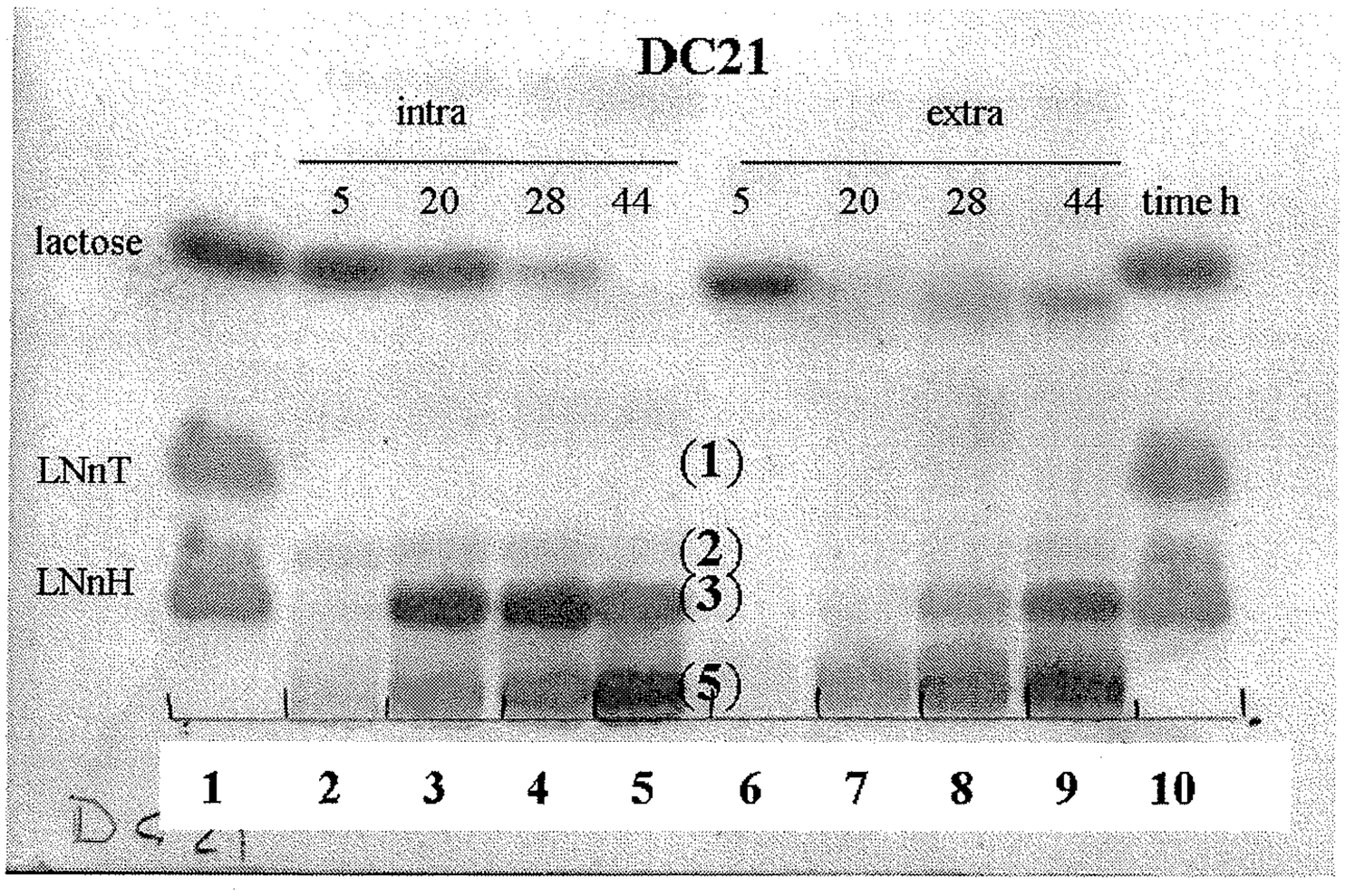


FIGURE 14

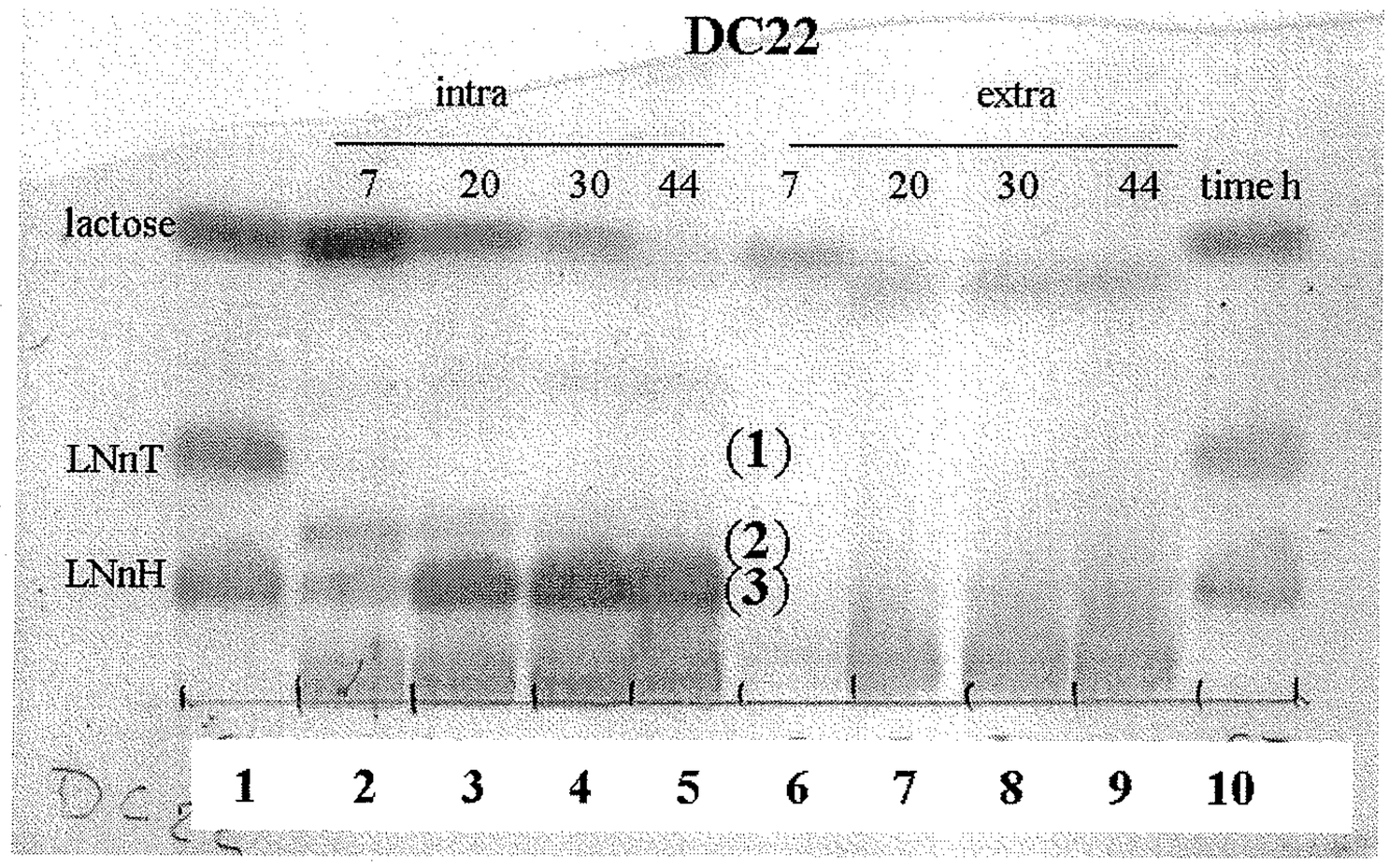


FIGURE 15

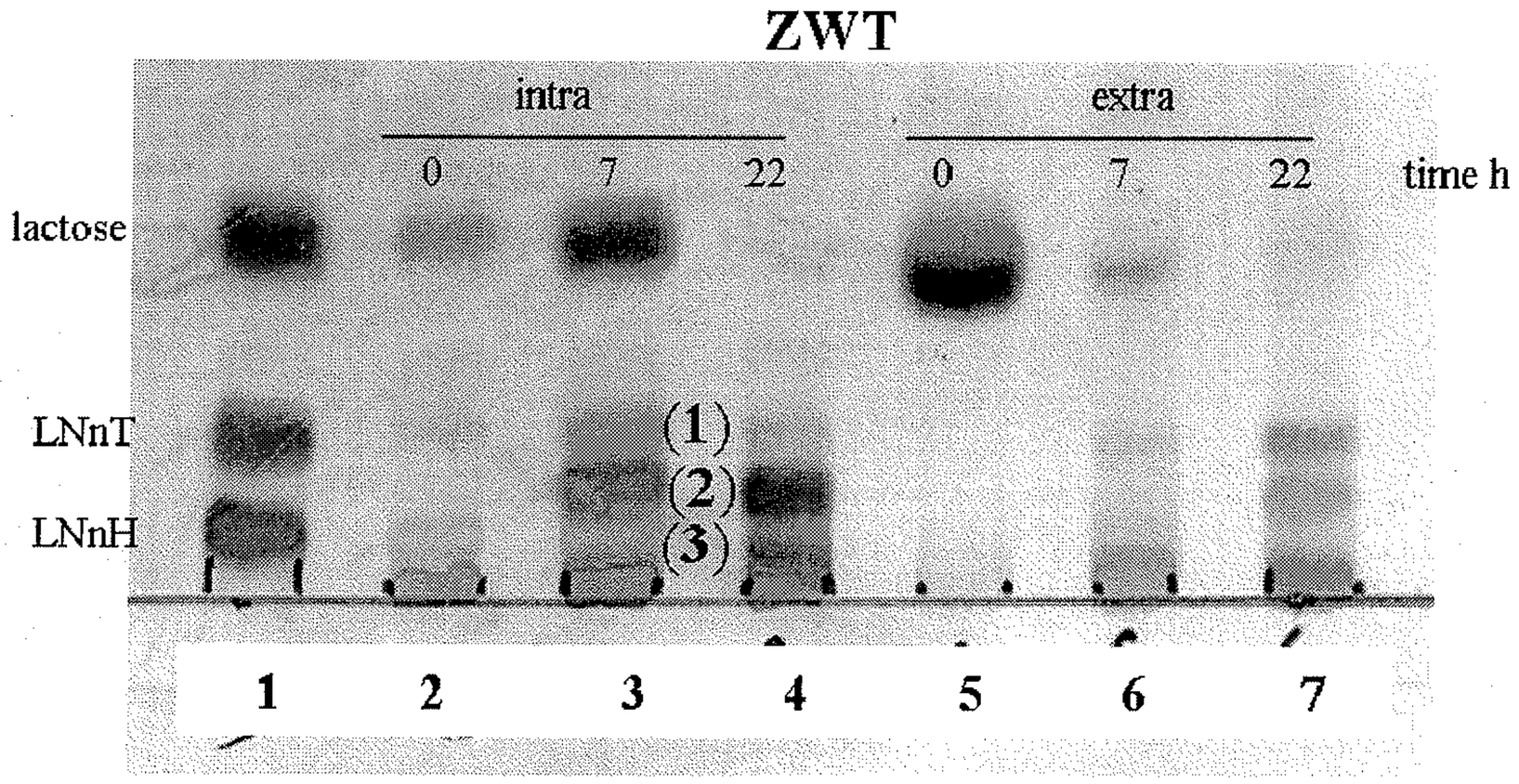


FIGURE 16

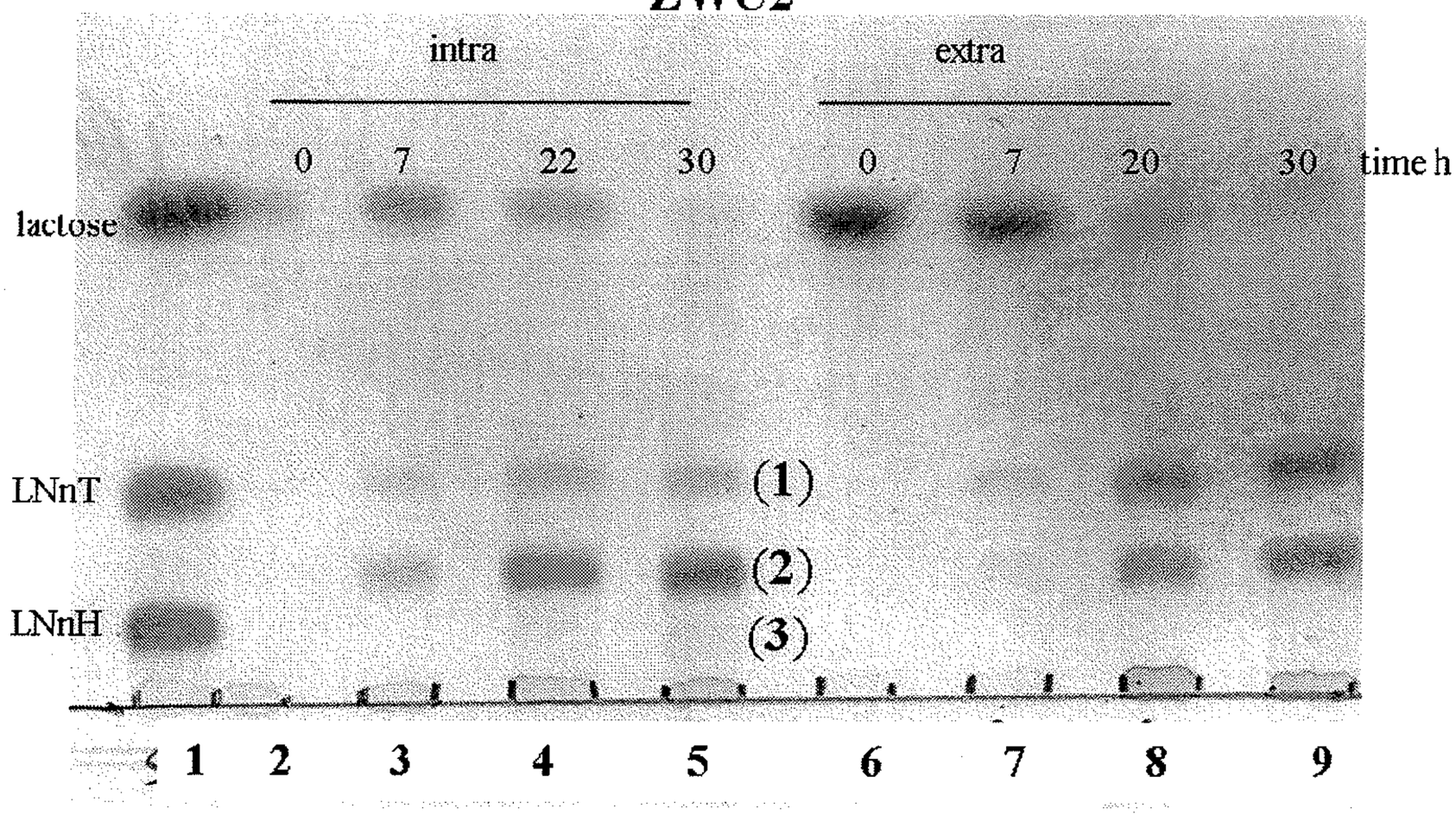
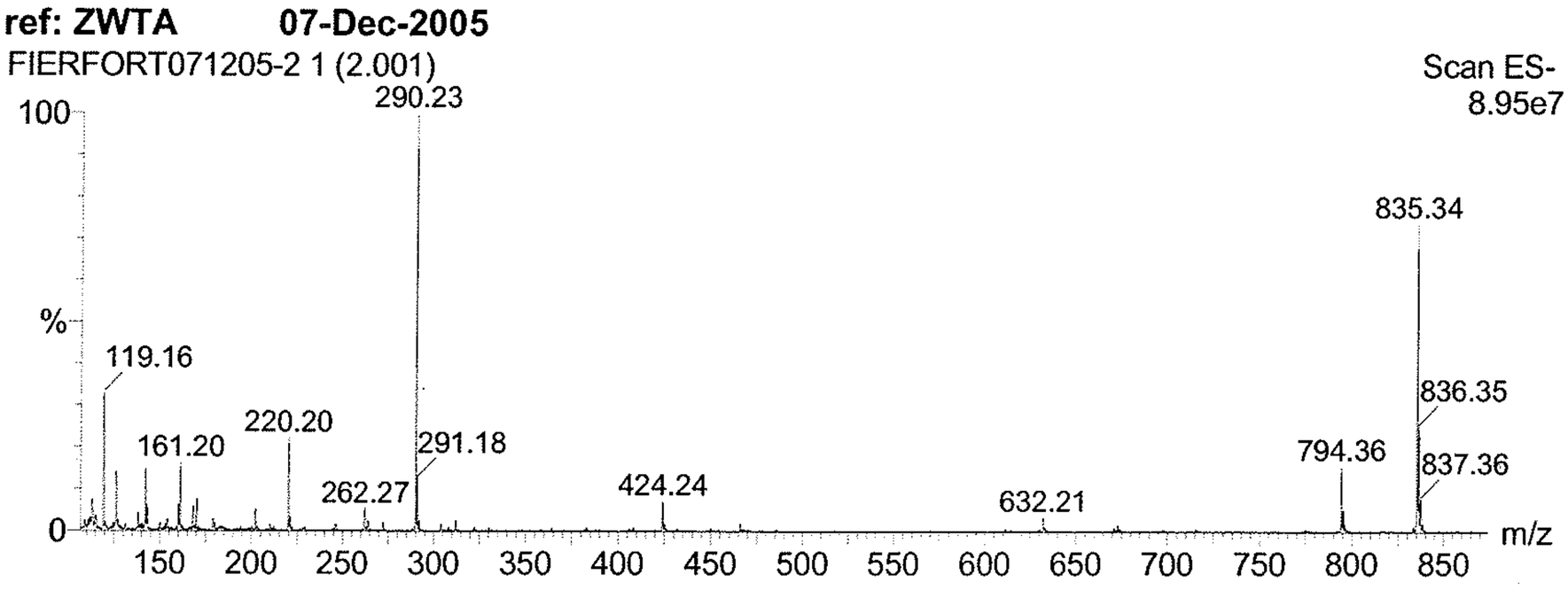


FIGURE 17

A



B

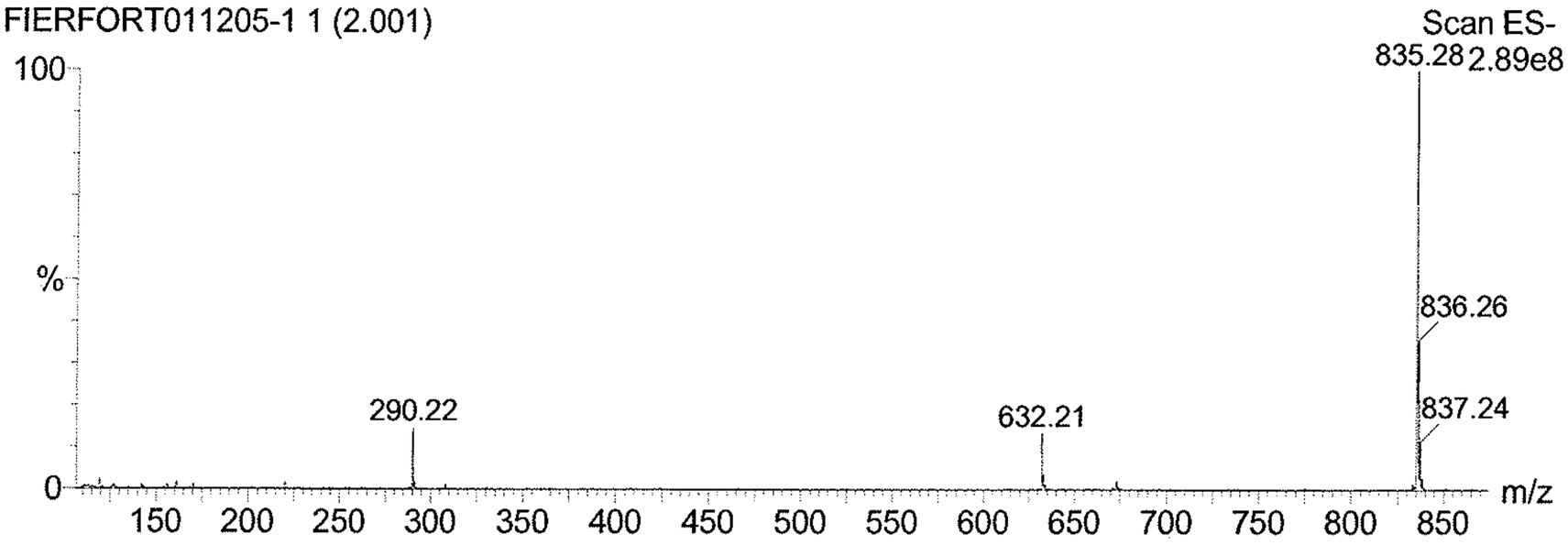


FIGURE 18

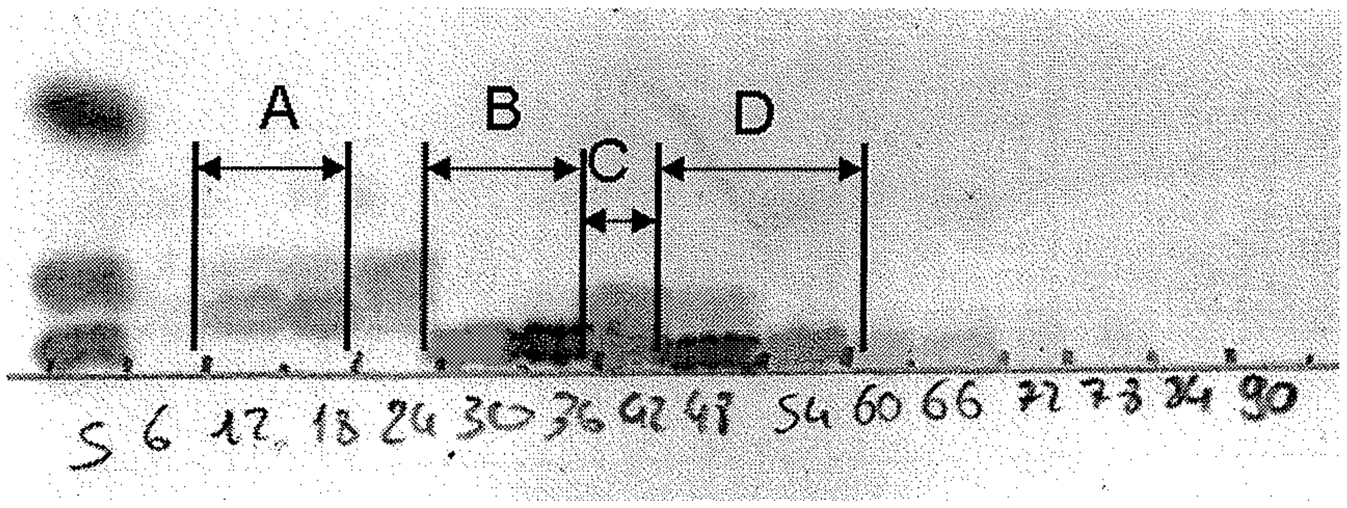


FIGURE 19

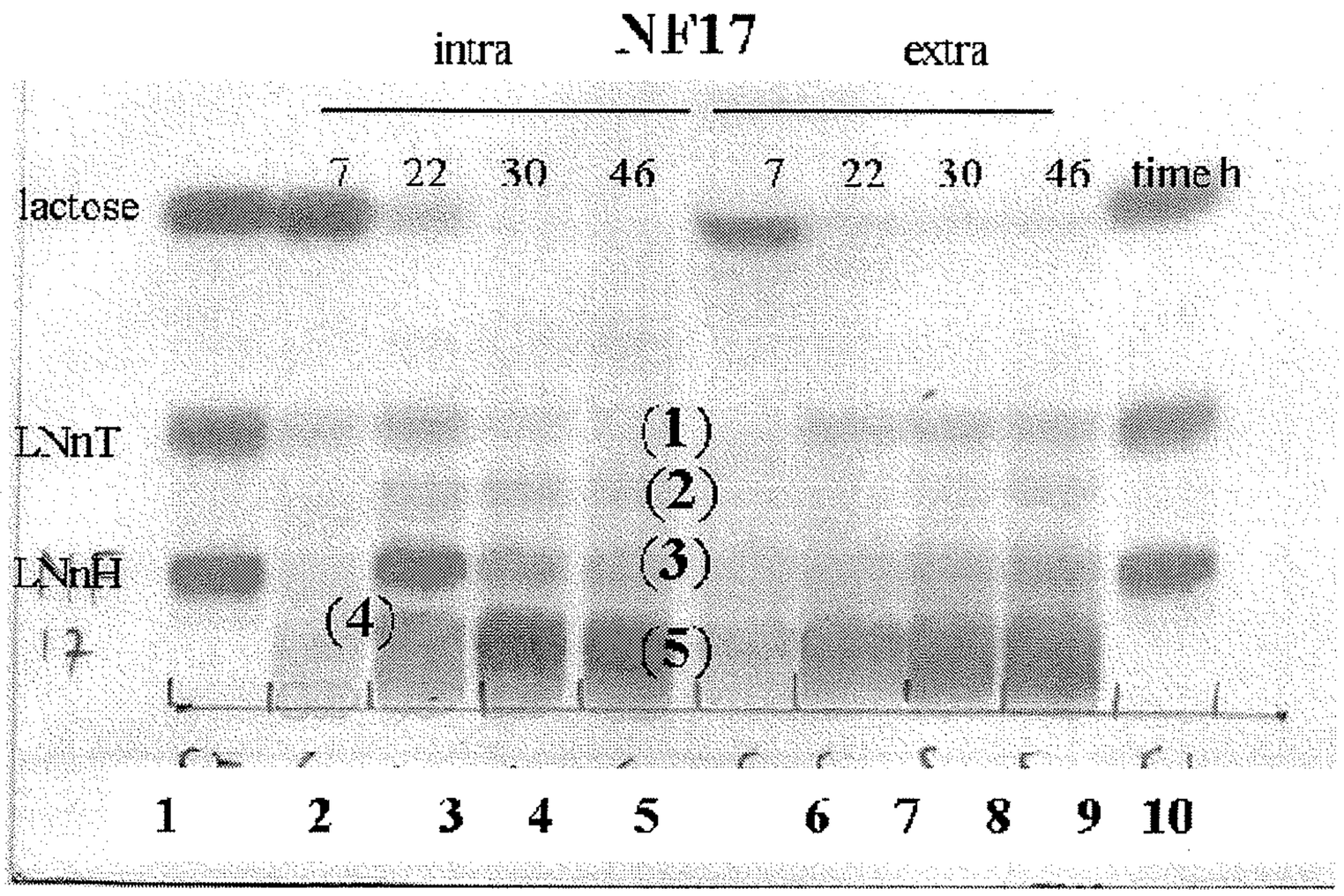


FIGURE 20