Pure Appl. Chem., Vol. 79, No. 12, pp. 2229–2242, 2007.
doi:10.1351/pac200779122229

Synthesis and biological relevance of *N*-acetylglucosamine-containing oligosaccharides*

El Sayed H. El Ashry[1,2,‡] and Mohamed R. E. Aly[3]

[1]International Center for Chemical and Biological Sciences (HEJ Research Institute for Chemistry), Karachi University, Karachi, Pakistan; [2]On leave from Chemistry Department, Faculty of Science, Alexandria University, Alexandria, Egypt; [3]Chemistry Department, Faculty of Science, Suez Canal University, Ismailia, Egypt

Abstract: The structural diversity as well as the biological significance of *N*-acetylglucosamine-containing glycans are exemplified. The problem of forming the respective glycosidic bonds of synthetic targets is addressed. Special emphasis has been given to human milk oligosaccharides (HMOs), in view of their biological relevance, and synthetic approaches of selected examples are reported.

Keywords: glycans; *N*-acetylglucosamine; glycoconjugates; glycoforms; human milk oligosaccharides; dimethylmaleoyl; protecting groups.

INTRODUCTION

Carbohydrates are amongst the most abundant biomaterials in living matter. These extraordinary biomolecules are involved in countless biologically significant events in all living organisms. Their role is essential for understanding the many biologically sophisticated pathways that have evolved into the science of glycobiology [1]. By virtue of diversity of monosaccharides, besides their linkage to proteins and lipids to give glycoconjugates and to sugars to give oligosaccharides, as well as the stereoselective nature of the glycosidic bonds, carbohydrates manifest a huge spectrum of structural diversity. They exist not only as polysaccharides but also as soluble ingredients in body fluids and as membrane-bound oligosaccharides.

Amino sugars are widely distributed in nature and exist in many glycoforms which play significant roles in vital processes with D-glucosamine being the most abundant one and existing in most of the cases as the *N*-acetyl derivative with β-glycosidic linkage. The cell membrane of living organisms is decorated with oligosaccharide, containing D-glucosamine, architectures known collectively as cell-surface oligosaccharides. These antennary molecules in the extracellular matrices are bound to the membrane via protein or lipid anchors, giving rise to glycoproteins and glycolipids, respectively. They are involved in host cell communication events, including cell–cell adhesion and proliferation, as signaling molecules, as ligands for proteins of all types, including enzymes, toxins, and antibodies. They are involved also in many immunological interactions.

The isolation of these materials from nature in pure form and in reliable amounts to evaluate their biological significance is really difficult. Therefore, synthetic approaches have been developed to prove

*Invited contribution to a collection of papers for the IUPAC project 2005-042-1-300 "Chemistry for Biology". Other contributions to the project are published in this issue, pp. 2179–2366.
‡Corresponding author

the structure, to provide sufficient material in pure form for biological investigations, and to modify or mimic these species for bioorganic applications. A cornerstone in any synthetic approach is the formation of β-linked *N*-acetylglucosamine (Glc*N*Ac) as well as the less abundant Gal*N*Ac linkages.

In this article, we throw some light on the structural diversity of β-linked Glc*N*Ac-containing glycans and their biological relevance, and address the problem of forming the glycosidic bond of 2-amino sugars during the synthesis of their glycosides. Special emphasis is given to the synthesis of human milk oligosaccharides (HMOs), as significant soluble glycans for humans in their first steps of life.

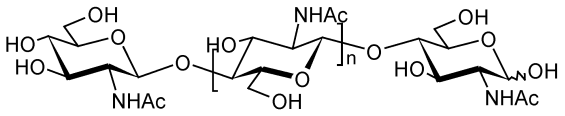
SELECTED EXAMPLES OF BIOLOGICAL RELEVANCE

This section shows the structural diversity of the complex *N*-acetyl-β-D-glucosamine-containing carbohydrates in biological systems. Examples are given below.

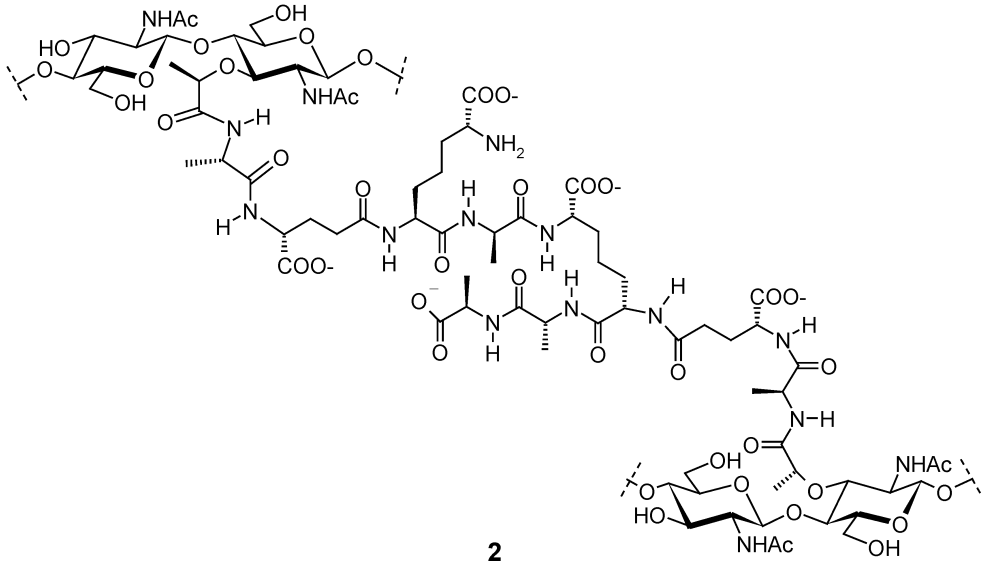
Structural polysaccharides

Glc*N*Ac is incorporated in a series of structural polysaccharides. Chitin **1**, for instance, is the most abundant one and exists in the exoskeleton of invertebrates like insects, arthropods, nematodes, *Plasmodium falciparum* and as a constituent of fungal cell walls. Chitin-hydrolyzing enzymes, known as chitinases, occur in organisms such as bacteria, fungi, plants, and vertebrates, but humans are deficient in these enzymes. Therefore, chitinase inhibitors are promising targets for development of insecticides and drugs against human pathogens such as *Candida albicans* and the human malaria parasite *P. falciparum* [2]. Natural chitinase inhibitors [3] as well as synthetic candidates [4] are rare, therefore, Glc*N*Ac-based inhibitors [2,5] are an active area of research.

Murein (**2**) is another structural polysaccharide constituting the basic component of bacterial cell wall. It consists of repeating Glc*N*Ac and muramic acid residues. The polysaccharide chains are cross-



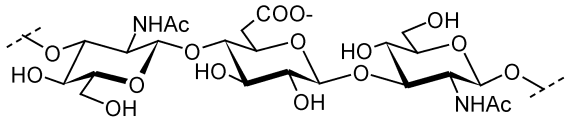
1



2

linked via peptide chains, forming a peptidoglycan frame via the DD-peptidase enzyme; synthetic fragments of murein are valuable tools for investigating this event [6,7].

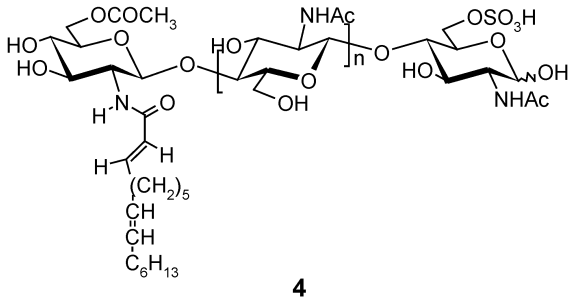
Hyaluronic acid (HA) (**3**) is another linear polysaccharide known as hyaluronan. This glycosaminoglycan is the major structural element in the extracellular matrix and responsible for viscoelastic properties of cartilages and synovial fluids. It is involved also in protein interactions whereby it influences natural processes such as angiogenesis, cell motility, and adhesion as well as wound healing and cancer [8]. Some pathogenic bacteria have extracellular HA as defensive coats against the host immune response. Synthetic Glc*N*Ac-β-(1-4)-GlcUA nucleotides and phosphates were used as probes to study the mechanism of hyaluronic acid synthases (HAS), which might lead to treatment of HA disorder-related diseases [9].



3

Signal transduction stimulants

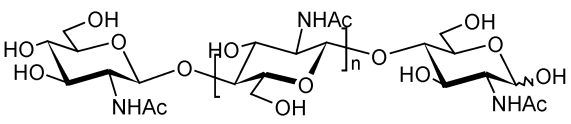
Nod factors (NFs), (e.g., **4**), are signaling molecules excreted by bacteria known as rhizobia that grow as nitrogen-fixing nodules on the roots of leguminosae. Upon nitrogen fixation in the nodules, the NFs are excreted and stimulate, as signaling molecules, the host plant through morphologic developments of the host root hairs, meristems in the root cortex, and nodule organogenesis [10,11].



4

Bacterial receptors

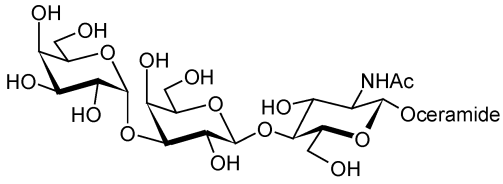
Cell-surface carbohydrates can be involved as receptors for adhesion of bacteria [12] and their toxins [13] to host cells. For instance, the Glc*N*Ac-β-(1-4)-Glc*N*Ac epitope of surface glycoproteins of brain microvascular endothelial cells (BMECs) is a ligand for *E. coli* K1 outer membrane protein A (OmpA). This recognition event mediates the traversal of *E. coli* across the blood–brain barrier causing meningitis during the neonatal period. Synthetic chitooligomers (**5**) were investigated as competitive soluble inhibitors for this recognition and might be a lead for Glc*N*Ac-based antibacterial agents.



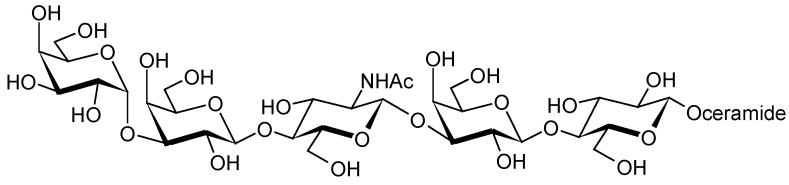
5

Antibody recognition

Another cell-surface carbohydrate interaction event is the recognition of cell-surface carbohydrates by antibodies. IgG, the most abundant antibody in human blood serum, recognizes terminal Gal-α-(1-3)-Gal residues, known as the (α-Gal) epitopes, of animal xenografts and mediates hyperacute rejection upon xenotransplantation [14]. Three cell-surface carbohydrates have this antigen; two of them (**6** and **7**) have a Glc*N*Ac moiety at the oligosaccharide core.

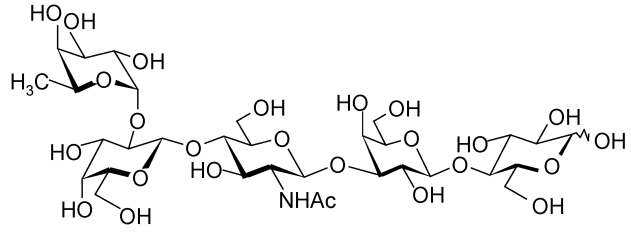


6

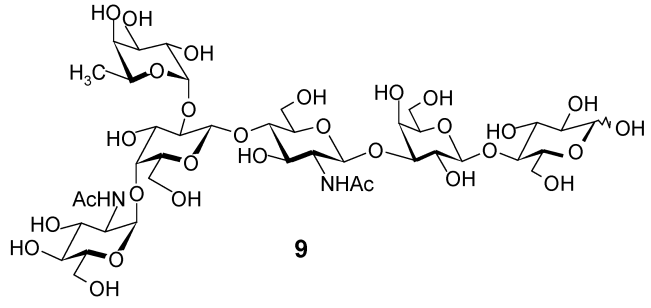


7

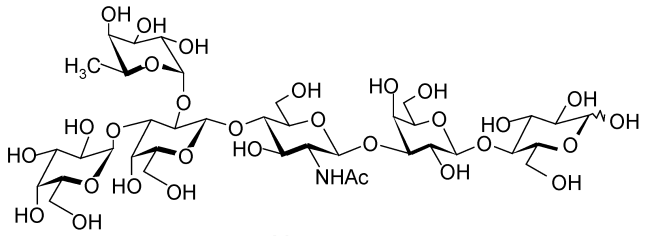
Blood group antigens **8**, **9**, and **10** are Glc*N*Ac-containing surface glycoproteins on red blood cells (erythrocytes). These antigenic dominants are characteristic for individuals and can recognize selective antibodies from persons having different blood groups H, A, and B [15].



8



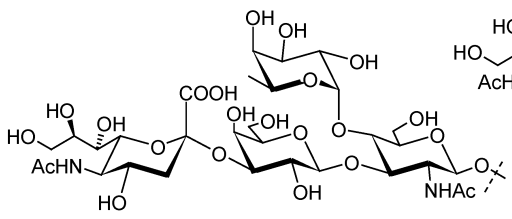
9



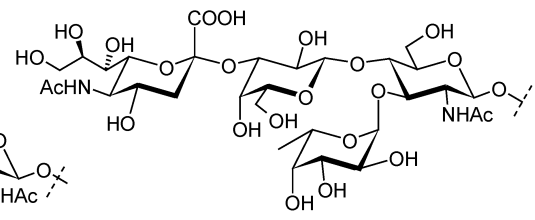
10

Cancer metastasis

Malignant tumor cells frequently express the surface glycolipids known collectively as tumor-associated antigens or simply tumor markers. These antigens play an important role in diagnosis of malignancy via immune detection with their selective antibodies. Sialyl-Lewis X (sLe^x) **12** and sialyl-Lewis A (sLe^a) **11** are Glc*N*Ac-containing epitopes on many cancer cells. These epitopes serve as ligands for natural host membrane-bound proteins known as selectins. Adhesion of these epitopes with selectins is thought to play significant roles in blood-born metastasis [16–18].



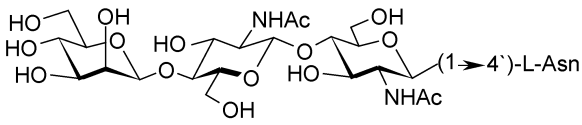
11



12

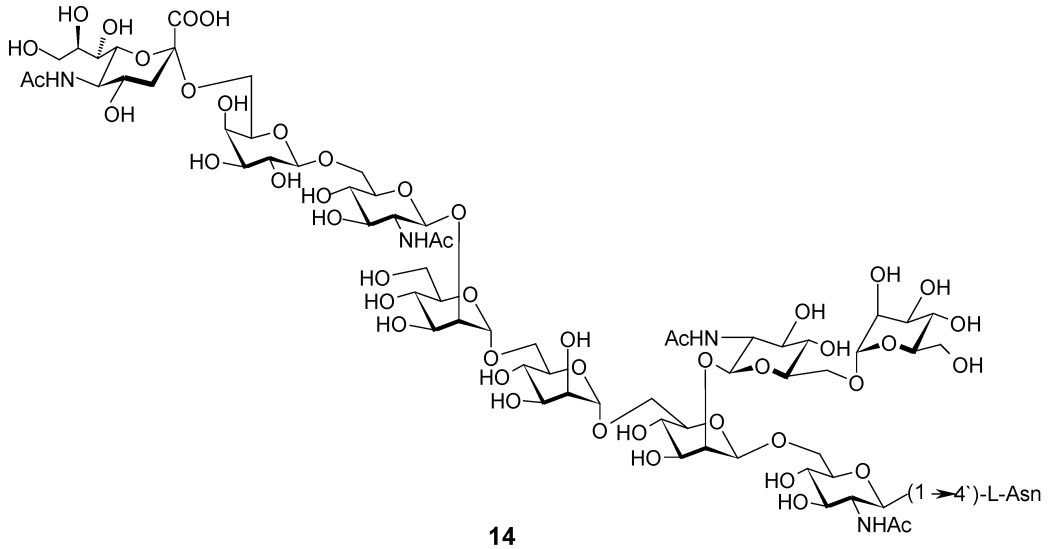
Hormonal and enzymatic glycoproteins

Glycoproteins having Glc*N*Ac residues are known to function as enzymes. Bromelain (**13**), for instance, is a glycoprotein enzyme found in the plant kingdom [19].



13

GlcNAc is found also as part of the structure of hormonal glycoproteins such as follicle-stimulating hormone (FSH), luteinizing hormone (LH), human menopausal gonadotropin (hMP), pregnant-mare serum gonadotropin, thyroid-stimulation hormone (TSH), and human chrionic gonadotropin (hCG) (**14**) [19].



14

Human milk oligosaccharides

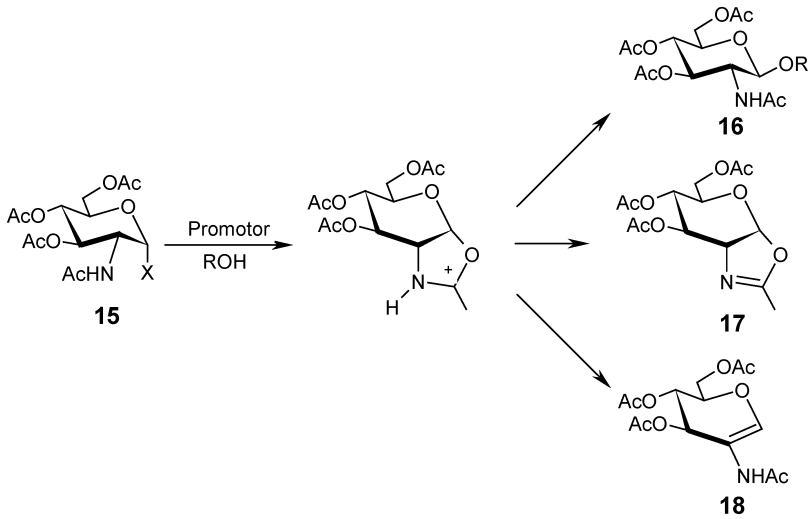
Human milk is characterized by inimitable soluble oligosaccharides known collectively as the HMOs, which are not available in the milk of other mammals. These unbound oligosaccharides consist mainly of a lactose moiety as reducing terminal attached to lactosamine and/or isolactosamine at the non-reducing terminal of the molecule in a linear or branched manner. They are frequently L-fucosylated and sialylated via α-glycosidic linkages [20,21]. The latter form the so-called acidic oligosaccharides in human milk [22]. HMOs exist in concentrations as high as 5–10 g/L, thus, they constitute the third largest ingredients after proteins and lipids [23]. Their contribution to the health of breast-fed vs. formula-fed infants was recognized as early as 1892 [24]. They are, structurally, divergent oligosaccharides, and over 130 types are identified right now [25]. They are resistant to hydrolytic enzymes in the gastrointestinal tract, therefore, they can reach the large intestine and cross into the bloodstream as soluble matter [26] where they act as prebiotics [27] and competitive ligands to pathogenic microorganisms and inhibit their adhesion to the epithelial cells [28]. Thus, they possess significant defensive roles, and they are associated with less reported gastrointestinal, respiratory, and ear infections and even reduced postneonatal death in breast-fed infants [24]. On the other hand, they act as a source of sialic acid and D-galactose, which are essential for brain development and improved learning ability [29]. At the same time, sialyl oligosaccharides in human milk possess binding epitopes of selectin ligands, thus, they impair or decoy the interaction of selectins with cellular ligands and modulate the migration of leukocytes from the blood stream into the inflammation regions. In this way, they serve as anti-inflammatory components in human milk that reduce inflammatory diseases such as necrotizing enterocolitis in breast-fed vs. formula-fed infants. They modulate also the formation of platelet-neutrophil complex (PNC) involved by the immune system [30]. The list of defensive roles gained by HMOs is growing, and, therefore, synthetic access to these hardly separable ingredients is of great demand as it might lead to new therapeutics and vaccines. We present here some synthetic probes of these oligosaccharides.

SYNTHESIS OF 2-ACETAMIDO-2-DEOXY-β-D-GLUCOPYRANOSIDES

One of the consequences of development of glycoscience is the need for the synthesis of naturally occurring oligosaccharides. This requirement allows modification of naturally occurring glycans to be used for specific biological purposes and for searching more effective analogs of these markers. The problem of preparing the above-mentioned materials or their analogs is mainly the configuration of the

anomeric center, which is generated during the glycosylation process, and which could afford either α- or β-linked glycosides or a mixture of both. The glycosylation processes require a glycosyl donor and acceptor where the substituents, particularly at position-2 on the donor, play a major role in the stereochemical outcome of the resulting glycoside. The protecting groups of the amino function have had a major impact in this regard, to ensure the *trans* 1,2-coupling, and they can be transformed to the acetyl moiety [31]. In general, these requirements still constitute one of the most compelling topics in carbohydrates synthesis [32–34].

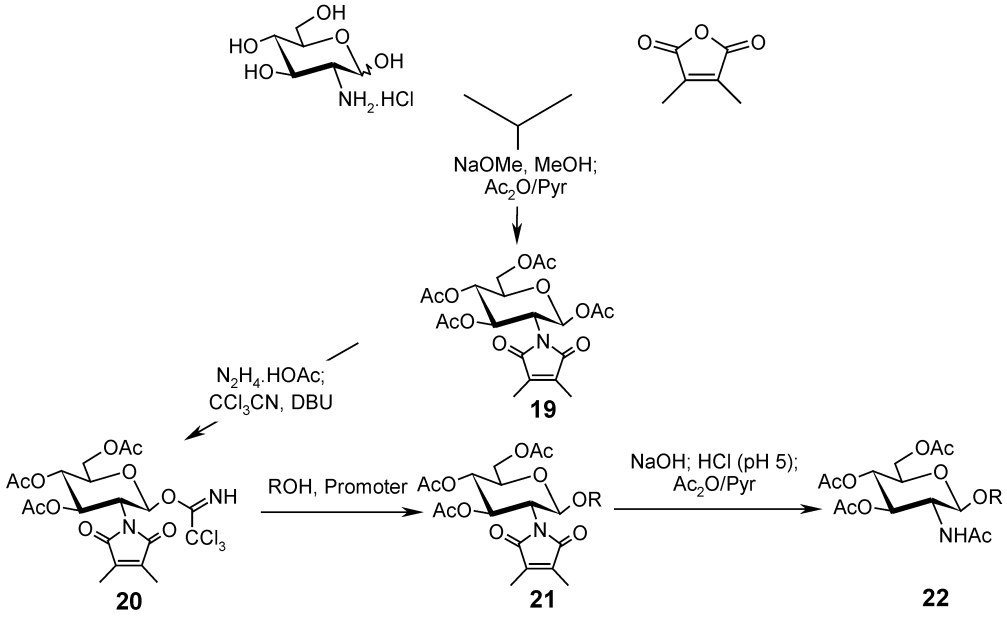
The early Koenigs–Knorr method was used, whereby 2-acetamidoglucopyranosyl halides (**15**), as glycosyl donors, were coupled with alcohols in the presence of halophilic promoters like mercuric salts. Of course, this donor should permit β-coupling due to participation by the acetyl moiety and afford directly the desired acetamido glycoside (**16**). However, this ideal hypothesis was recognized only with reactive acceptors having primary OH groups, while, with weak acceptors having secondary OH-group(s), a stable oxazoline (**17**) was mainly formed and to a lesser extent glucal (**18**) due to 1,2-elimination, Scheme 1.



Scheme 1 Koenigs–Knorr glycosylation reactions.

Several alternatives for Koenigs–Knorr donors were continuously developed to solve the problem of oxazoline formation. A crucial condition is the availability of a protecting group for the amine functionality of glucosamine that permits β-coupling and then can be substituted easily with the acetyl moiety in high yields. In addition, it should be stable to the conditions involved along the whole synthetic protocol and it should be symmetric if possible to avoid problems related to NMR analysis. Of these, we mention here the phthalimido (Phth) and trichlorethoxycarbonyl (Troc) amino protecting groups. These groups in combination with thioalkyl and trichloroacetimidate as reactive anomeric leaving groups have gained wide application in this field and led successfully to a wide range of glycans and glycoconjugates. Also, the 2-azido derivatives were used for directing the glycosylations towards the β-face, which then reduced and acetylated to give the target glycosides. Apparently, the Phth as well as the Troc groups support the β-coupling via neighboring group participation with a carbonyl group. The azido group does the same under the influence of acetonitrile that blocks the α-face of the intermediate oxocarbenium ion via formation of an intermediate α-nitrilium ion, known as the nitrile effect. Recently, the dimethylmaleoyl (DMM) group was developed as a successful amino protecting group in oligosaccharide synthesis [35]. Scheme 2 is a model scheme for introducing the DMM group as in **19**,

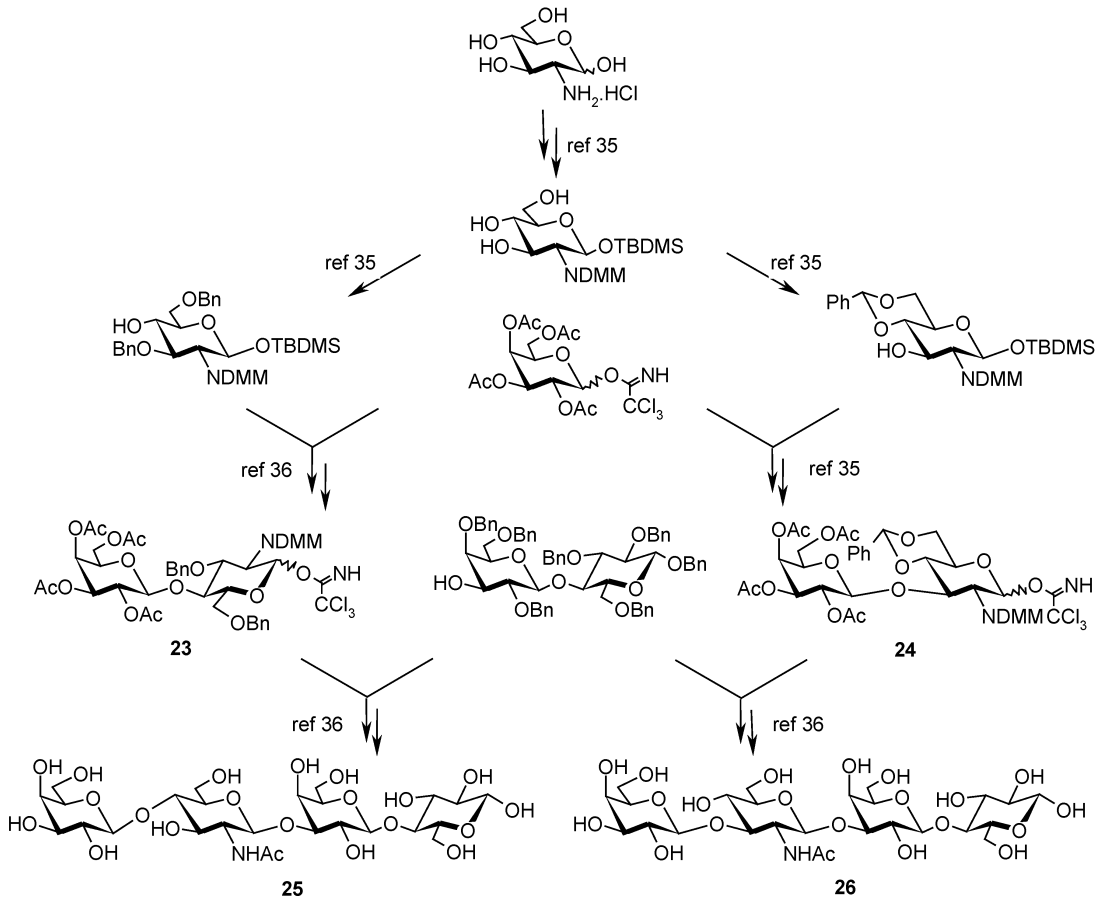
tuning it as a model donor **20**, and then reaction with a general acceptor to afford glycoside **21** followed by converting it to the acetyl group as in **22**. In the next section, the contribution of this approach is described for the synthesis of some HMOs in solution as well as on solid supports.



Scheme 2 DMM glycosylation method.

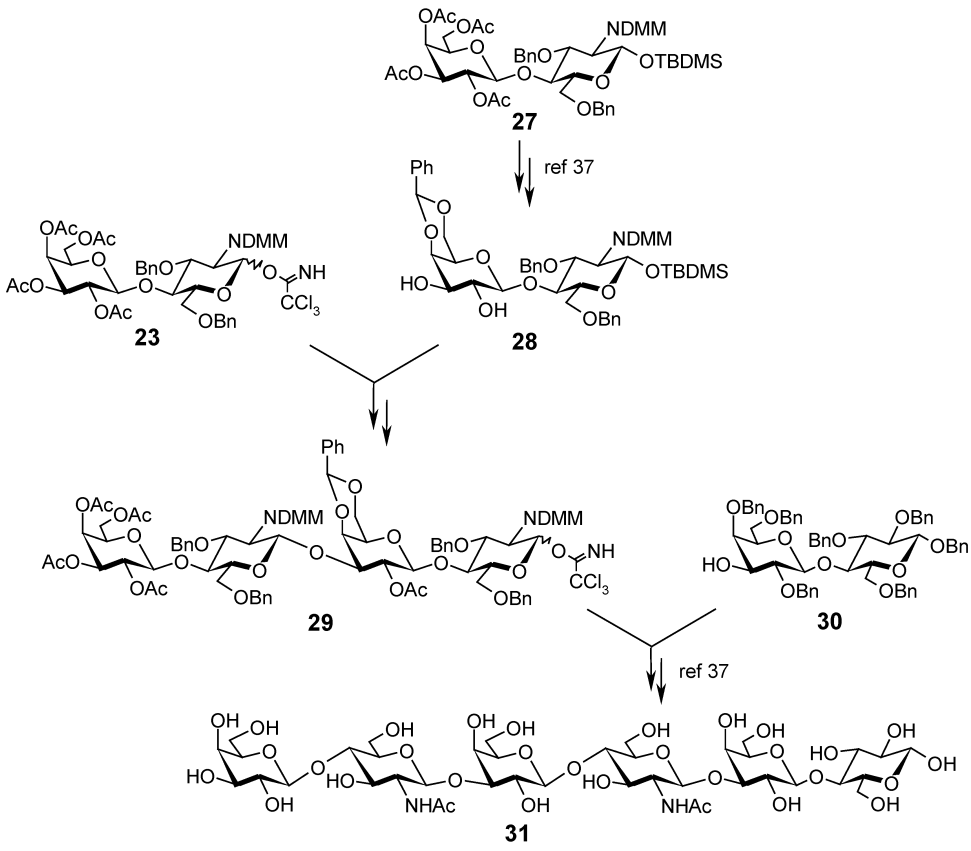
SELECTED SYNTHESES OF HUMAN MILK OLIGOSACCHARIDES

By virtue of the stability of the DMM group, compatibility to various protecting group manipulations and good reactivity of NDMM-protected glycosyl donors as well as acceptors, both of the donors lactosamine **23** and isolactosamine **24** could be synthesized easily. These disaccharide donors are the essential building blocks of the backbone of HMOs, therefore, lacto-*N*-neotetraose **25** and lacto-*N*-tetraose **26** could be synthesized efficiently, Scheme 3 [35,36].

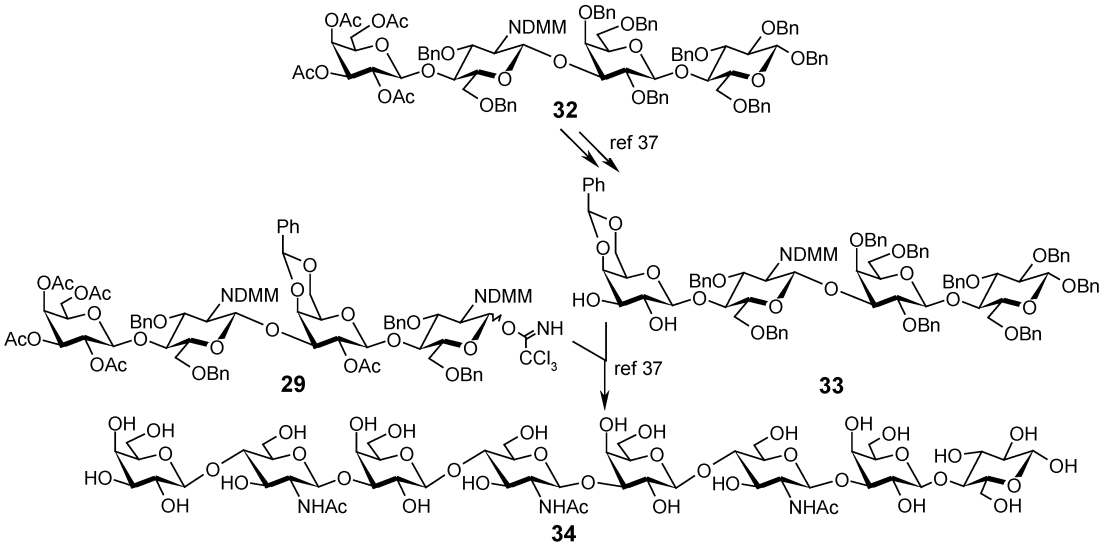


Scheme 3 Synthesis of lacto-*N*-neotetraose and lacto-*N*-tetraose.

Another advantage of the DMM approach is the performance of efficient coupling in a regioselective manner. Thus, coupling of the NDMM-protected glycosyl donor **23** with diol (**28**), prepared from **27** by deacetylation and subsequent benzylidenation, having two (2b,3b)-OH free groups, Scheme 4, afforded regioselectively the β-(1-3) linked glycoside in high yield. Desilylation allowed the synthesis of the trichloroacetimidate **29**. Reaction of the tetrasaccharide donor with the disaccharide acceptor **30** gave the HMOs lacto-*N*-neohexaose **31**, Scheme 4. Similarly, **32** was transformed into **33**, whose reaction with **29** gave lacto-*N*-neooctaose **34**, Scheme 5 [37].

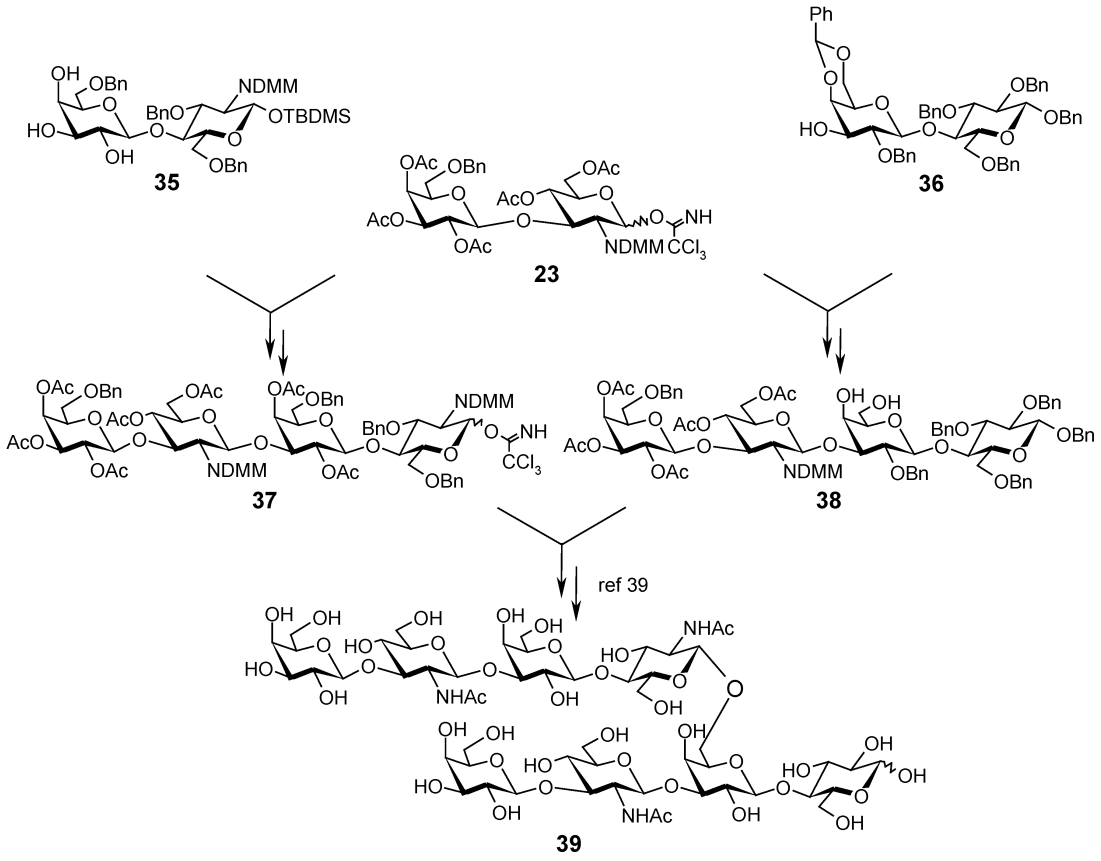


Scheme 4 Synthesis of lacto-*N*-neohexaose.



Scheme 5 Synthesis of lacto-*N*-neooctaose.

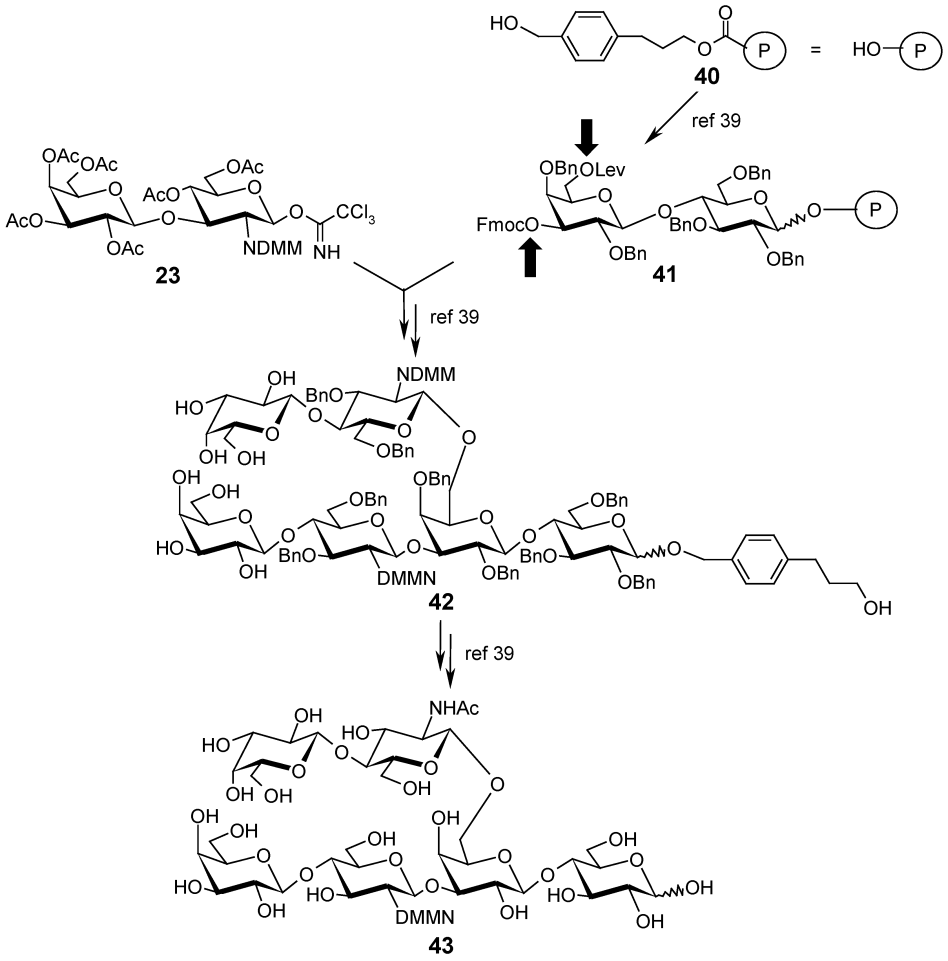
Branched HMOs containing lactosamine as well as isolactosamine building blocks could also be easily synthesized. Again, the regioselective β-(1-3) glycosylation with NDMM-protected glycosyl donors even with triol **35** led to tetrasaccharide donor **37**, which upon reaction with acceptor **38**, prepared from **36**, led to the octasaccharide **39**, Scheme 6 [38].



Scheme 6 Synthesis of a branched octasaccharide.

Finally, the general ease of purification after DMM deprotection was an advantage of this approach that permitted separation of target HMOs in rather high yields.

HMOs were also accessible via a solid-phase synthesis (SPS). Thus, a precursor (**41**), prepared from benzylic-type linker (**40**), as latent acceptor after suitable deprotection, was coupled with donor **23** to give **42**, which can be transformed into the anomeric free-branched saccharide lacto-*N*-neohexaose **43**, Scheme 7. The SPS of HMOs has been proposed to be of industrial importance as a lead for these important oligosaccharides to modify the defensive properties of formula milk [39].



Scheme 7 SPS of branched HMOs.

ACKNOWLEDGMENTS

The valuable help and encouragement of Prof. Dr. Richard R. Schmidt, AvH foundation in Germany and HEC in Pakistan, are highly appreciated.

REFERENCES

1. (a) A. Kobata. *Proc. Jpn. Acad.* **81B**, 418 (2005); (b) C. R. Bertozzi, L. L. Kiessling. *Science* **291**, 2357 (2001); (c) A. Helenius, M. Aebi. *Science* **291**, 2364 (2001); (d) J. Alper. *Science* **291**, 2338 (2001); (e) J. C. McAuliffe, O. Hindsgaul. *Chem. Ind.* 170 (1997); (f) E. F. Hounsel. *Carbohydr. Res.* **300**, 47 (1997); (g) P. M. Simon. *Drug Discov. Today* **1**, 522 (1996); (h) R. A. Dwek. *Chem. Rev.* **96**, 683 (1996); (i) A. Varki. *Glycobiology* **3**, 97 (1993); (j) A. Kobata. *Acc. Chem. Res.* **26**, 319 (1993); (k) R. Kornfeld, S. Kornfeld. *Ann. Rev. Biochem.* **54**, 631 (1985); (l) H. Herzner, T. Reipen, M. Schultz, H. Kunz. *Chem. Rev.* **100**, 4495 (2000).
2. G. Vaaje-Kolstad, A. Vasella, M. G. Peter, C. Netter, D. R. Houston, B. Westereng, B. Synstad, V. G. H. Eijsink, D. M. F. van Aalten. *J. Biol. Chem.* **279**, 3612 (2004).
3. (a) F. V. Rao, D. R. Houston, R. G. Boot, J. M. F. G. Aerts, S. Sakuda, D. M. F. van Aalten. *J. Biol. Chem.* **278**, 20110 (2003); (b) N. Asano. *Glycobiology* **13**, 93R (2003).

4. F. V. Rao, O. A. Andersen, K. A. Vora, J. A. DeMartino, D. M. F. van Aalten. *Chem. Biol.* **12**, 973 (2005).
5. A. R. Yeager, N. S. Finney. *J. Org. Chem.* **69**, 613 (2004).
6. D. Hesek, M. Suvorov, K.-I. Morio, M. Lee, S. Brown, S. B. Vakulenko, S. Mobashery. *J. Org. Chem.* **69**, 778 (2004).
7. D. Hesek, M. Lee, K.-I. Morio, S. Mobashery. *J. Org. Chem.* **69**, 2137 (2004).
8. P. L. DeAngelis. *J. Biol. Chem.* **274**, 26557 (1999).
9. H. Takaku, H.-K. Ishida, M. Fujita, T. Inazu, H. Ishida, M. Kiso. *Synlett* 818 (2007).
10. J.-C. Promé, J. Denarié, G. Truchet. *Pure Appl. Chem.* **70**, 55 (1998).
11. G. E. D. Oldroyd, R. M. Mitra, R. J. Wais, S. R. Long. *Plant J.* **28**, 191 (2001).
12. N. V. Parasadarao, C. A. Wass, K. S. Kim. *Infect. Immun.* **64**, 154 (1996).
13. M. R. E. Aly, P. Rochaix, M. Amessou, L. Johannes, J.-C. Florent. *Carbohydr. Res.* **341**, 2026 (2006).
14. J.-Q. Wang, X. Chen, W. Zhang, S. Zacharek, Y. C. Chen, P. G. Wang. *J. Am. Chem. Soc.* **121**, 8174 (1999).
15. W. M. Watkins. *Adv. Human Genet.* **10**, 1 (1980).
16. M. Izawa, K. Kumamoto, C. Mitsuoka, A. Kanamori, K. Ohmori, H. Ishida, S. Nakamura, K. Kurata-Miura, K. Sasaki, T. Nishi, R. Kannagi. *Cancer Res.* **60**, 1410 (2000).
17. M. Ugorski, A. Laskowska. *Acta Biochim. Pol.* **49**, 303 (2002).
18. K. D. Patel, S. L. Cuvelier, S. Wiehler. *Semin. Immunol.* **14**, 73 (2002).
19. J. F. Kennedy, C. A. White. *Bioactive Carbohydrates in Chemistry, Biochemistry and Biology*, Ellis Horwood, Chichester (1983).
20. P. Chaturvedi, C. D. Warren, M. Altaye, A. L. Morrow, G. Ruiz-Palacios, L. K. Pickering, D. S. Newburg. *Glycobiology* **11**, 365 (2001).
21. M. Gudiel-Urbano, I. Goni. *Arch. Latinoam. Nutr.* **51**, 332 (2001).
22. S. Martin-Sosa, M. J. Martin, L. A. Garcia-Pardo, P. Hueso. *J. Dairy Sci.* **86**, 52 (2003).
23. P. McVeagh, J. B. Miller. *J. Paediatr. Child Health* **33**, 281 (1997).
24. M. B. F. Stepans, S. L. Wilhelm, M. Hertzog, T. K. C. Rodehorst, S. Blaney, B. Clemens, J. J. Polak III, D. S. Newburg. *Breastfeeding Med.* **1**, 207 (2006) and refs. cited therein.
25. L. Bode. *J. Nutr.* **136**, 2127 (2006).
26. M. B. Engfer, B. Stahl, B. Finke, G. Sawatzki, H. Daniel. *Am. J. Clin. Nutr.* **71**, 1589 (2000).
27. G. R. Gibson, M. Roberfroid. *J. Nutr.* **125**, 1401 (1995).
28. G. V. Coppa, L. Zampini, T. Galeazzi, B. Facinelli, L. Ferrante, R. Capretti, G. Orazio. *Pediatr. Res.* **59**, 377 (2006).
29. B. Wang, J. Brand-Miller, P. McVeagh, P. Petocz. *Am. J. Clin. Nutr.* **74**, 510 (2001).
30. L. Bode, J. Rudloff, C. Kunz, J. Strobel, N. Klein. *J. Leukoc. Biol.* **76**, 820 (2004).
31. (a) J. Banoub, P. Boullanger, D. Lafont. *Chem. Rev.* **92**, 1167 (1992); (b) J. Debenham, R. Rodebaugh, B. Fraser-Reid. *Liebigs Ann.* 791 (1997); (c) G.-J. Boons. In *Carbohydrate Chemistry*, Blackie A & P, Glasgow (1998); (d) S. Hanessian. In *Preparative Carbohydrate Chemistry*, Marcel Dekker, New York (1997); (e) A. F. G. Bongat, A. V. Demchenko. *Carbohydr. Res.* **342**, 374 (2007); (f) S. E. Zurabyan, A. Ya. Khorlin. *Russ. Chem. Rev.* **43**, 887 (1974).
32. E. S. H. El Ashry, N. Rashed, E. S. I. Ibrahim. *Curr. Org. Synth.* **2**, 175 (2005).
33. E. S. H. El Ashry, A. El Nemr. *Synthesis of Naturally Occurring Nitrogen Heterocycles from Carbohydrates*, Blackwell, Oxford (2005).
34. R. R. Schmidt, J. Castro-Palomino, O. Retz. *Pure Appl. Chem.* **71**, 729 (1999).
35. M. R. E. Aly, J. Castro-Palomino, E. I. Ibrahim, E. S. H. El Ashry, R. R. Schmidt. *Eur. J. Org. Chem.* 2305 (1998).
36. M. R. E. Aly, E. I. Ibrahim, E. S. H. El Ashry, R. R. Schmidt. *Carbohydr. Res.* **316**, 121 (1999).
37. (a) M. R. E. Aly, E. I. Ibrahim, E. S. H. El Ashry, R. R. Schmidt. *Eur. J. Org. Chem.* 319 (2000); (b) M. R. E. Aly, E. I. Ibrahim, E. H. El Ashry, R. R. Schmidt. *Carbohydr. Res.* **331**, 129 (2001).

38. P. Knuhr, J.-C. Palomino, M. Grathwohl, R. R. Schmidt. *Eur. J. Org. Chem*. 4239 (2001).
39. F. Roussel, M. Takhi, R. R. Schmidt. *J. Org. Chem.* **66**, 8540 (2001).