Enzymatic cascade synthesis provides novel linear human milk oligosaccharides as reference standards for xCGE-LIF based high-throughput analysis[†]

Thomas Fischöder[1,†], Samanta Cajic[2,†], Udo Reichl[2,4], Erdmann Rapp[2,3] and Lothar Elling[1]

[1]Laboratory for Biomaterials and Institute for Biotechnology and Helmholtz-Institute for Biomedical Engineering RWTH Aachen University, Aachen, Germany

[2]Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany

[3]glyXera GmbH, Magdeburg, Germany

[4]Chair of Bioprocess Engineering Otto-von-Guericke-University, Magdeburg, Germany

[†]These authors contributed equally to this work

Correspondence: 1) Prof. Dr. rer. nat. Lothar Elling, Laboratory for Biomaterials, Institute for Biotechnology and Helmholtz Institute for Biomedical Engineering RWTH Aachen University, Pauwelsstrasse 20, D-52074 Aachen, Germany.
E-mail: l.elling@biotec.rwth-aachen.de

2) Dr. rer. nat. Erdmann Rapp, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse 1, D-39106 Magdeburg, Germany.
E-mail: rapp@mpi-magdeburg.mpg.de

Abbreviations:

3GALase, β(1-3) galactosidase; **4GALase**, β(1-4) specific galactosidase; **AP**, alkaline phosphatase; **APTS**, 8-aminopyrene-1,3,6-trisulfonic acid; **β4GalT**, β1,4-galactosyltransferase-1; **β3GalT**, β1,3-galactosyltransferase; **β3GlcNAcT**, β1,3-*N*-acetylglucosaminyltranserase; **CE-LIF**, capillary electrophoresis with laser-induced fluorescence detection; **CGE-LIF**, capillary gel electrophoresis with laser-induced fluorescence detection; **xCGE-LIF**, multiplexed capillary gel electrophoresis with laser-induced fluorescence detection; **GlcNAc**, *N*-acetylglucosamine; **GlcNAcase**, β(1-4,6) *N*-Acetylglucosaminidase; **HILIC**, hydrophilic-interaction-chromatography **HMOS**, human milk oligosaccharides; **HTS**, high throughput screening; **ID**, inner diameter; **LacNAc**, *N*-acetyllactosamine; **LNT**, Lacto-*N*-biose-type; **LNnT**, Lacto-*N*-neo-type; **MALDI-TOF-MS**, matrix assisted laser desorption ionization – time of flight - mass spectrometry; **MBP**, maltose binding protein; **MTU**, migration time units; **NQ**, not quantifiable; **PVA**, poly vinyl alcohol; **RFU**, relative fluorescence units

Abstract

A rising amount of known health benefits leads to an increased attention of science and nutrient industry to human milk oligosaccharides (HMOS). The unique diversity of HMOS includes several rare, complex and high molecular weight structures. Therefore, identification and elucidation of complex structures, which may occur only in traces, poses a daunting analytical challenge, further complicated by the limited access to suitable standards. Regarding this, inherent diversity of HMOS and their structural complexity make them difficult to synthesize. The use of recombinant *Leloir*-glycosyltransferases offers a common strategy to overcome the latter issues. In this study, we were able to tailor linear long-chained Lacto-*N*-biose-type (LNT) and Lacto-*N*-neo-type (LNnT) HMOS far beyond the known naturally occurring length. Thereby we provide novel well-defined reference standards for screening HMOS composition by high performance and high throughput analytics. We show here for the first time the synthesis of LNT oligomers up to 26 and LNnT oligomers up to 30 sugar units in a semi-sequential one-pot synthesis as analyzed by high performance multiplexed capillary gel electrophoresis with laser-induced fluorescence detection (xCGE-LIF). While being a high-throughput method, xCGE-LIF can also handle long chained linkage isomers of challenging similarity, some of them even present only in trace amounts.

1 Introduction

From time immemorial midwives and mothers have known about the benefits of breastfeeding for the newborn infants. In the last decades, several important health benefits far beyond the simple nutrition were revealed and dedicated to HMOS.[1] HMOS provide only a part of the human milk composition ranging from 5 to 23 g/L,[2] but they present a nearly unique diversity in the mammalian class.[2b, 3] The extreme diverse HMOS composition includes several complex high molecular structures, only present in very small amounts.[4] The HMOS composition depends on the phenotype, and it changes during lactation.[5] A straightforward technique, which can handle only traces of compounds, is capillary gel electrophoresis with laser-induced fluorescence detection (CGE-LIF).[6] Development and improvement of novel high throughput methods and databases offer novel opportunities like fingerprinting.[7] Nevertheless, structural elucidation of HMOS compounds that are present in human milk only in traces stays an extremely challenging task. In particular, very rare structures or only putative existing structures are usually not available as reference standards. Without any reference and the possibility to annotate these structures in common databases, the identification and structural elucidation of these structures in native samples are even more challenging. Providing tailored HMOS and HMOS like structures beyond the known existing structural diversity as references and standards can be a promising strategy to identify and elucidate novel HMOS structures. The enzymatic synthesis of HMOS is a common strategy.[4a, 8] Several *in vivo* and *in vitro* systems can be found in the literature.[4a, 9] In that respect *in vitro* systems provide the benefit to deliver well-defined products or product distributions and can be directed to form high molecular structures. HMOS are generally based on lactose and are propagated with repeating *N*-acetyllactosamine (LacNAc) units in a linear or a branched pattern.[3b, 4a, 4c] Enzymatic cascade reactions including recombinant human β1,4-galactosyltransferase-1 (β4GalT),[10] bacterial β1,3-*N*-acetylglucosaminyltranserase (β3GlcNAcT),[11] and

4

bacterial β1,3-galactosyltransferase (β3GalT)[12] have been used for one-pot and sequential poly-LacNAc synthesis.[12a, 13] Here we present, to the best of our knowledge for the first time, the integration of these enzymatic tools in semi-sequential setups for the synthesis of linear Lacto-*N*-biose-type (LNT) and Lacto-*N*-neo-type (LNnT) HMOS, far beyond the known naturally occurring length, up to 26 and 30 sugar units, respectively. In this way, the efficient method provides a wide range of LNT, LNnT HMOS and HMOS like oligosaccharides as reference standards for human milk structural analysis, e.g. by multiplexed capillary gel electrophoresis with laser-induced fluorescence detection (xCGE-LIF), a powerful tool for glycoanalysis.[7c] We further demonstrate the high resolution of this high-performance tool by separating challenging linkage isomers of the Lacto-*N*-biose-Type and Lacto-*N*-neo-type. Our approach allows monitoring of product formation in complex enzymatic reaction cascades including the verification of desired isomers. Pushing the xCGE-LIF based analytics to its limit reveals the formation of up to 30mer LNnT and 26mer LNT like oligomers and gives new insights into the substrate specificity of the involved enzymes.

2 Materials and methods

2.1 Production of recombinant enzymes

Protein expression and purification were carried out as reported previously.[10, 12a, 13-14] The β1,3-*N*-acetylglucosaminyltransferases from *Helicobacter pylori* was produced as a fusion with an N-terminal maltose binding protein (MBP-β3GlcNAcT, designated as β3GlcNAcT) in *E. coli* BL21 (DE3) (Novagen/Merk) as described previously.[11-12, 13, 14c] The human β1,4-galactosyltransferase-1 fusion protein (β4GalT) carrying an N-terminally fused lipase pre-propeptide from *Staphylococcus hyicus* with His$_6$-tag and was expressed in *E. coli* SHuffle T7 Express (New England Biolabs).[10, 12a, 13] The β1,3-galactosyltransferase from *E. coli* O55:H7[12b] with a C-terminal fused pp and His$_6$-tag

(β3GalT) was produced in *E.coli* BL21 (DE3) (Novagen/Merk) as described previously.[12a] Cells were harvested by centrifugation after 24h cultivation. The cells were disintegrated by sonificating a 40% w/v cell suspension. After centrifugation (13,000 rpm, 5 min) the supernatant was used for purification. His_6-tagged proteins were isolated by a HisTrapTM HP 5 mL column (GE Healthcare) and MBP-tagged proteins by an MBPTrap HP 5 mL column also from GE Healthcare as recommended by the manufacturer. For purification of the β3GalT, 2% TritonX was added to the elution buffer. For the β3GlcNAcT and the β3GalT, a buffer exchange was carried out by dialysis. The β3GlcNAcT dialysis buffer contains 20 mM Tris-HCl at pH 7.4, 200 mM NaCl, 1 mM EDTA and 1 mM DTT. The β3GalT dialysis buffer contains 100 mM NaH_2PO_4 at pH 7.5 500 mM NaCl and 5 mM DTT.

2.2 Capillary electrophoresis - laser-induced fluorescence detection (CE-LIF)

The CE-LIF analysis was carried out on an Agilent 7100 CE instrument with a Picometrics LIF detector (ex450nm) using a 50 µm ID PVA capillary with 65 cm total length and 45 cm effective length. Samples were injected with 30 mbar for 10 s. As separation buffer 40 mM ε-aminocaproic acid (pH 4.5, acetic acid) with 0.02 % (w/v) hydroxypropylmethylcellulose was used with -20 kV voltages at 20°C.

2.3 Enzyme activity assay by CE-LIF

The activity of β3GlcNAcT for 5 mM acceptor substrate **1** was carried out at 30°C in 100 mM HEPES-NaOH, pH 7.5, 25 mM KCl containing 6.5 mM $MgCl_2$, 6.5 mM $MnCl_2$, 6.5 mM UDP-GlcNAc, 6 U alkaline phosphatase (AP) and 10 µg purified β3GlcNAcT in a final volume of 50 µL. The activity of the β3- and the β4GalT for 5 mM acceptor substrate **2** was determined under the same conditions with 6.5 mM UDP-Gal. The reaction was stopped at different time points by heating (95°C for 5 min). The supernatant was used for a fluorescent labeling step with 8-aminopyrene-1,3,6-trisulfonic acid (APTS) and a HILIC-SPE post derivatization sample clean-up according to Hennig et al. as briefly described in the corresponding

section.[7b] Samples were measured by CE-LIF. The volumetric activity (U/mL) was calculated from the linear slope area.

2.4 APTS labeling and HILIC-SPE clean-up

Sample labeling and clean-up was carried out as previously described previously.[7b] Briefly, 2 µL supernatant was mixed with 2 µL 20 mM APTS in 3.6 M citric acid and 2 µL 0.2 M 2-picoline borane complex in dimethyl sulfoxide (DMSO) and incubated over 16 h at 37°C in darkness. The reaction was stopped by adding 94 µL of an 80:20 % (v/v) ACN-water solution (ACN solution). The HILIC-SPE clean-up was performed in a 96-well format. 200 µL of 100 mg/mL Bio-Gel P-10 slurry solved in 70:20:10 % (V/V) water-EtOH-ACN was added in each well of an AcroPrep™ 96-well filter plate (0.45 µm, GHP Membrane, Pall Corporation). The solvent was removed by applying vacuum (< 5 inHg) to the plate. After washing three times with water and equilibrating with ACN solution the sample was applied on top of the Bio-Gel phase. After 5 min. shaking the solvent was removed by vacuum. A 100 mM triethylamine (pH 8.5, acetic acid) ACN solution (80:20 % v/v) was used to wash the samples five times followed by three additional washing steps with ACN solution. For the sample elution, 150 µL water (ultrapure type 1) was used after 5 min shaking (450 rpm) by vacuum three times. Samples were dried using a vacuum concentrator and solved in 10 -50 µL water.

2.5 Synthesis of LNnT oligomers (mainly GlcNAc terminated)

The synthesis of LNnT oligomers was carried out in a 10 mL scale over 48 h at 30°C. The reaction was buffered by 100 mM HEPES (pH 7.5) and contains 25 mM KCl, 15 mM MgCl$_2$, 15 MnCl$_2$, 15 mM UDP-Gal, 15 mM UDP-GlcNAc and 5 mM Lactose (Compound **1**). 150 mU β4GalT, 150 mU of the β3GlcNAcT and 20 U alkaline phosphatase (AP) were added to start the reaction. The reaction was stopped by denaturation of the enzymes by a 5 min. heat treatment at 95°C. Precipitated enzymes were removed by centrifugation for 5 min at 15.000 rpm.

7

2.6 Synthesis of LNnT oligomers (Gal terminated)

For the synthesis of β4Gal terminated LNnT oligomers, a 5 mM Lactose solution was treated as described for the GlcNAc terminated oligomers. After stopping the reaction by a heat treatment additional 150 mU β4GalT, 20 U AP and 5 mM UDP-Gal were added and incubated over 24 h at 30°C. The enzyme was removed as already described.

2.7 Synthesis of LNT oligomers (Gal terminated)

Synthesis of β3Gal terminated LNT oligomers was performed as described for the β4Gal terminated LNnT oligomers. A 5 mM Lactose solution was treated as described for the GlcNAc terminated oligomers. After stopping the reaction by a heat treatment 150 mU β3GalT, 20 U AP and 5 mM UDP-Gal was added and incubated for 48 h at 30°C. The enzyme was stopped and removed as described before.

2.8 xCGE-LIF

Sample preparation for xCGE-LIF based glycoanalysis was performed as previously described:[7b] oligosaccharides were first labeled with APTS (Sigma-Aldrich), and the excess label was then removed using polyacrylamide based stationary phase in hydrophilic-interaction-chromatography (HILIC) mode. xCGE-LIF measurement of purified fluorescently labeled oligosaccharides was performed as published.[15] Briefly, one µL of HILIC-SPE eluate (or previously diluted sample eluate) was mixed with one µL GeneScan™ 500 LIZ™ Size Standard (1:50 dilution in HiDi) and ten µL Hi-Di™ Formamide (both Life Technologies). The mixture was transferred to a 384-well plate, septa sealed and centrifuged for 1 min at 200 g to avoid air bubbles at the bottom of the wells. The xCGE-LIF measurement was performed on a modified 3130xl Genetic Analyzer, equipped with a 50 cm capillary array, filled with the POP-7™ polymer (Life Technologies). The samples were electrokinetically injected and analyzed with a running voltage of 15 kV. Data was collected for 60 min. For processing of xCGE-LIF generated data, as well as for normalization of migration times to an internal standard the software glyXtool™ (glyXera) was used. As a

result, out of an electropherogram an HMOS "fingerprint" is created: signal intensity in relative fluorescence units (RFU) on the y-axis is plotted over the normalized migration time in normalized migration time units (MTU) on the x-axis. Structure identification was performed based on the migration time matching to the commercially available oligosaccharide standards or the products of their exoglycosidase digests. Identification was further confirmed by means of spiking with authentic oligosaccharide standards into the samples.[7c]

2.9 HMOS Digest

We further validated our annotation (sequence, monosaccharide types and linkages) by extensive exoglycosidase digests and subsequent reevaluation of HMOS fingerprints by xCGE-LIF. Following specific enzymes were used: β(1-4) galactosidase (recombinant from *Bacteroides fragilis*, expressed in *E.coli*; New England Biolabs), β(1-4,6) galactosidase (from *Canavalia ensiformis* (Jack bean); Prozyme), β(1-3) galactosidase (recombinant from *Xanthomonas manihotis*, expressed in *E.coli*; New England Biolabs), β(1-2,3,4,6) *N*-Acetylglucosaminidase (recombinant from *Streptococcus pneumoniae*, expressed in *E.coli*; New England Biolabs). Exoglycosidase digestions were carried out at 37 °C in buffers and under conditions recommended by the suppliers of the enzymes. Samples were purified by HILIC-SPE prior to xCGE-LIF analysis. Specific activity and possible side activity was carefully tested for each exoglycosidase enzyme used.

2.10 Cotton HILIC-SPE clean-up

Purification of unlabeled oligosaccharides used for MALDI-TOF-MS analysis was performed by cotton hydrophilic interaction liquid chromatography (HILIC) solid phase extraction (SPE) as described previously.[16] Briefly, samples removed from -20 °C were allowed to return to room temperature before proceeding. One µL of sample was mixed with 40 µL of 85% ACN (Optima LC/MS Grade, Fisher Scientific) prior to purification. Twenty microliter pipet tips packed with cotton thread were conditioned and equilibrated by pipetting three

times 20 µL of Milli-Q water, followed by three times 20 µL of 85% ACN. The sample was loaded by pipetting 20 times into the reaction mixture. The tips were washed three times with 20 µL of 85% ACN 1% TFA (Thermo Scientific), and three times with 20 µL of 85% ACN. Subsequent elution was performed in 10 µL of Milli-Q water.

2.11 MALDI-TOF-MS

For MALDI-TOF-MS analysis, 0.5 µL of oligosaccharide sample purified by cotton HILIC-SPE was spotted on a MTP AnchorChip 384 BC MALDI target (Bruker Daltonics), mixed on plate with 0.5 µL 2,5-dihydroxybenzoic acid (5 mg/mL; Bruker Daltonics) 1 mM NaOH (Sigma-Aldrich) in 50% ACN (Optima LC/MS Grade, Fisher Scientific) and dried under cold air. Analyses were performed in positive-ion reflector mode on an UltraFlextreme MALDI-TOF-MS equipped with a Smartbeam-II laser and flexControl 3.3 software Build 108 (Bruker Daltonics). The instrument was calibrated using a dextran hydrolysate. A 25 kV acceleration voltage was applied after a 120 ns extraction delay. A mass window of m/z 320 to 3500 was used for oligosaccharide samples. For each spectrum, 20 000 laser shots were accumulated at a laser frequency of 1000 Hz, using a complete sample random walk with 200 shots per raster spot. High laser intensity was used for sample profiling, while the monoisotopic peak was still clearly defined for all detectable oligosaccharide masses. Data analysis was performed by flexAnalysis 3.3 software Build 80 (Bruker Daltonics).

3 Results

In order to synthesize linear, long-chained Lacto-*N*-neo-type and Lacto-*N*-biose-type oligomers, a one-pot cascade reaction with human β1,4-galactosyltransferase-1 (β4GalT) and a β1,3-*N*-acetylglucosaminyltransferase (β3GlcNAcT) from *Helicobacter pylori* was carried out (**Scheme 1**). We obtained mainly *N*-acetylglucosamine terminated uneven LNnT oligomers (**Figure 1 A**) besides partially converted lactose as analyzed by xCGE-LIF. Galactose terminated LNnT structures were present only in traces. These results confirmed

previous studies on the substrate preference of β3GlcNAcT for longer LacNAc (Galβ1-4GlcNAc) glycan repeats.[13, 17] After enzyme removal, human β4GalT and β1,3-galactosyltransferase (β3GalT) from *E.coli* were used for the synthesis of LNnT oligomers (**Figure 1 B**) or LNT oligomers (**Figure 1 C**), respectively.

While human β4GalT1 fully converts long-chained poly LacNAc structures (**Figure 1B**) we obtained with bacterial β3GalT a mixture of LNT type oligomers with remaining GlcNAc terminated- and LNnT type oligomers (**Figure 1 C**). We confirmed our previous studies that the bacterial enzyme is less specific for long poly-LacNAc glycan substrates.[12a, 14b] Quantification of relative product distribution and MALDI-TOF analyses shows that LNT and LNnT glycan structures with up to 14 to 15 sugar units are detectable (**Table S1, Figure S1-S3**). Moreover, even the challenging linkage isomer mixture of LacNAc type 1 and type 2 terminated glycans (**Figure 1 C**) was separated and detected by xCGE-LIF.

These synthesis strategies (**Figure 1 A-C**) gave a broad product distribution for LNnT and LNT based glycan as analyzed by xCGE-LIF (Supplementary Materials, **Table S1**).

3.1 Novel structure evaluation

xCGE-LIF has been demonstrated as versatile HTS method for glycan analysis of complex biological samples to decipher even minor changes in *N*-glycan/oligosaccharide fingerprints.[7c, 15, 18] We here utilize xCGE-LIF as an ideal method for high throughput monitoring of LNT and LNnT type oligomers from synthetic enzyme reactions. Identification and annotation of oligosaccharides were carried out based on our internal HMOS database (comparison of migration times). The structural assignment was additionally confirmed by spiking defined amounts of purchased HMOS standards to the unknown samples (data not shown). However, only oligosaccharides up to octaose were commercially available for spiking-in and database build-up. The identity of long linear LNTs and LNnTs produced in this study needed, therefore, to be further examined by exoglycosidase digests, cleaving specific monosaccharide residues from nonreducing termini (Supplementary Materials, **Figure DS1-3**). Samples were then re-analyzed by xCGE-LIF and, based on the characteristic peak shifts that depend on the type of terminal monosaccharide cleaved, its linkage and even position in an oligosaccharide, the

12

oligosaccharide structures could be identified. In this study we used β(1-2,3,4,6) *N*-Acetylglucosaminidase to trace GlcNAc terminated LNnT oligomers (**Figure 2**), β(1-4,6) galactosidase for all oligosaccharides with galactose at nonreducing termini (Supplementary Materials, **Figure DS1**), β(1-4) galactosidase more specifically for even LNnT oligomers (Supplementary Materials, **Figure DS2**), and β(1-3) galactosidase for even LNT oligomers (Supplementary Materials, **Figure DS3**). As pinpointed above, this high resolution method allows detecting even minor structural differences between oligosaccharides, like in the case of structures differing only in the linkage type of terminal galactose (LNT vs. LNnT; **Figure 1** (**C**)). As such, the method has been proven to be competitive to MS-based methods used for HMOS identification that generally only provide structural information on a compositional level (Supplementary Materials, **Figure S1-3**).

3.2 Detection of the larger low abundant oligosaccharides

Chemical synthesis of oligosaccharides beyond 18mer is very rare and extremely challenging, involving many complex synthesis reactions.[19] Therefore, we further investigated what is the longest oligosaccharide structure we can enzymatically synthesize and detect by xCGE-LIF. For this purpose, samples were concentrated 240 times and re-injected. Linear oligosaccharide chains up to 30mer could be detected. In detail, even numbered LNnTs 30mer, uneven numbered LNnTs 29mer, and even numbered LNTs 26mer were synthesized (**Figure 3**). Electrokinetic injection applied in xCGE-LIF is sensitive to the salt content. Sample concentration increased the salt content in the loading sample, which might be the limiting step and reason why we do not observe even longer oligosaccharides. Additional sample purification to reduce the salt content and pre-fractionation to reduce the amount of small linear HMOS could make even larger structures detectable. Nevertheless, this is still the largest linear HMOS like structure ever synthesized by any method.

13

4 Discussion

In natural samples linear LNTs and LNnT HMOS structures up to octaose oligomers were identified.[3a, 3b, 20] Long chained LNT and LNnT oligomers are basically lactose-based poly LacNAc oligomers. Their chemical synthesis is practically limited in the accessible length by relatively low overall yields and high technical effort due to several tedious synthetic protection and deprotection steps.[21] Nevertheless, chemical poly LacNAc syntheses up to octaose oligomers were reached in previous studies.[21a] Also, HMOS LNT and LNnT tetra-, LNnT hexa- and octaose oligomers were already chemically synthesized.[22] The development of automated solid-phase oligosaccharide synthesis has been a milestone for efficient and fast chemical oligosaccharide synthesis.[23] Automated solid-phase based synthesis of HMOS like structures has been demonstrated for Lewis X-Lewis Y (Le[X]-Le[Y]) nonasaccharide structures.[24] In contrast, *Leloir*-glycosyltransferase utilized for LNT HMOS [4a] and poly LacNAc oligomer synthesis reaches quantitative yields.[12a, 13, 14b, 14c, 25] The enzymatic synthesis of long chained poly LacNAc oligomers has been carried out in previous studies up to dodecaoses.[13, 26] Nevertheless, the enzymatic *in vitro* synthesis of linear LNT and LNnT HMOS was demonstrated only for tetra – and hexaose structures before.[3a, 3j, 4a, 12b, 27] Whole cell biotransformation, which can reach multigram scales easily, seems to be actually very limited to reach mainly LNnT tetraoses, low levels of hexa- and octaoses.[9, 28] With our synthesis strategy we achieved large linear HMOS like structures far beyond so far published. We applied optimized conditions published previously by Rech et al. where the influence of ratios of enzymes, donor and acceptor substrates were thoroughly investigated.[13] Dodecaose poly-LacNAc structures could be detected by HPLC-MS.[13] However, with the high resolution and highly sensitive xCGE-LIF based glycoanalysis, even longer glycan structures could be observed.

14

Besides pushing the glycoanalysis to new limits, we demonstrate here for the first time β4GalT1 activity with poly-LacNAc substrates up to 30 sugar units. The human β4GalT1 is of special interest, since it was isolated from human milk samples indicating its role in the natural HMOS synthesis.[29] We may conclude from the catalytic efficiency of human β4GalT1 that some of the novel long chained HMOS like structures may occur also in traces in nature.[30]

Concluding, we here demonstrate for the first time the synthesis of linear LNnT oligomers up to 30 sugar units and LNT oligomers up to 26 sugar units by the use of recombinant *Leloir*-glycosyltransferases. Compared to published data, we obtain a length far beyond the known, natural occurring octaoses for linear LNnT and LNT oligomers.[3a, 3b, 20] Since already the abundance of the known linear long chained HMOS in human milk is relatively low, the chance to identify putative bigger structures occurring in traces without any references, is imaginable low. The semi one-pot strategy was key to achieve synthesis of such novel tailored HMOS reference standards with defined broad structural distribution. Combination with a high performance high throughput system is a promising strategy to identify and elucidate even very rare glycan structures in natural samples. The presented xCGE-LIF system allows a database annotation (structure identification via migration time matching) and thereby enables the possibility for future screening or fingerprinting projects to identify even rare novel structures in native samples much easier. In this connection, we rate the demonstrated principle to provide enzymatic tailored reference standards as a very general and straightforward strategy.

Acknowledgement

Conflict of interest

The authors declare no financial or commercial conflict of interest.

5 References

[1]   a) L. Bode, *Glycobiology* **2012**, *22*, 1147-1162; b) J. B. German, C. J. Dillard, R. E. Ward, *Curr. Opin. Clin. Nutr. Metab. Care* **2002**, *5*, 653-658.

[2]   a) G. V. Coppa, O. Gabrielli, P. Pierani, C. Catassi, A. Carlucci, P. L. Giorgi, *Pediatrics* **1993**, *91*, 637-641; b) C. Kunz, S. Rudloff, W. Schad, D. Braun, *Br. J. Nutr.* **1999**, *82*, 391-399; c) L. R. Ruhaak, C. B. Lebrilla, *BMB Rep.* **2012**, *45*, 442-451; d) D. Viverge, L. Grimmonprez, G. Cassanas, L. Bardet, M. Solere, *J. Pediatr. Gastroenterol. Nutr.* **1990**, *11*, 361-364.

[3]   a) X. Chen, *Adv Carbohydr Chem Biochem* **2015**, *72*, 113-190; b) A. Kobata, *Proc. Jpn. Acad., Ser. B* **2010**, *86*, 731-747; c) A. Kobata, K. Yamashita, Y. Tachibana, *Methods Enzymol.* **1978**, *50*, 216-220; d) C. Kunz, S. Rudloff, *Acta Paediatr.* **1993**, *82*, 903-912; e) C. Kunz, S. Rudloff, W. Baier, N. Klein, S. Strobel, *Annu. Rev. Nutr.* **2000**, *20*, 699-722; f) A. Lucas, T. J. Cole, *The Lancet* **1990**, *336*, 1519-1523; g) R. P. McEver, *Ann. N. Y. Acad. Sci.* **1994**, *714*, 185-189; h) C. C. Nwosu, D. L. Aldredge, H. Lee, L. A. Lerno, A. M. Zivkovic, J. B. German, C. B. Lebrilla, *J. Proteome Res.* **2012**, *11*, 2912-2924; i) R. Peterson, W. Y. Cheah, J. Grinyer, N. Packer, *Glycobiology* **2013**, *23*, 1425-1438; j) G. A. Sprenger, F. Baumgärtner, C. Albermann, *J. Biotechnol.* **2017**, *258*, 79-91.

[4]   a) N. S. Han, T.-J. Kim, Y.-C. Park, J. Kim, J.-H. Seo, *Biotechnol. Adv.* **2012**, *30*, 1268-1278; b) M. R. Ninonuevo, Y. Park, H. Yin, J. Zhang, R. E. Ward, B. H. Clowers, J. B. German, S. L. Freeman, K. Killeen, R. Grimm, C. B. Lebrilla, *J. Agric. Food Chem.* **2006**, *54*, 7471-7480; c) S. Thurl, M. Munzert, G. Boehm, C. Matthews, B. Stahl, *Nutr. Rev.* **2017**, *75*, 920-933.

[5]   a) C. Marx, R. Bridge, A. K. Wolf, W. Rich, J. H. Kim, L. Bode, *Journal of human lactation : official journal of International Lactation Consultant Association* **2014**, *30*, 54-61; b) G. Xu, J. C. Davis, E. Goonatilleke, J. T. Smilowitz, J. B. German, C. B. Lebrilla, *J. Nutr.* **2017**, *147*, 117-124.

[6]   a) A. Guttman, F.-T. A. Chen, R. A. Evangelista, N. Cooke, *Anal. Biochem.* **1996**, *233*, 234-242; b) A. Guttman, T. Pritchett, *Electrophoresis* **1995**, *16*, 1906-1911; c) J. Liu, O. Shirota, M. V. Novotny, *Anal. Chem.* **1992**, *64*, 973-975.

[7]   a) D. Blank, H. Geyer, K. Maass, K. Yamashita, R. Geyer, *Anal. Biochem.* **2012**, *421*, 680-690; b) R. Hennig, E. Rapp, R. Kottler, S. Cajic, M. Borowiak, U. Reichl, *Methods Mol. Biol.* **2015**, *1331*, 123-143; c) R. Kottler, M. Mank, R. Hennig, B. Muller-Werner, B. Stahl, U. Reichl, E. Rapp, *Electrophoresis* **2013**, *34*, 2323-2336.

[8]   a) A. R. Prudden, L. Liu, C. J. Capicciotti, M. A. Wolfert, S. Wang, Z. Gao, L. Meng, K. W. Moremen, G.-J. Boons, *Proc. Natl. Acad. Sci.* **2017**, *114*, 6954-6959; b) H. Yu, Y. Li, Z. Wu, L. Li, J. Zeng, C. Zhao, Y. Wu, N. Tasnima, J. Wang, H. Liu, M. R. Gadi, W. Guan, P. G. Wang, X. Chen, *Chem. Commun.* **2017**, *53*, 11012-11015; c) H. Zhu, Z. Wu, M. R. Gadi, S. Wang, Y. Guo, G. Edmunds, W. Guan, J. Fang, *Bioorg. Med. Chem. Lett.* **2017**, *27*, 4285-4287.

[9]   F. Baumgartner, L. Jurzitza, J. Conrad, U. Beifuss, G. A. Sprenger, C. Albermann, *Bioorg. Med. Chem.* **2015**, *23*, 6799-6806.

[10]   B. Sauerzapfe, D.-J. Namdjou, T. Schumacher, N. Linden, K. Krenek, V. Kren, L. Elling, *J. Mol. Catal. B: Enzym.* **2008**, *50*, 128-140.

[11] S. M. Logan, E. Altman, O. Mykytczuk, J.-R. Brisson, V. Chandan, F. S. Michael, A. Masson, S. Leclerc, K. Hiratsuka, N. Smirnova, J. Li, Y. Wu, W. W. Wakarchuk, *Glycobiology* **2005**, *15*, 721-733.

[12] a) T. Fischöder, D. Laaf, C. Dey, L. Elling, *Molecules (Basel, Switzerland)* **2017**, *22*, 1320; b) X.-w. Liu, C. Xia, L. Li, W.-y. Guan, N. Pettit, H.-c. Zhang, M. Chen, P. G. Wang, *Bioorg. Med. Chem.* **2009**, *17*, 4910-4915.

[13] C. Rech, R. R. Rosencrantz, K. Křenek, H. Pelantová, P. Bojarová, C. E. Römer, F.-G. Hanisch, V. Křen, L. Elling, *Adv. Synth. Catal.* **2011**, *353*, 2492-2500.

[14] a) S. Böcker, D. Laaf, L. Elling, *Biomolecules* **2015**, *5*, 1671-1696; b) D. Laaf, P. Bojarová, H. Pelantová, V. Křen, L. Elling, *Bioconjugate Chem.* **2017**, *28*, 2832-2840; c) B. Sauerzapfe, K. Křenek, J. Schmiedel, W. W. Wakarchuk, H. Pelantová, V. Křen, L. Elling, *Glycoconjugate J.* **2009**, *26*, 141-159; d) A. Šimonová, C. E. Kupper, S. Böcker, A. Müller, K. Hofbauerová, H. Pelantová, L. Elling, V. Křen, P. Bojarová, *J. Mol. Catal. B: Enzym.* **2014**, *101*, 47-55.

[15] R. Hennig, S. Cajic, M. Borowiak, M. Hoffmann, R. Kottler, U. Reichl, E. Rapp, *Biochim. Biophys. Acta* **2016**, *1860*, 1728-1738.

[16] K. R. Reiding, D. Blank, D. M. Kuijper, A. M. Deelder, M. Wuhrer, *Anal. Chem.* **2014**, *86*, 5784-5793.

[17] Y. Li, M. Xue, X. Sheng, H. Yu, J. Zeng, V. Thon, Y. Chen, M. M. Muthana, P. G. Wang, X. Chen, *Bioorg. Med. Chem.* **2016**, *24*, 1696-1705.

[18] a) M. Abeln, K. M. Borst, S. Cajic, H. Thiesler, E. Kats, I. Albers, M. Kuhn, V. Kaever, E. Rapp, A. Münster-Kühnel, B. Weinhold, *ChemBioChem* **2017**, *18*, 1305-1316; b) C. B. Karsten, F. F. Buettner, S. Cajic, I. Nehlmeier, B. Neumann, A. Klippert, U. Sauermann, U. Reichl, R. Gerardy-Schahn, E. Rapp, C. Stahl-Hennig, S. Pohlmann, *J. Virol.* **2015**, *89*, 11727-11733; c) S. A. Konze, S. Cajic, A. Oberbeck, R. Hennig, A. Pich, E. Rapp, F. F. R. Buettner, *ChemBioChem* **2017**, *18*, 1317-1331; d) J. Schwarzer, E. Rapp, U. Reichl, *Electrophoresis* **2008**, *29*, 4203-4214; e) C. T. Thiesler, S. Cajic, D. Hoffmann, C. Thiel, L. van Diepen, R. Hennig, M. Sgodda, R. Weibetamann, U. Reichl, D. Steinemann, U. Diekmann, N. M. Huber, A. Oberbeck, T. Cantz, A. W. Kuss, C. Korner, A. Schambach, E. Rapp, F. F. Buettner, *Mol. Cell. Proteomics* **2016**, *15*, 1435-1452.

[19] a) O. Calin, S. Eller, P. H. Seeberger, *Angew. Chem., Int. Ed.* **2013**, *52*, 5862-5865; b) H. S. Hahm, M. Hurevich, P. H. Seeberger, *Nat. Commun.* **2016**, *7*, 12482; c) K. Naresh, F. Schumacher, H. S. Hahm, P. H. Seeberger, *Chem. Commun.* **2017**, *53*, 9085-9088.

[20] a) J. Amano, M. Osanai, T. Orita, D. Sugahara, K. Osumi, *Glycobiology* **2009**, *19*, 601-614; b) D. Blank, S. Gebhardt, K. Maass, G. Lochnit, V. Dotz, J. Blank, R. Geyer, C. Kunz, *Anal. Bioanal. Chem.* **2011**, *401*, 2495-2510; c) W. Chai, V. E. Piskarev, Y. Zhang, A. M. Lawson, H. Kogelberg, *Arch. Biochem. Biophys.* **2005**, *434*, 116-127; d) S. Haeuw-Fievre, J. M. Wieruszeski, Y. Plancke, J. C. Michalski, J. Montreuil, G. Strecker, *Eur. J. Biochem.* **1993**, *215*, 361-371; e) H. Kogelberg, V. E. Piskarev, Y. Zhang, A. M. Lawson, W. Chai, *Eur. J. Biochem.*

2004, *271*, 1172-1186; f) K. Yamashita, Y. Tachibana, A. Kobata, *J. Biol. Chem.* **1977**, *252*, 5408-5411.

[21] a) J. Alais, A. Veyrieres, *Carbohydr. Res.* **1990**, *207*, 11-31; b) T. K. Mong, C. Y. Huang, C. H. Wong, *J. Org. Chem.* **2003**, *68*, 2135-2142.

[22] a) M. R. E. Aly, E. S. I. Ibrahim, E. El-Ashry, R. R. Schmidt, *Synthesis of lacto-N-neohexaose and lacto-N-neooctaose using the dimethylmaleoyl moiety as an amino protective group*, **2000**; b) M. R. E. Aly, E.-S. I. Ibrahim, E.-S. H. El Ashry, R. R. Schmidt, *Carbohydr. Res.* **1999**, *316*, 121-132; c) T. Takamura, T. Chiba, H. Ishihara, S. Tejima, *Chem. Pharm. Bull.* **1980**, *28*, 1804-1809.

[23] O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Science* **2001**, *291*, 1523-1527.

[24] a) K. Routenberg Love, P. H. Seeberger, *Angew. Chem., Int. Ed.* **2004**, *43*, 602-605; b) P. H. Seeberger, D. B. Werz, *Nat. Rev. Drug Discovery* **2005**, *4*, 751.

[25] a) S. Di Virgilio, J. Glushka, K. Moremen, M. Pierce, *Glycobiology* **1999**, *9*, 353-364; b) K. I. Hidari, T. Murata, K. Yoshida, Y. Takahashi, Y. H. Minamijima, Y. Miwa, S. Adachi, M. Ogata, T. Usui, Y. Suzuki, T. Suzuki, *Glycobiology* **2008**, *18*, 779-788.

[26] K. Naruchi, T. Hamamoto, M. Kurogochi, H. Hinou, H. Shimizu, T. Matsushita, N. Fujitani, H. Kondo, S. Nishimura, *J. Org. Chem.* **2006**, *71*, 9609-9621.

[27] a) T. Miyazaki, T. Sato, K. Furukawa, K. Ajisaka, *Methods Enzymol.* **2010**, *480*, 511-524; b) T. Murata, T. Inukai, M. Suzuki, M. Yamagishi, A. T. Usui,

Glycoconjugate J. **1999**, *16*, 189-195; c) L. Renaudie, R. Daniellou, C. Augé, C. Le Narvor, *Carbohydr. Res.* **2004**, *339*, 693-698.

[28]   a) F. Baumgartner, J. Conrad, G. A. Sprenger, C. Albermann, *ChemBioChem* **2014**, *15*, 1896-1900; b) F. Baumgartner, L. Seitz, G. A. Sprenger, C. Albermann, *Microb. Cell Fact.* **2013**, *12*, 40; c) C. Dumon, B. Priem, S. L. Martin, A. Heyraud, C. Bosso, E. Samain, *Glycoconjugate J.* **2001**, *18*, 465-474; d) B. Priem, M. Gilbert, W. W. Wakarchuk, A. Heyraud, E. Samain, *Glycobiology* **2002**, *12*, 235-240.

[29]   a) N. L. Shaper, M. Charron, N.-W. Lo, J. H. Shaper, *Journal of Mammary Gland Biology and Neoplasia* **1998**, *3*, 315-324; b) M. Ujita, A. K. Misra, J. McAuliffe, O. Hindsgaul, M. Fukuda, *J. Biol. Chem.* **2000**, *275*, 15868-15875.

[30]   a) H. E. Appert, T. J. Rutherford, G. E. Tarr, N. R. Thomford, D. J. McCorquodale, *Biochem. Biophys. Res. Commun.* **1986**, *138*, 224-229; b) T. Endo, J. Amano, E. G. Berger, A. Kobata, *Carbohydr. Res.* **1986**, *150*, 241-263.
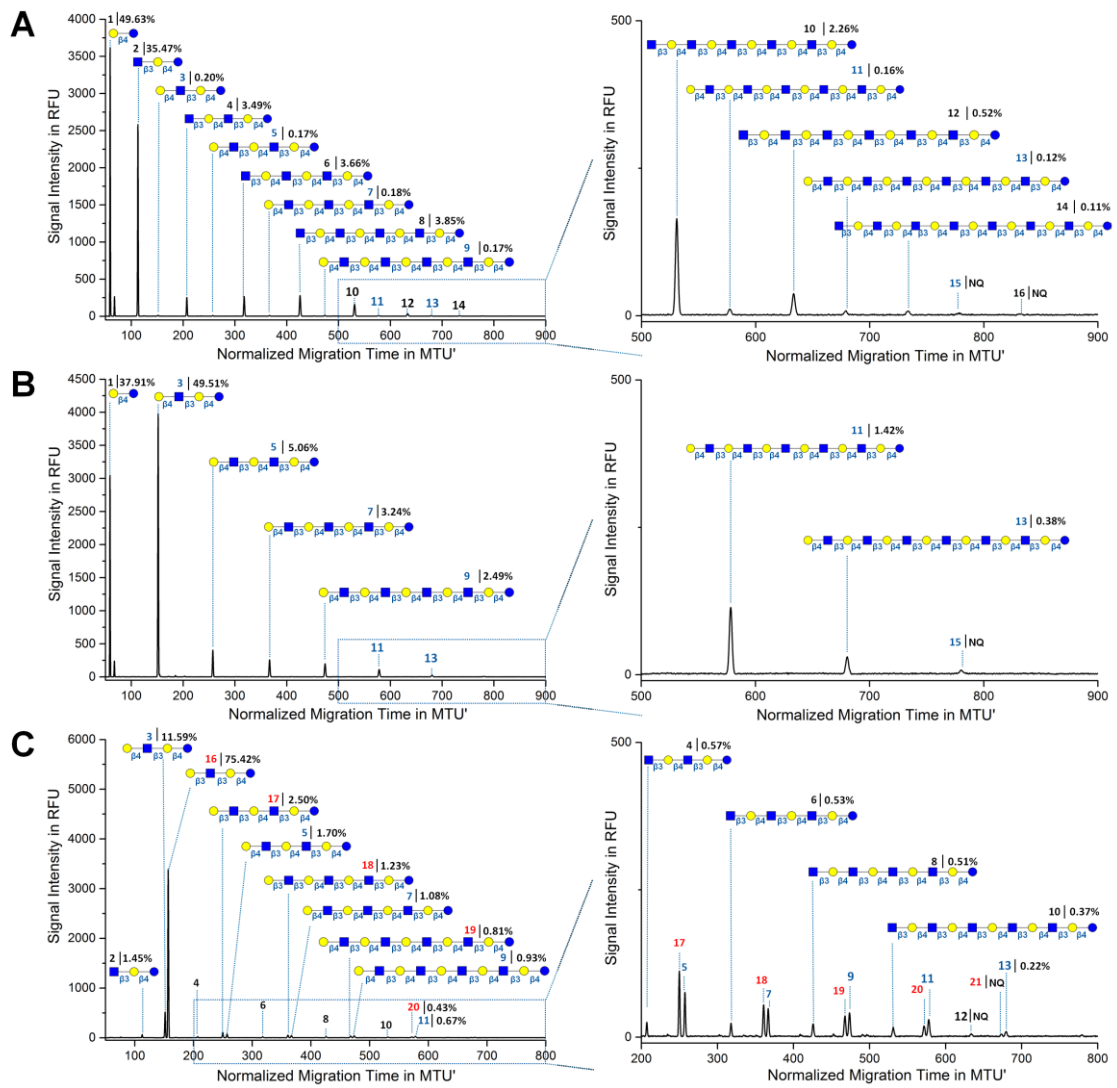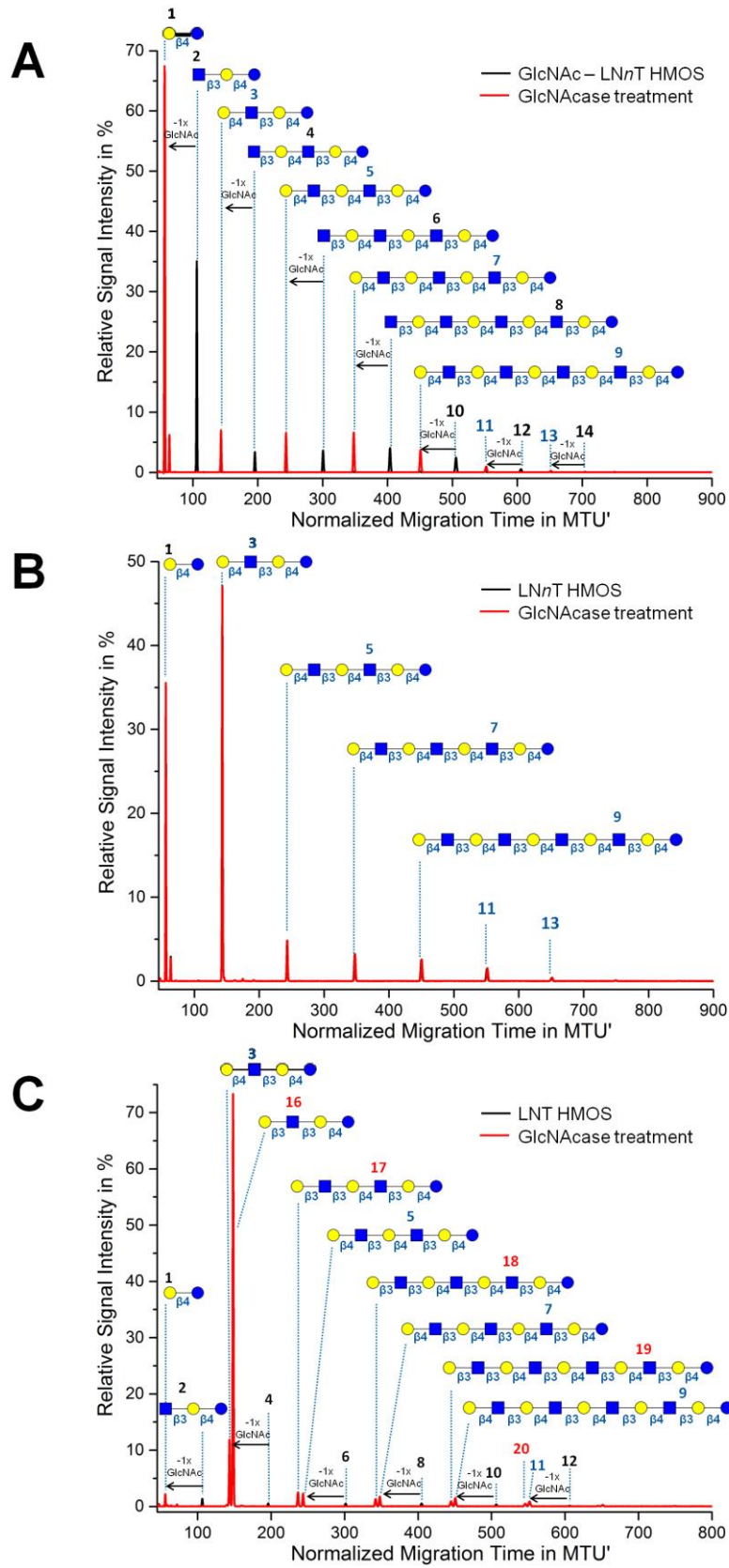
Figures
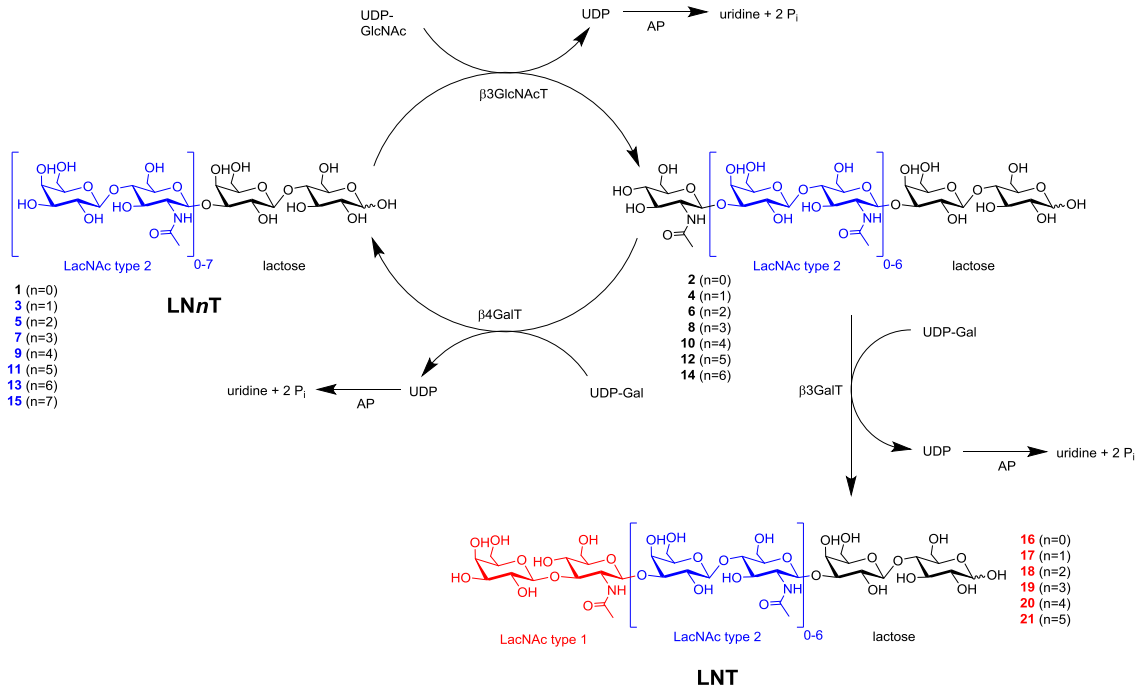
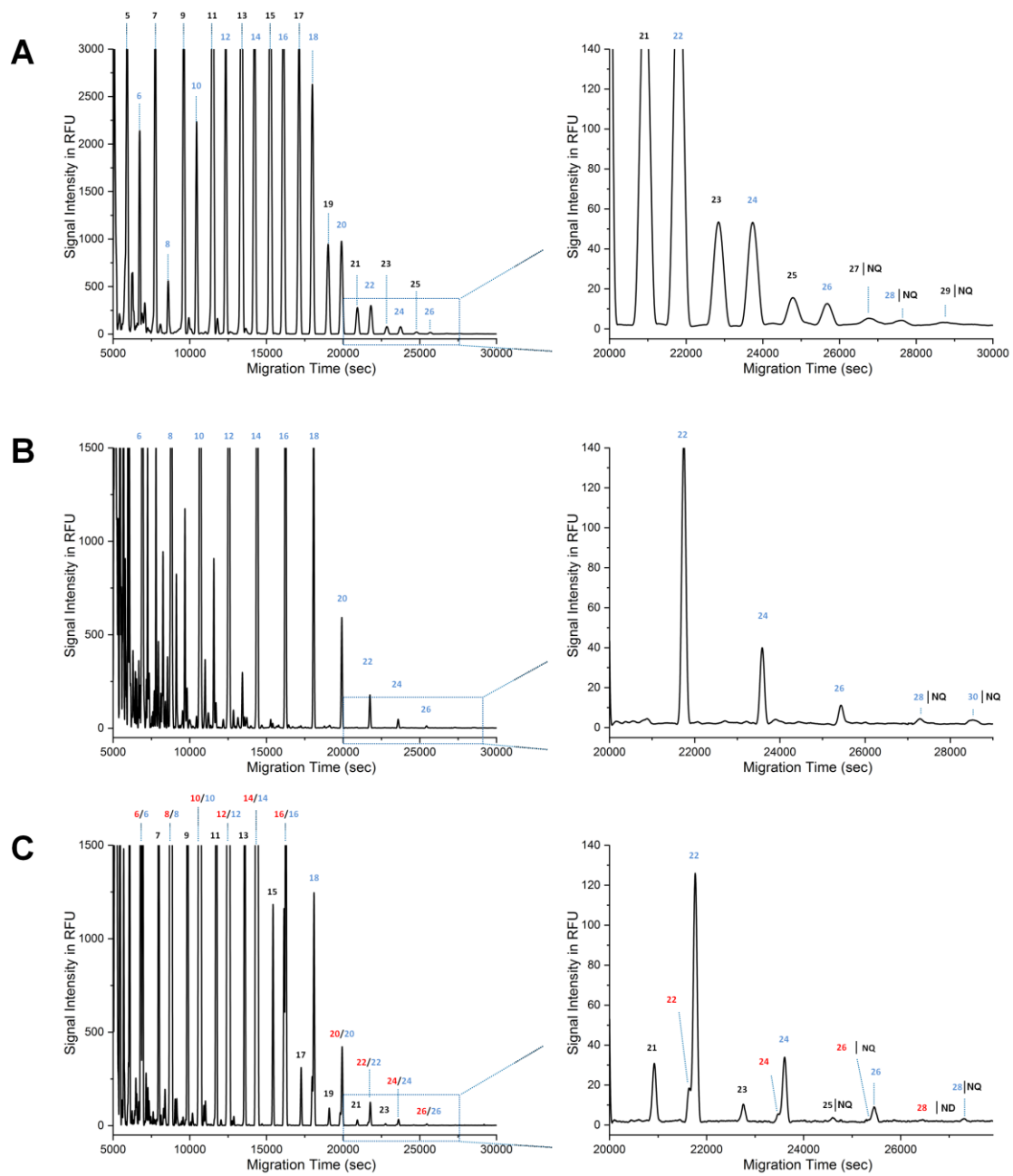
Scheme 1.

Figure 1.

Figure 2.

Figure 3.

Figure legends

Scheme 1. Enzymatic, semi-sequential one-pot synthesis of human milk oligosaccharides **2-21** by utilizing human β1,4-galactosyltransferase-1 (β4GalT), β1,3-*N*-acetylglucosaminyltransferase from *Helicobacter pylori* (β3GlcNAcT), *E.coli* β1,3-galactosyltransferase (β3GalT), and alkaline phosphatase (AP).

Figure 1. Semi-sequential one-pot synthesis of GlcNAc terminated LNnT oligomers (**A**), even LNnT oligomers (**B**) and LNT oligomers (**C**) analyzed by xCGE-LIF. Peaks with a signal to noise ratio 3 < S/N < 9 were defined as not quantifiable (NQ).

Figure 2. β(1-2,3,4,6) specific *N*-Acetylglucosaminidase treatment (*Streptococcus pneumoniae*, NEB).

Figure 3. Semi-sequential one-pot synthesis of GlcNAc terminated LNnT (black) oligomers (**A**), even LNnT (blue) oligomers (**B**) and LNT (red) oligomers (**C**) analyzed by xCGE-LIF concentrated 240 times and re-injected. Here, differing from the previous figures, peaks were marked regarding to their sugar units and not to their compound numbers. Peaks with S/N ≤ 3 were marked as not detectable (ND). Peaks with a signal to noise ratio 3 < S/N < 9 were defined as not quantifiable (NQ).