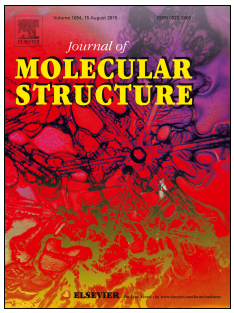
Accepted Manuscript



Isolation, structure elucidation and DFT study on two novel oligosaccharides from yak milk

Meenakshi Singh, Alok Kumar, Gaurav Srivastava, Desh Deepak, M.P.V.V. Singh

PII: S0022-2860(16)30261-7

DOI: 10.1016/j.molstruc.2016.03.065

Reference: MOLSTR 22376

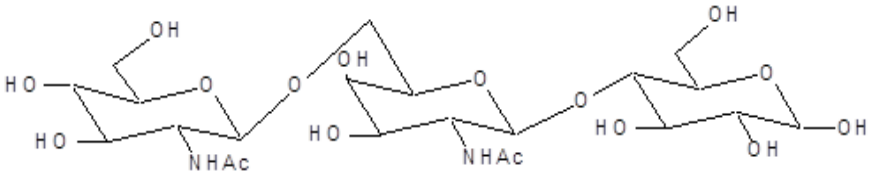
To appear in: *Journal of Molecular Structure*

Received Date: 31 December 2014

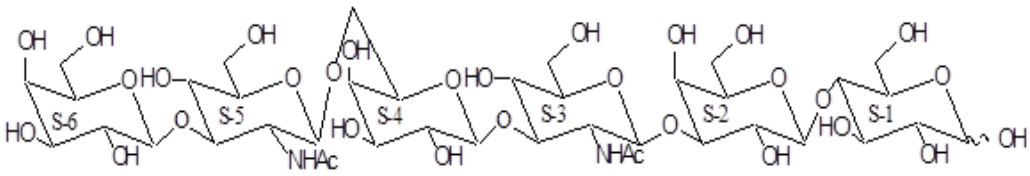
Revised Date: 15 March 2016

Accepted Date: 20 March 2016

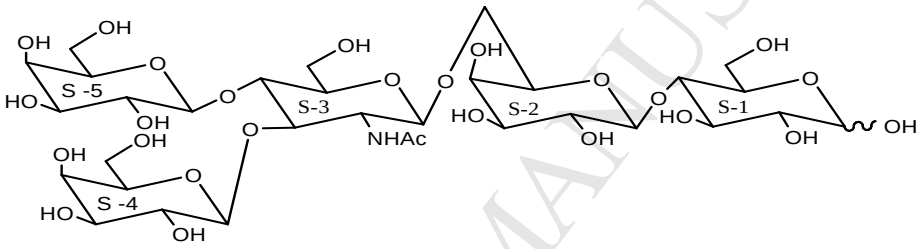
Please cite this article as: M. Singh, A. Kumar, G. Srivastava, D. Deepak, M.P.V.V. Singh, Isolation, structure elucidation and DFT study on two novel oligosaccharides from yak milk, *Journal of Molecular Structure* (2016), doi: 10.1016/j.molstruc.2016.03.065.



(Bosiose)



(Bovisose)



(Yakalose)

Isolation, Structure elucidation and DFT Study on Two novel oligosaccharides from yak milk

Meenakshi Singh[a, b], *Alok Kumar*[b], *Gaurav Srivastava*[a], *Desh Deepak*[a]* *and M.P.V.V. Singh*[c]

[a]*Department of Chemistry, University of Lucknow, Lucknow 226007, India*

[b]*Department of Chemistry, BBD, University, Faizabad Road, Lucknow, Uttar Pradesh 226028*

[c]*Department of Botany, Shri Jai Narain Post Graduate (KKC) College, Lucknow 226001*

Corresponding author, Email: singhmeenakshi144@gmail.com

Abstract

Two novel oligosaccharides were isolated from yak milk. The milk was processed by the method of Kobata and Ginsberg involving deproteination, centrifugation and lyophilization followed by gel filtrate chromatography acetylation and silica gel column chromatography of derivatized oligosaccharides while their homogeneity was confirmed by HPLC. The structures of these isolated oligosaccharides were elucidated by chemical transformation, chemical degradation, ^{1}H, ^{13}C NMR, 2D NMR (COSY, TOCSY and HSQC) and mass spectrometry. The geometry of compound A (Bosiose) and B (Bovisose) have been optimized at B3LYP method and 6-311+G(d,p) basis set. The difference between the energies of A and B is 1.269 a.u. or 796.309 kcal/mol.

Keywords-Oligosaccharides, Yak milk, Bosiose (A), Bovisose (B), HPLC

1. Introduction

Milk has co-evolved with mammals and mankind to nourish their offspring and is a biological fluid of unique complexity and richness [1]. It contains all necessary nutrients for the growth and development of the newborn [2]. The development of carbohydrates and glycoconjugates as therapeutics continues to gain interest when biological relevance of these biopolymers were further revealed and understood [3]. Oligosaccharides have established themselves as an effective class of organic biomolecules impacting various physiological and pathological processes such as molecular recognition, signal transaction, differentiation and developmental events and exhibit varied biological activities such as antitumor [4], immunostimulant [5], anticancer [6], anticomplementry, anticoagulant, antiinflamatory, hypoglycemic, antiviral & immunological activities [7].

Oligosaccharides isolated from various milk sources are categorized in two classes i.e. sialylated & nonsialylated. Both these class of oligosaccharide have been tested for their varied biological activities [8]. Human milk oligosaccharides (HMOs) are important parts of these functional ingredients [9, 10]. Numbers of biologically active oligosaccharides have been isolated from human, buffalo, donkey, cow, mare, sheep and goat milk [11]. The milk oligosaccharides recognize cancer associated antigens, used as antimicrobial agents; tumor associated antigens and has physiological significance in infants [12]. Buffalo and Donkey milk oligosaccharides have shown promising immunostimulant activity [1]. Glycosylation of human milk lactoferrin exhibits dynamic changes during early lactation enhancing its role in pathogenic bacteria-host interactions. Oligosaccharides are the third most abundant component in human milk [13, 8]. Human milk oligosaccharides (HMOs) containing α1, 2-linked fucose inhibits the stable toxin-producing Escherichia coli *in vitro* and its toxin-induced secretary diarrhea *in vitro* and

in vivo [14, 15]. Recently it was found that the content of 2-linked fucosyl oligosaccharides in human milk is significantly associated with lower risk of diarrhea in breastfed infants, suggesting a major role for these oligosaccharides in immunity [16]. Functions of HMOs are not fully understood but it has been postulated that HMOs plays an important role in the development of immune system, the prevention of pathogenic infection and, in the modulation of infant GI tract to bifidogenic microbiota [17-19]. Milk is not only food for neonates; it also contains many nutritional as well as medicinal values [20, 21].

People living on the Tibetan Plateau rely for survival upon the yak, the region's native cattle [22]. The yak milk was used in Tibetan medicinal system for the treatment of enema therapy, in which yak milk was used as enema solution along with other drugs [23]. It has been seen in the literature survey that the amino acid, calcium and vitamin A in yak milk are comparatively higher than in cow's milk and yak milk caseins could become a resource to generate antihypertensive peptides and be used as multifunctional active ingredients for many value-added functional foods as well as a traditional food protein [24-26]. The China Nutrition society claimed that the amino acid, calcium and vitamin A in yak milk are comparatively higher than in cow's milk [27].

In view of above facts and observations, therefore isolation, structure elucidation, chemical degradation, chemical transformation and spectroscopic characterization of new three oligosaccharides from yak (*Bos grunniens*) milk have been presented in the present paper. The completed analysis of structure has been evaluated using spectroscopic technique (2D (COSY, TOCSY, and HSQC) and ^{13}C NMR data. Other techniques like deacetylation, methylation, hydrolysis, chemical degradation and ESI MS (mass spectrometry) have also been used for structure elucidation of

oligosaccharides. Furthermore, quantum chemical calculation has been performed to determine the lower geometry structure and stability of the isolated and characterized compounds.

2 Theoretical studies

All computations were performed using the Gaussian 09 program package [28]. The quantum chemical calculation have been performed on B3LYP functional and 6-311+G(d,p) basis set. Geometries of compound A and B have been first optimized and the presence of positive wavenumbers values for all the optimized geometry indicates stability of the compounds [29].

2.1 Experimental

General procedures

Optical rotations were measured using an AA-5 automatic polarimeter in a 1 dm tube for water solutions whose concentrations are expressed in g/100mL.

The ^{1}H and ^{13}C NMR spectra were recorded in $CDCl_3$ and D_2O on Bruker DRX-300 spectrometer at 300 and 400 MHz using TMS as an internal reference.

The DART Mass spectra were recorded on a MICROMASS QUATTRO II triple quadruple mass spectrometer. The C, H and N analyses were recorded on CARDO-ELBA 1108 elemental analyzer.

The detection of new spots for carbohydrates was monitored by thin layer chromatography (TLC) using Silica Gel 60 F254 plates. TLC plates were visualized by exposure to them in with 50% aq. H_2SO_4 reagent.

A new spot for carbohydrates was also monitored by paper chromatography (PC) with acetyl acetone and p-dimethyl amino benzaldehyde reagents. PC was

performed on Whatman No.1 filter paper using ethylacetate-pyridine (2:1) saturated with H_2O as solvent system. Authentic samples of TLC and PC were same and purchased from Aldrich Chemicals (glucosamine (GlcN), galactosamine (GalN), glucose (Glc), galactose (Gal), N-acetylglucosamine (GlcNAc) and N-acetylgalactosamine (GalNAc)).

2.2. Isolation of yak milk oligosaccharides by Kobata and Ginsberg method

10 liter milk was collected from a yak and was stored at -20°C until use. The milk was processed by the method of Kobata and Ginsberg [29]. It was centrifuged for 15 min. at 6500 rpm at -4°C. The solidified lipid layer was removed by filtration through glass wool column in cold atmospheric condition (0°C). Ethanol was added to the clear filtrate (supernatant) to a final concentration of 68% for precipitating out the lactose and proteins and the resulting solution was left overnight at 0°C. The white precipitate of lactose and protein was formed and removed by centrifugation for 15 min. at 6500 rpm at -2°C and washed twice with 68% ethanol. Further for complete remove of remaining lactose the supernatant was passed through a microfilter (0.24 μm) and lyophilized to get the crude oligosaccharide mixture (12.0 grams).

The lyophilized material responded positively to Morgan-Elson test [30] and thiobarbituric-acid assay suggesting the presence of N-acetyl sugars and sialic acid in oligosaccharide mixture. The detailed of the test has been given in *supplementary file 1*.

2.3. Yak milk oligosaccharide mixture chromatographed over Sephadex G-25 (1.6x40 cm) column

The gel filtration was performed by Sephadex G-25 chromatography of crude yak milk oligosaccharide mixture. The oligosaccharide mixture was packed in a column (1.6 x 40 cm) (void volume = 25 mL) equilibrated with glass double distilled water and

left for 10-12 h to settle down. The 18.0 gm of the crude oligosaccharide mixture was obtained from 10 liter milk. From which 200 mg material in 150 mL distilled water was applied onto a Sephadex G-25 column and was eluted for separation of protein and glycoprotein from oligosaccharide mixture (low molecular weight component). The flow rate was adjusted to 3 mL / min. Then the eluent was reapplied on the same Sephadex G25 column. The process was repeated twice for further purification. 1 ml fractions were collected in all eluted fractions and tested for the presence of neutral sugars by phenol-sulphuric acid test [31]. Same above procedure (approximately 88 x 2 times) was applied for the remaining 16.0 gm of the crude oligosaccharide mixture to obtain 12.0 gm of pool I and pool II. The sephadex G-25 chromatography of yak milk oligosaccharide mixture which was monitored by UV (absorbance 280 nm) spectrophotometry showed four peaks i.e. I, II, III and IV. A substantial amount of proteins, glycoproteins and serum albumin were eluted in the void volume which was confirmed by positive coloration with p-dimethyl amino benzaldehyde reagent and phenol-sulphuric acid reagent. Fractions under peaks II and III gave a positive phenol-sulphuric acid test [31] which showed the presence of oligosaccharide mixture in yak milk. They were pooled together and lyophilized.

2.4. Confirmation of homogeneity of yak milk oligosaccharide by reverse phase HPLC

Fraction II and III obtained from Sephadex G-25 column, containing oligosaccharide mixture were qualitatively analyzed by reverse phase HPLC. The HPLC system was equipped with Perkin-Elmer 250 solvent delivering system, 235 diode array detector and G.P. 100 printer plotter. The column used for this purpose was C18 Purosphere 25 cm x 0.4 cm x 5 μm. A binary gradient system of acetonitrile: 0.5% trifluro acetic acid (5:95) in distilled water to CH_3CN: 0.5% TFA (60:40) within 25 mm at a flow rate of 1 mL / min was used. The eluants were detected at 220 nm. Eleven

peaks were noticed in the sample (pooled fractions II and III) at the varied retention times from 2.272 min. to 5.717 min. The peaks were numbered in increasing order of retention time i.e. 2.272 (R_1), 3.115 (R_2), 3.637 (R_3), 3.968 (R_4), 4.331 (R_5), 4.693 (R_6), 5.077 (R_7), 5.227 (R_8), 5.419 (R_9), 5.557 (R_{10}), 5.717 (R_{11}).

2.5. Acetylation of oligosaccharide mixture

The pooled fraction (II and III) (12.0 g) which gave positive phenol-sulphuric acid test [31] was acetylated with pyridine (15 mL) and acetic anhydride (15 mL) at 60°C and the solution was stirred overnight. The mixture was evaporated under reduced pressure and the viscous residue was taken in $CHCl_3$ (250 mL) and washed in sequence with 2N HCI (1x25 mL), ice cold 2N $NaHCO_3$ (2x25 mL) and finally with H_2O (2x25 mL). The organic layer was dried over anhydrous Na_2SO_4, filtered, and evaporated to dryness yielding the acetylated mixture. By acetylation the oligosaccharides were converted into their non-polar acetyl derivatives which resolved very nicely on TLC.

2.6. Purification of acetylated milk oligosaccharide on silica gel column

Separation of the acetylated products (12.0 g) was carried over silica gel using varying proportions of hexane: $CHCl_3$, $CHCl_3$ and $CHCl_3$: MeOH as eluants. The acetylated products was packed in a column (1.6 x 40 cm) (void volume = 25 mL) equilibrated with $CHCl_3$ and left for 10-12 h to settle down. Repeated column chromatography (7 column) led to the isolation of three chromatographically pure compounds A (858 mg), B (46 mg) and C (125 mg). The flow rate adjusted to 3 mL / min and the detailed purification processes description of polarity ration of eluent and volume of each fractions of compound A, and B is given in *Supplementary material Table S1-S7.*

2.7. Deacetylation of compounds

Compound A (52 mg) obtained from repeated column chromatography of acetylated oligosaccharide mixture was dissolved in acetone (3 mL) and NH_3 (3.5 mL) was added and left overnight in a stoppered hydrolysis flask. Ammonia was removed under reduced pressure and the product was washed with $CHCl_3$ (3x3 mL) and was finally freeze dried giving the deacetylated oligosaccharide 'A' (46 mg) $[\alpha]_D$+62.4°(c, 0.2, H_2O). Compound B (32 mg) was deacetylated by the same method as for compound A, resulting in deacetylated compound B (24 mg) $[\alpha]_D$-41.3° (c, 0.2, H_2O).

2.8. Description of compounds

The two compounds Compound 'A' (Bosiose) and Compound B (Bovisose) have been isolated and characterized with the help of spectroscopic techniques.

2.8.1. Compound 'A' **(Bosiose)**

Compound 'A' $C_{22}H_{38}O_{16}N_2$, $[\alpha]_D$ +62.4° (c, 2.0, H_2O). For elemental analysis, this compound was dried over P_2O_5 at 100°C and 0.1 mm pressure for 8 h. Elemental analysis: Calcd. C, 45.05; H, 6.48; N, 4.77; Found C, 45.01; H, 6.47 ; N, 4.76; It gave purple-violet spot on PC for amino sugars when sprayed with acetyl-acetone reagent I and heated, the dry strips were then sprayed with p-dimethyl amino benzaldehyde reagent II and returned to the oven for a further short hot treatment.

Methyl glycosidation/Acid hydrolysis: Compound A (8 mg) was ref1uxed with absolute MeOH (2 ml) at 70°C for 18 h in the presence of cation exchange IR-l20 (H) resin. The reaction mixture was filtered while hot and filtrate was concentrated. In the solution of methylglycoside of B, 1,4-dioxane (1 ml), and 0.1N H_2S0_4 (1 ml) was added and the solution was warmed for 30 minutes at 50°C. The hydrolysis was complete after 24 h. The hydrolysate was neutralized with freshly prepared $BaCO_3$ filtered and

concentrated under reduced pressure to afford α-and β-methylglucosides along with the Glc, GalNAc and GlcNAc. Their identification was confirmed by comparison with authentic samples (TLC, PC).

Kiliani Hydrolysis: For Kiliani hydrolysis [32], the compound A (5 mg) was dissolved in 2 ml Kiliani mixture (AcOH-H_2O-HCI, 7:11:2) and heated at 100°C for 1 h followed by evaporation under reduced pressure. Then 2 ml of H_2O was added and extracted twice with 3 ml $CHCl_3$. The aqueous residual solution was made neutral by addition of 1-2 drops of 2N NaOH and was evaporated under reduced pressure to afford Glc, GalNAc and GlcNAc on comparison with authentic samples of Glc, GalNAc and GlcNAc, respectively. This hydrolysis confirming the presence of Glc, GalNAc and GlcNAc sugar units in it. The presence of these sugar units have been further confirmed by NMR and Mass spectrometry given in *supplementary file 1*.

Mannich-Siewart hydrolysis: For Mannich-Siewart hydrolysis [33] 6 mg of the compound A was dissolved in 2.5 ml acetone and conc. HCl (0.02 ml). The solution was kept under carbon dioxide in dark room at room temperature. After five days paper chromatogram showed three spots, mobility of one spot was identical in mobility with authentic sample of GlcNAc, and the other spot with lowest mobility was identical with unreacted compound A. Further the compound with intermediate mobility may be the disaccharide. After eleven days two new spots with faster mobility than disaccharide were observed on TLC and PC which were found identical to authentic sample of GalNAc and Glc. The hydrolysis was completed in eleven days showing three spots on TLC which were found identical with Glc, GalNAc and GlcNAc (PC, TLC) on comparison with authentic samples.

^{1}H NMR values of compound A in D_2O: 5.27[d, 1H, J=3.6 Hz, α-Glc(S-1) H-1], 4.57[d, 2H, J=8.0Hz, β-Glc(S-1) & β-GlcNAc(S-3)H-1], 4.50[d, 1H, J=8.0Hz,

βGalNAc(S2), 3.82[βGalNAc(S-2), H-4], 3.32[t, 1H, J=7.2 Hz, β-Glc(S-1), H-2], 2.05[s, 3H, $NHCOCH_3$, β-GlcNAc(S-3), 2.04[s, 3H, $NHCOCH_3$, β-GalNAc(S-2)].

^{13}C NMR values of compound A in D_2O:

177.43[s,$NHCOCH_3$], 173.71[s,$NHCOCH_3$], 102.84[β-GlcNAc(S-3),C-1], 101.09[β-GalNAc(S-2),C-1], 95.74[β-Glc(S-1),C-1], 88.79[α-Glc(S-1),C-1], 21.81[s,$NHCOCH_3$], 21.25[s,$NHCOCH_3$].

^{1}H NMR values of acetylated A in $CDCl_3$: 6.25 [d, IH, J=3.3 Hz, β-Glc (S-1), H-l], 5.67 [d,1H, J=8.1 Hz, β-Glc (S-1), H-1], 4.49 [d, 1H, J=8.1 Hz, β-GlcNAc (S-3), H-1], 4.45 [d, 1H, J=9.9 Hz, β-GalNAc (S-1), H-1], 3.87 [β-Glc (S-1), H-4], 3.84 [α-Glc (S-1), H-4], 3.78 [β-GalNAc (S-2) H-6].

^{13}C NMR values of acetylated A in $CDCI_3$: 101.15 [β-GlcNAc (S-3), C-l], 100.87 [β -GalNAc (S-2), C-1], 91.51 [β-Glc (S-l), C-1], 88.94 [α-Glc (S-1), C-1], 75.63 [α-Glc (S-1), C-4], 73.47 [β-GalNAc (S-2), C-6].

ES Mass of compound A found 586$[M]^+$

2.8.2. Compound B (**Bovisose)**

Compound B, $C_{40}H_{68}O_{31}N_2$, $[\alpha]_D$ +41.3° (c, 2.0, H_2O). Elemental analysis, Calcd. C, 44.57; H, 6.34; N 2.61; Found C, 44.75; H, 6.33, N, 2.60; It gave purple-violet spot on PC for amino sugars when sprayed with acetyl-acetone reagent I and heated, the dry strips were then sprayed with p-dimethyl amino benzaldehyde reagent II and returned to the oven for a further short hot treatment.

Methyl glycosidation/Acid hydrolysis: 8 mg of the compound was ref1uxed with absolute MeOH (2 ml) at 70°C for 18 h. The similar procedure was opted for acid hydrolysis like compound A. After neutralized, filtrate, concentrated under reduced pressure of hydrolysate, afford α-and β-methylglucosides along with the Glc, Gal and GlcNAc. Their identification was confirmed by comparison with authentic samples

(TLC, PC).

Kiliani Hydrolysis: Compound B (5 mg) was used for Kiliani hydrolysis and the hydrolysate was checked by paper chromatography with authentic samples which inferred the presence of Glc, Gal and GlcNAc in compound B and further sugar units have been confirmed by NMR and Mass spectrometry given in *supplementary file 1*.

Mannich-Siewart hydrolysis: For this hydrolysis compound B (5 mg) was dissolved in 2 ml acetone and conc. HCl (0.02 ml) was added. The solution was kept under carbon dioxide in dark room at room temperature. After two days paper chromatogram showed three spots, the compound with faster mobility was identical in mobility with authentic sample of Gal while the spot with lowest mobility was identical with unreacted compound C, further the compound with the intermediate mobility may be the pentasaccharide. Further after five days two new spots were observed of which one was identical in mobility with authentic sample of GlcNAc and other, which was faster in mobility than the pentasaccharide, may be the tetrasaccharide. After seven days one more spot was observed which was having faster mobility than the tetrasaccharide, so it may be trisaccharide, which was formed by loss of another Gal moiety. Further after ten days a new spot was observed which was faster in mobility than the trisaccharide and having same mobility as the authentic sample of lactose and it was formed by the loss of GlcNAc moiety. The hydrolysis was partially completed in twelve days and showed a new spot, which was found identical with authentic sample of Glc on TLC and PC. Thus the hydrolysis was completed in twelve days showing three spots on TLC which were found identical with Glc, Gal, and GlcNAc (PC, TLC) on comparison with authentic samples.

^{1}H NMR values of compound B in D_2O: 5.25 [d, 1H, J=3.6Hz, α-Glc (S-1) H-1], 4.69 [d, 1H, J=8.0Hz, β-GlcNAc (S-3) H-1], 4.61 [d, 1H, J=7.6Hz, β-GlcNAc (S-5) H-l],

4.59 [d, 1H, J=7.6Hz, β-Glc (S-1) H-1], 4.55 [d, 1H, J=8.0Hz, β–Gal (S-6) H-1], 4.48 [d, 2H, J=8.0Hz, β-Gal (S-2)& β-Gal (S-4) H-1], 4.16 [β-Gal (S-2), H-4], 3.32 [t, 1H, J=8. 1 Hz, β-Glc(S-l), H-2], 2.03 [s, 3H, $NHCOCH_3$, β-GlcNAc (S-5)], 2.02 [s, 3H, $NHCOCH_3$, β-GlcNAc (S-3)]

^{13}C NMR values of compound B in D_2O: 177.38 [s, $NHCOCH_3$], 173.66 [s, $NHCOCH_3$], 103.8 [β-Gal (S-6), C-1], 102.9 [β-Gal (S-4), C-1], 102.8 [β-Gal (S-2), C-1], 101.1 [β-GlcNAc (S-5), C-1], 100.1 [β-GlcNAc (S-3), C-1], 95.8 [β–Glc (S-1), C-1], 91.7 [α-Glc (S-1), C-1], 21.9 [s, $NHCOCH_3$], 21.3 [s, $NHCOCH_3$]

^{1}H NMR values of acetylated B in $CDCl_3$: 6.27 [d, 1H, J=3.6Hz, α-Glc (S-1), H-1], 5.68 [d, 1H, J=8.1Hz, β-Glc (S-1), H-1], 4.64 [d, 1H, J=7.8Hz, β-GlcNAc (S-3) H-l], 4.62 [d, 1H, J=7.8Hz, β-GlcNAc (S-5), H-1], 4.40 [d, 1H. J=7.8 Hz, β-Gal (S-4), H-l], 4.38 [d, 1H, J=7.5 Hz, β-Gal (S-2), H-1], 4.36 [d, 1H, J=9.0 Hz, β-Gal (S6), H-1], 3.99 [β-Gal (S-2), H-3], 3.88 [β-Gal (S-4), H-6], 3.81 [β-Glc (S-1), H-4], 3.81[β–GlcNAc (S-3), H-3], 3.75 [β-GlcNAc (S-5), H-3].

^{13}C NMR values of acetylated B in $CDCl_3$:103.6 [β-Gal (S-4), C-1], 102.5 [β-Gal (S-2) & β-Gal (5-6) C-1], 101.1 [β-GlcNAc (S-5), C-1], 100.8 [β-GlcNAc (S-3), C-1], 91.7 [β-Glc (S-1), C-1], 89.0 [α-Glc (S-1), C-1].

ES Mass of compound B: 1073 [M+H]

3 Result and discussion

3.1 Stability of Molecular geometries of the isolated compounds

The geometry of compound A (Bosiose) and B (Bovisose) have been optimized at B3LYP method and 6-311+G(d,p) basis set and given in Figure 1. As we know that

molecular geometry and conformational analysis play a very important role in determining the structure-activity relationship [34-39]. The optimized geometries of two compounds show positive wavenumbers [34-39] values indicate the stability of the compound A and B. The optimized geometry of ground state structure of A and B are shown in **Figure 1**. The compound B is lower in energy as compared to compound A and C. The conformer of compound A and B has energies -1878.730 and -1879.999 a.u. at room temperature. The difference between the energies of A and B is 1.269 a.u. or 796.309 kcal/mol. The Molecular structures of all three conformers possess C_1 point group symmetry.

3.2 Structure elucidations of the isolated yak milk oligosaccharide

TLC of acetylated oligosaccharide mixture at different polarity proportions is given in *Supplementary Figure S2*. The Sephadex G-25 chromatography of yak milk oligosaccharide mixture showed four peaks i.e. I, II, III and IV as shown in **Figure 2** and collected fractions is shown in *supplementary Table S1*. A positive phenol-sulphuric acid test for sugars has been given by fractions under peaks II and III which indicate the presence of oligosaccharide in yak milk. These fractions were pooled and lyophilized. Pooled fractions (peaks II and III) obtained after Sephadex G-25 column were further qualitatively examined by reverse phase HPLC as shown in *supplementary Figure S2*. The reverse phase HPLC shows eleven peaks at varied retention times from 2.272 min. to 5.717 min as shown in Table 1.

3.2.1 NMR spectroscopy

The isolated compounds have been identified with the help of 1H, 13C NMR, 2D NMR (COSY, TOCSY and HSQC) and mass spectrometry. The experimental

spectra of the isolated compounds have been given in *Figure S3-S17 of supplementary file 1*. Most of milk oligosaccharides are made up of basic core units which are comprised of D-Glucose, D-Galactose, D-GlcNAc, D-GalNAc, Fucose and Sialic acid. The structure of milk oligosaccharides were elucidation by chemical degradation and spectroscopic techniques (NMR and ES-MS). Keeping above mentioned basic core units in mind, the structures of various milk oligosaccharides were established by comparing the chemical shift data (1H and 13C NMR) of anomeric signals and other important signals of unknown milk oligosaccharides with the chemical shifts of known milk oligosaccharides. In the present study analogies between chemical shift of certain 'structural reporter group resonances' (SRG) [40-52] were used to make proton resonance assignments as well as structural assignments of the oligosaccharides. All chemical shifts of anomeric proton signals of milk oligosaccharides were further confirmed by 2D (COSY, TOCSY and HSQC) NMR experiments, which were earlier assigned with the help of 1H and 13C NMR data. Other techniques like deacetylation, methylation, hydrolysis, chemical degradation and mass spectrometry were also used in the structural elucidation of the oligosaccharides.

Compound A **(Bosiose)** $C_{22}H_{38}O_{16}N_2[\alpha]_D$ +62.4° (c, 0.2, H_2O) and B (**Bovisose)** $C_{40}H_{68}O_{31}N_2$ $[\alpha]_D$ +41.3° (c, 0.2, H_2O) gave positive phenol-sulphuric acid test [31], Fiegl test [53] and Morgon-Elson [30] test showing the presence of normal and amino sugar moiety in the compound A. The HSQC spectrum of acetylated compound A and B given in *supplementary file 1* show the presence of four and seven cross peaks of respective protons and carbons in the anomeric region at δ(6.25x88.94, 5.67x91.51, 4.49x100.89, 4.45x101.15) and δ(6.27x89.0, 5.68x91.7, 4.64x100.8, 4.62x101.1, 4.40x103.6, 4.38x103.5, 4.36x103.5), respectively, suggested the presence of four and seven anomeric protons. The presence of four and seven anomeric protons were

confirmed by the presence of doublets at δ{6.25 (1H), 5.67(1H), 4.49(1H), 4.45(1H)},{ δ 6.27 (J=3.6 Hz), 5.68 (J=8.1 Hz), 4.64 (J=7.8 Hz), 4.62 (J=7.8 Hz), 4.40 (J=7.5 Hz), 4.38 (J=7.5 Hz), 4.36 (J=9.0 Hz)} in the ^{1}H NMR spectrum of acetylated A and B in $CDCl_3$ respectively. Again presence of four and seven anomeric carbons were confirmed by the presence of four and seven anomeric signals at {δ101.15, 100.89, 91.51 and 88.94} and {δ 104.35(2C), 104.16(1C), 99.95(1C), 91.51(1C), 89.19(1C)} in the ^{13}C NMR spectrum of acetylated A and B respectively. These four anomeric protons could be interpreted for the presence of a trisaccharide in its reducing form. The reducing nature of compound A and B were confirmed by the methyl glycosylation of compound A and B by $MeOH/H^{+}$ followed by its acid hydrolysis which led to the isolation of α and β methyl glucosides (3) which suggested the presence of glucose at the reducing end as seen in **Scheme 1** and **2**. The three and six monosaccharide unit have been present in compound A and B and designated as (S_1, S_2 and S_3) and (S_1, S_2, S_3, S_4, S_5, S_6,) respectively. The monosaccharide constituents in compound A and B were confirmed by Kiliani hydrolysis [32] under strong acidic conditions, followed by PC and TLC. In this hydrolysis three spots were found on PC and TLC which were identical with (Glc, GalNAc and GlcNAc) in compound A and (Glc, Gal and GlcNAc) in compound B by co-chromatography with authentic samples.

Thus the compound A (trisaccharide) and compound B (hexasaccharide) contained three types of sugar moieties i.e. (Glc, GalNAc, GlcNAc) and (Glc, Gal and GlcNAc) in it, respectively, and given in **Scheme 1 and 2.**

The monosaccharide constituents in compound A and B have been further confirmed by the hydrolyzing the compounds in under mild acidic conditions (Mannich-Siewart hydrolysis [33]) followed by PC and TLC. As shown in **Scheme A** and **B** after five days paper chromatogram showed three spots, mobility of one spot was identical in

mobility with authentic sample of GlcNAc, and the other spot with lowest mobility was identical with unreacted compound 'A' (I). Further the compound with intermediate mobility may be the disaccharide (II). After eleven days two new spots with faster mobility than disaccharide were observed on TLC and PC which were found identical to authentic sample of GalNAc and Glc.

The Mannich-Siewart hydrolysis [33] of compound B supported the monosaccharide sequence deduced by ^{1}H NMR. In this hydrolysis after two days paper chromatogram showed three spots, the compound with faster mobility was identical in mobility with authentic sample of Gal while the spot with lowest mobility was identical with unreacted compound B(I), further the compound with the intermediate mobility may be the pentasaccharide (II). Further after five days two new spots were observed which was identical in mobility with authentic sample of GlcNAc and other, and faster in mobility than the pentasaccharide, may be the tetrasaccharide (III). After seven days one more spot was observed which was having faster mobility than the tetrasaccharide, so it may be trisaccharide (IV), which was formed by loss of another Gal moiety. Further after ten days a new spot was observed which was faster in mobility than the trisaccharide and having same mobility as the authentic sample of lactose (V) and it was formed by the loss of GlcNAc moiety. The hydrolysis was partially completed in twelve days and showed a new spot, which was found identical with authentic sample of Glc on TLC and PC.

The reducing nature of glucose was confirmed by the presence of two doublets at δ5.27 (J=3.6Hz) , 4.57 (J=8.0Hz) and 5.25(J=3.6Hz), 4.59 (J=7.6Hz) for α and β glucose [54, 55] respectively, in the ^{1}H NMR spectrum of compound A and B in D_2O solvent, respectively, showing the presence of glucose at the reducing end. Another anomeric proton signal which appeared as doublet at δ4.50 along with a singlet of amide

methyl at δ 2.04 in the ^{1}H NMR spectrum of compound 'A' was due to the presence of GalNAc moiety [56]. Since βGlc H-2 signal appeared as a triplet at δ3.32 in the ^{1}H NMR spectrum of compound A, it indicated that the equatorially oriented hydroxyl group at C-4 of the reducing Glc was involved in glycosidation and the linkage was 1→4 between S_2 and S_1 (SRG) [40-52] showing the presence of a lactose type of structure with addition al signal of NHAc group of GalNHAc at 2.04. This linkage was further supported by the presence of β-Glc H-4 resonance which appeared at δ3.87 in the ^{1}H NMR spectrum of acetylated A.

3.4 Mass Spectrometry

Additional supports for the isolated structures have been obtained by mass spectrometry. The fragmentation pattern of compound A and B is shown in **Figure 3** and **4**. The appearance of molecular ion peak in compound A at m/z = 586 which was due to M^+ $C_{22}H_{38}O_{16}N_2$ $[\alpha]_D$ +62.4°. Further the mass fragments were formed by repeated H transfer in the oligosaccharide and was accompanied by the elimination of terminal sugar less water which gave a fragment at m/z = 383 for disaccharide fragment. On further fragmentation the disaccharide fragment at m/z = 383 gave a monosaccharide fragment at m/z = 180 which attributed to the loss of GalNAc (S_2). This confirmed that trisaccharide was comprised of one Glc, one GalNAc and one GlcNAc moieties and the second sugar in sequence from the non-reducing end of trisaccharide was GalNAc and glucose was present at the reducing end. Thus mass spectrometry not only confirmed the assigned structure but also confirmed the sequence of monosaccharide in oligosaccharide. Thus on the basis of all the above data the structure of isolated trisaccharide, 'A' was deduced as- GlcNAc-(1→6)GalNAc(1→4)Glc and shown in **Figure 5**.

In compound B observed molecular ion peak at m/z = 1073 due to [M+H],

confirmed the molecular weight of compound as 1072. The hexasaccharide m/z = 1072 on fragmentation gave pentasaccharide at m/z = 910(I), which was due to loss of S-6 sugar unit i.e. Gal (S-6) sugar unit linked to the S-5 of hexasaccharide. The peak m/z = 910 further fragmented to give a fragment peak at m/z = 707(II), corresponding to tetrasaccharide unit, which was due to loss of S-5 sugar unit i.e. Gl cNAc(S-5) sugar unit linked to the S-4 of pentasaccharide. The tetrasaccharide on fragmentation gave a fragment ion peak at m/z = 545(III), which was due to loss S-4 sugar unit i.e. Gal(S-4) sugar unit linked to the S-3 of tetrasaccharide unit. The trisaccharide on fragmentation gave fragment peak at m/z = 342(IV), which was due to loss of S-3 sugar unit i.e. GlcNAc(S-3) sugar unit linked to the S-2 of trisaccharide unit. This disaccharide on further fragmentation gave a fragment peak at m/z = 180(V), which was due to loss of S-2 sugar unit i.e. Gal(S-2) sugar unit linked to the S-l unit of disaccharide. Thus based on the above data the structure of compound B was deduced and shown in **Figure 6.**

4. Conclusion

The geometries of compound A (Bosiose) GlcNAc-(1→6)GalNAc(1→4)Glc and B (Bovisose) Gal-β(l→3)GlcNAc-β(1→6)Gal-β(1→3)GlcNAc-β(1→3)Gal-β (1→4)Glc have been optimized at B3LYP method and 6-311+G(d,p) basis set. The optimized geometries of two compounds show positive wavenumbers values indicated the stability of the compound A and B. The difference between the energies of A and B is 1.269 a.u. or 796.309 kcal/mol. The two novel oligosaccharides have been isolated by yak milk. The isolated compound A (**Bosiose**) and B (**Bovisose**) have been characterized with the help of ^{1}H, ^{13}C, 2D NMR spectroscopy, further the molecular ion peaks in compound A and B have been confirmed by the analysis of mass spectrometry. The isolated compound may possess the variety of biological activity. The testing of

biological activity of these isolated compounds has been continued in one laboratory.

Acknowledgments

Authors acknowledge CDRI for spectral recording and Department of Chemistry for computational help.

References

1) B. Casado, M. Affolter, M. Kussmann, J. Proteomics 73 (2009) 196–208

2) S. Divya, B. S. Subramonian, Intern. J. Engin. Trends Tech. 2(2013) 86-89.

3) J. C. Gildersleeve, S. M. Muthana, C. T. Campbell, ACS Chem. Biol. 7(2012) 31–43

4) M. Schwonzen, R. Schmits, S. E. Baldus, M. Vierbuchen, F. G. Hanish, Br J Cancer. 65(1992) 559–565.

5) K. Abe, J. M. J. Mckibbin, S. Hakomori, Eur J Biochem Bio Chem 258(1983) 11793–11797.

6) J N Fang, H. Proksch, H. Wagner, Phytochemistry 24(1985) 2619-22.

7) R Srivastava, D. K. Kulshretha, Phytochem 28(1989) 2877-2883.

8) A. Ferdouse, Md. Monirujjaman, Nova J. Med. Bio. Sci. 2(2014) 1-7

9) R. Peterson, Wai Y. Cheah, J. Grinyer, N. Packer, Glycobiology 23(2013) 1425–1438

10) L. R. Ruhaak, J. Huang, B. Lönnerdal, J. B. German, Mol Cell Proteomics. 11(2012) 015248.

11) W.L. Claeys, C. Verraes, S. Cardoen, J. De Block, A. Huyghebaert, K. Raes, K. Dewettinck, L. Herman, Food Control 42 (2014) 188-201

12) Kei E Fujimura, Nicole A Slusher, Michael D Cabana, and Susan V Lynch, Expert Rev Anti Infect Ther. 8(2010) 435–454

13) R. Mehra, D. Barile, M. Marotta, C. B. Lebrilla, C. Chu, J. B. German, PLoS One. 9(2014) 96040-96047.

14) M. Guillermo, R. Palacios, L. E. Cervantes, P. Ramos, B. Chavez-Munguia, D. S. Newburg, J. Biol. Chem. 278(2003) 14112-14120.

15) L.E. Cervantes, D.S. Newburg, G.M. Ruiz-Palacios, Pediatr. Res. 37(1995) 171-178.

16) Maciej Chichlowski, J. Bruce German, Carlito B. Lebrilla, David A. Mills, Annu. Rev. Food Sci. Technol. 2(2011) 331-351

17) C. Kunz, S. Rudloff, Adv. Exp. Med. Biol. 606(2008) 455-65.

18) A. M. Zivkovic, J. B. German, C. B. Lebrilla, D. A. Mills, Proc Natl Acad Sci U S A. 108(2011) 4653-4658.

19) F. Lara-Villoslada, E. Debras, A. Nieto, A. Concha, J. Galvez, E. Lopez-Huertas, J. Boza, C. Obled, Xaus, Clin Nutr. 25(2006) 477-88.

20) R. Saksena, D. Deepak, A. Khare, R. Sahai, L. M. Tripathi, Biophy Acta, 1428(1999) 433-445

21) T. Urashima, S. Asakuma, F. Leo, K. Fukuda, M. Messer, Oftedal, T. Olav Amer. Soc. Nut. Adv. Nutr. 3(2012) 473S–482S

22) D.d Rhode, D. B. Madsen, P. J. Brantingham, T. Dargye, Develop. Quat. Sci. 9(2007) 1571-0866

23) Y. W. Park, Bioactive Components in Milk and Dairy Products, John Wiley & Sons, Ltd., Georgia, Athens (2009).

24) J. Jiang, S. Chen, F. Ren, Z. Luo, S. S. Zeng, J. Dairy Res. 74(2007) 18−25

25) S. Kumar, V. S. Chouhan, A. Sanghi, U. V.S. Teotia, Veterinary World, 6(2013) 2231-0916

26) S. Kumar, U. V. S. Teotia, A. P. S. Aswal, Int. J. Livest. Res. 3(2013) 130-134

27) E. Muehlhof, A. Bennett, D. McMahon, Milk and dairy products in human nutrition, (2013), Nutrition Division, Food and Agriculture Organization of the United Nations (FAO), Viale delle Terme di Caracalla, 00153 Rome, Italy

28) M.J. Frisch et al,. Gaussian 09, (2009) Gaussian, Inc., Pittsburgh.

29) E. Taufik, K. Fukuda, A. Senda, T. Saito, C. Williams, C. Tilden, Glycocon. J. 1(2012) 012-037

30) J. A. Cifonelli, D. Albert, Biol. Chem. 231(1958) 11-18.

31) M. Dubois, K.A. Gilles, J. K. Hamilton, P. A. Rebers, F. Smith, Analytical Chemistry. 28(1956) 350-356.

32) H. Kiliani, U. D. Verum. Ber. Deutsch Chem.Ges. 63(1930) 2866-2869.

33) C. Mannich, G. Siewert, Ber. Deutsch. Chem. Ges. 75(1942) 737-750.

34) R.N. Singh, P. Rawat, S. Sahu Spectrochimica Acta Part A 135 (2015) 1162–1168

35) R.N. Singh, A. Kumar, R.K. Tiwari, P. Rawat Spectrochimica Acta Part A 113(2013) 378–385

36) P. Rawat, R.N. Singh, Spectrochimica Acta Part A 140(2015) 344–355

37) P. Rawat, R. N. Singh, J. Mol. Struc. 1082(2015) 118–130

38) R.N. Singh, A. Kumar, P. Rawat, A. Srivastsva, J. Mol. Struc. 1049(2013) 419–428

39) P. Rawat, R.N. Singh, J. Mol. Struc. 1081(2015) 293–303

34) V. K. Dua, C. A. Bush, Analy. Biochem.133(1983) 1-8.

35) I.M.V. Nieuwenhof, W.E. Schiphorst, D.H.V. Eijden, FEBS Letters. 459(1999) 377-380.

36) G. Strecker, S. Fievre, J. Micheal, J. Montreuil, Carbohydrate Research. 226(1992) 1-14.

37) Y. Uemura, S. Asakuma, L. Yon, T. Saito, K. Fuuda, I. Arai, T. Urashima, Comp. Biochem. Phy. Part A Mol. Int. Phy. 145(2006) 468-478.

44) T. Urashima, T. Nakamura, K. Teramoto, I. Arai, T. Saito, T. Komatsu, Comp. Biochem. Phy. Part A Mol. Int. Phy. 139(2004) 587-595.

45) V.K. Dua, K. Goso, V.E. Dube, C.A. Bush, J. Chromat. 259(1985) 259-269.

46) P. Chaturvedi, C.B. Sharma, Biophysica Acta. 967(1988) 115-121

47) M.T. Tarrago, K.H. Tucker, H.V. Halbeak, D.V. Smith, Archieves Biochem. Biophys. 207(1988) 352-362.

48) D. Viverge, L. Grimmonprez, M. Solere, Biochimica et Biophysica Acta 1336(1997) 157-164.

43) T. Urashima, T. Saito, K. Ohmisya, K. Shimazak, Biochim. Biophys. Acta. 1073(1991) 295-300.

50) T. Urashima, M. Kobayashi, S. Asakuma, Y. Uemura, I. Arai, Comp. Biochem. Phy. Part B: Biochem. Mol. Bio. 146(2007) 153-159.

51) G. Gunnar, P. Lipniunas, T. Lundgren, F. Lindh, N. Bo, Archives Biochem. Biophy. 278(1990) 546-555

52) T. Urashima, T. Nakamura, K. Teramoto, T. Mizukami, I. Arai, T. Saito, Comp. Biochem. Phy. Part B: Biochem. Mol. Bio. 135(2003) 439-448.

53) F. Fiegl, Spots tests in organic analysis. Elsvier Publication, Amsterdam, (1975) pp. 337.

54) F. Dorland, B. L. Schut, J.F.G. Vliegenthart, G. Strecker, B. Fournet, G. Spik, J. Montreuil, Eur. J. Biochem. 73(1977) 93-97.

55) P. Chaturvedi, B. C. Sharma, Carbohydrate Research. 203(1990) 91-101

56) Y. Uemura, S. Asakuma, T. Nakamura, I. Arai, M. Taki, T. Urashima, Biochimica et Biophysica Acta. 1725(2005) 290-297.

Captions

Scheme 1 The trisaccharide contained three types of sugar moieties Glc(4), GalNAc(5) and GlcNAc (6)

Scheme 2 The hexasacharride contained three types of sugar moieties Glc(4), Gal (5) and GlcNAc (6)

Scheme-A Mannich-Siewart hydrolysis of compound A **(Bosiose)**

Scheme-B Mannich-Siewart hydrolysis of compound B **(Bovisose**)

Figure 1. The optimized geometries of A **(Bosiose)** and B **(Bovisose**)

Figure 2. Spehadex G-25 chromatography of yak milk oligosaccharides detected by phenol sulphuric acid method. Elution was made with distilled water

Figure 3. Mass fragmentation of oligosaccharide 'A' (Bosiose) (a) by repeated H transfer (b) ES-MS mass fragments

Figure 4. ES Mass fragmentation of oligosaccharide B (Bovisose) (a) by repeated H transfer (b) ES-MS mass fragments

Figure 5. The elucidate structure of isolated compound A **(Bosiose)**

Figure 6. The elucidate structure of isolated compound B **(Bovisose**)

Table 1: Reverse phase HPLC of yak milk oligosaccharides

Supplementary material

Figure S1. TLC **(a)** acetylated oligosaccharide mixture at different polarity proportions and **(b)** Isolated acetylated oligosaccharide at $CHCl_3$:MeOH (95:5).

Figure S2. Reverse phase HPLC

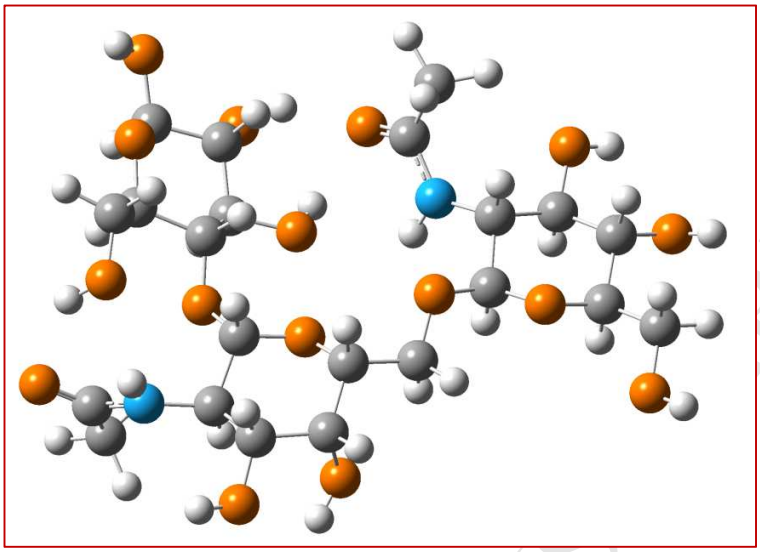
Figure S3-S17. NMR and Mass Spectral data of compound A and B

Table S1. Combined fractions (200 mg (88 x 2 = 18.09 gm) of yak milk oligosaccharide mixture chromatographed over Sephadex G-25 chromatography.

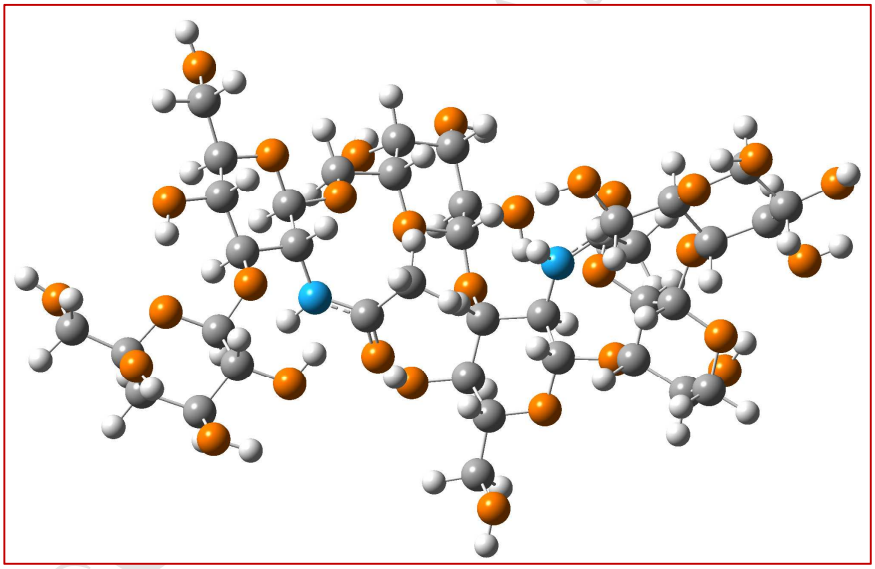
Table S2-S7. Purification of acetylated milk oligosaccharides on silica gel column

Table 1: Reverse phase HPLC of yak milk oligosaccharides

S.No	Retention time	Area	Area (%)	Height	Height (%)
1	2.272	9822	0.67	1688	7.17
2	3.115	40360	2.74	4827	20.51
3	3.637	13282	0.90	1551	6.57
4	3.968	542	0.04	91	0.39
5	4.331	16671	1.13	1388	5.90
6	4.693	746	0.05	104	0.44
7	5.077	312	0.02	52	0.22
8	5.227	170	0.01	40	0.17
9	5.491	122	0.01	52	0.22
10	5.557	206	0.02	54	0.23
11	5.717	284	0.02	64	0.27



A



B

Figure 1 The optimized geometries of A **(Bosiose)** and B **(Bovisose**)

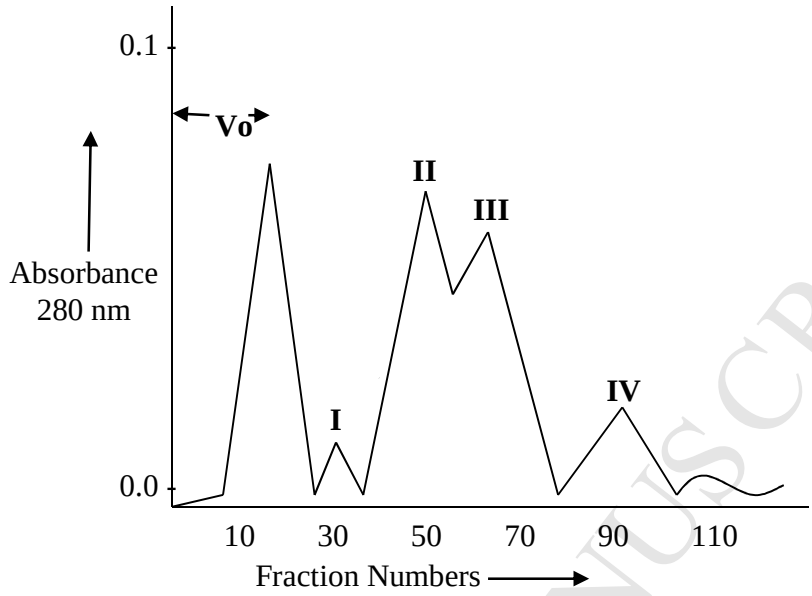
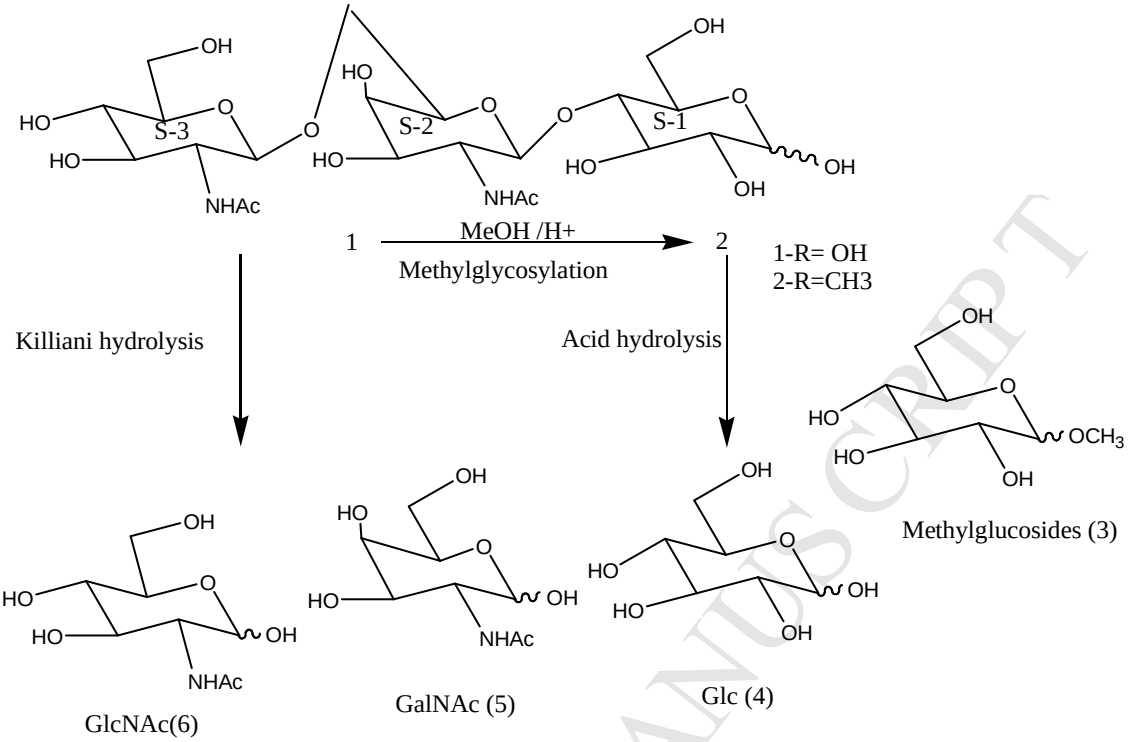
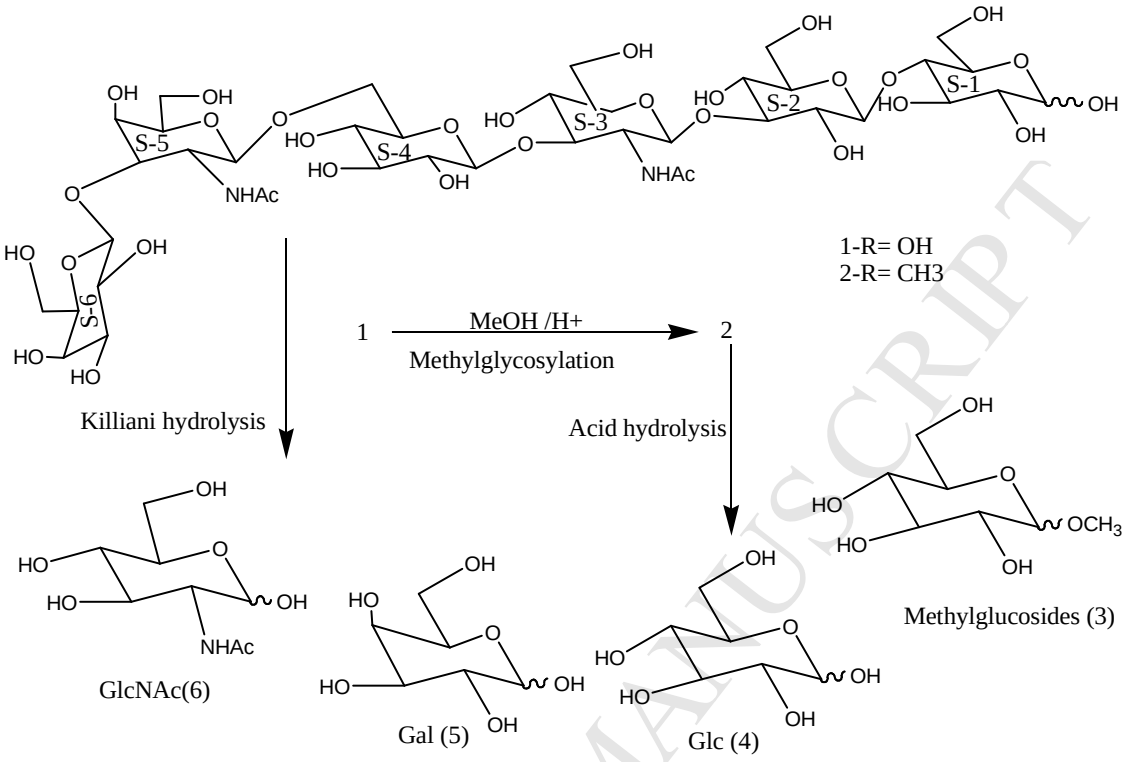


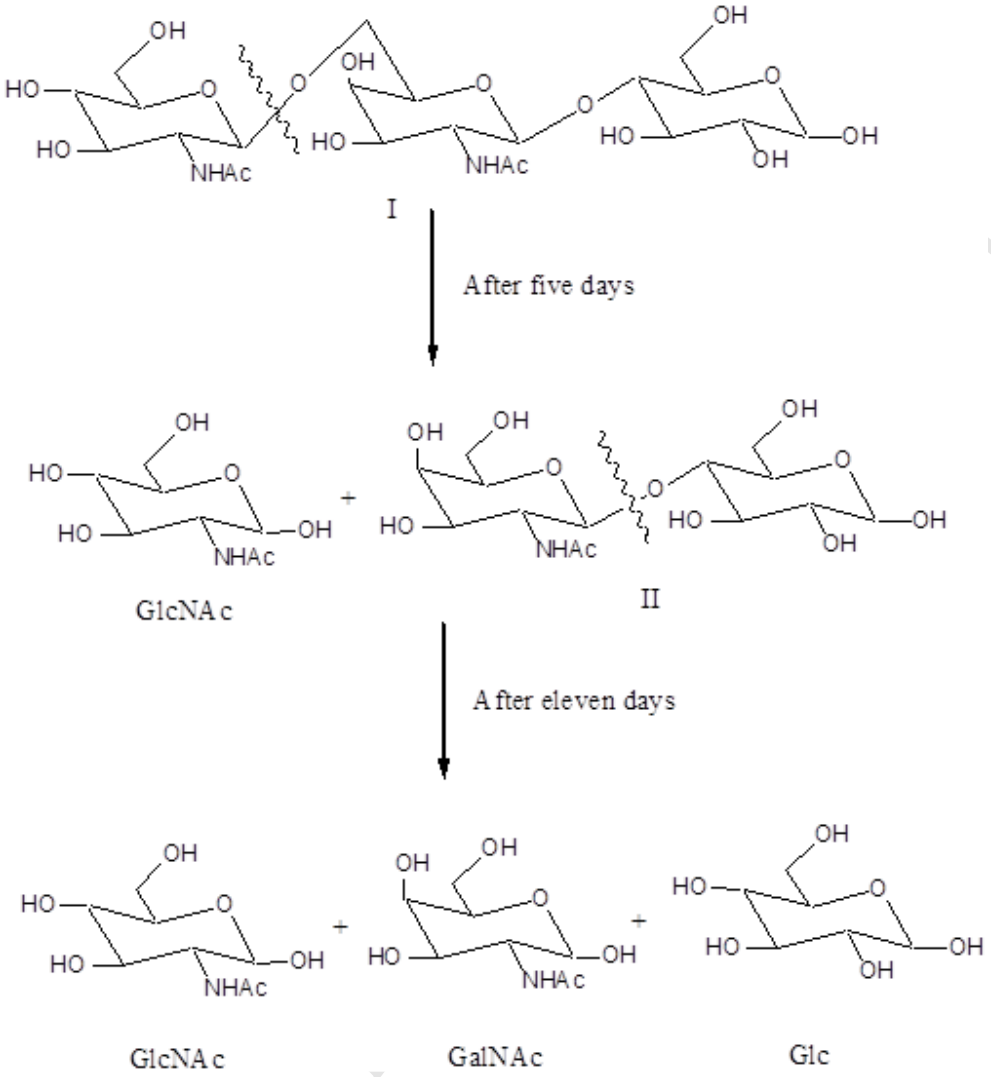
Figure 2. Spehadex G-25 chromatography of yak milk oligosaccharides detected by phenol sulphuric acid method. Elution was made with distilled water



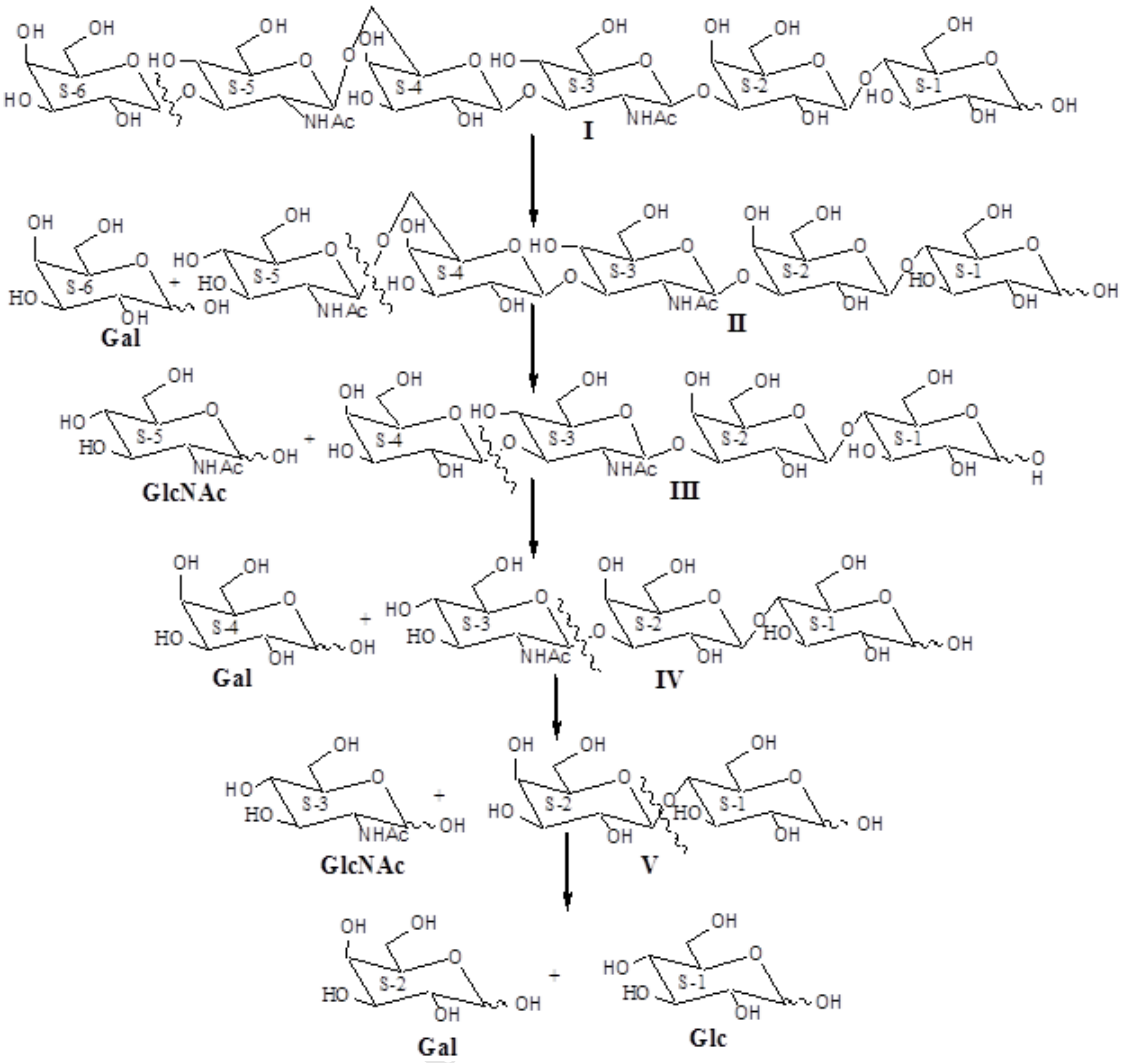
Scheme 1 The trisaccharide contained three types of sugar moieties Glc(4), GalNAc(5) and GlcNAc (6)



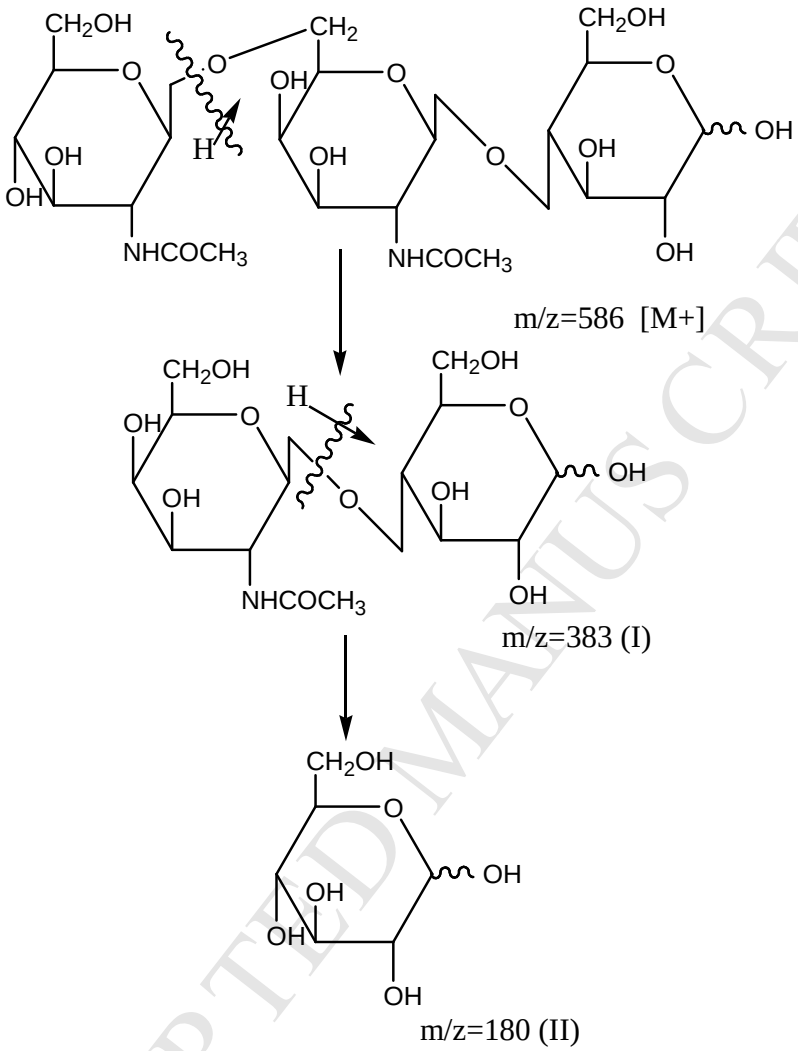
Scheme 2 The hexasacharride contained three types of sugar moieties Glc(4), Gal (5) and GlcNAc (6)



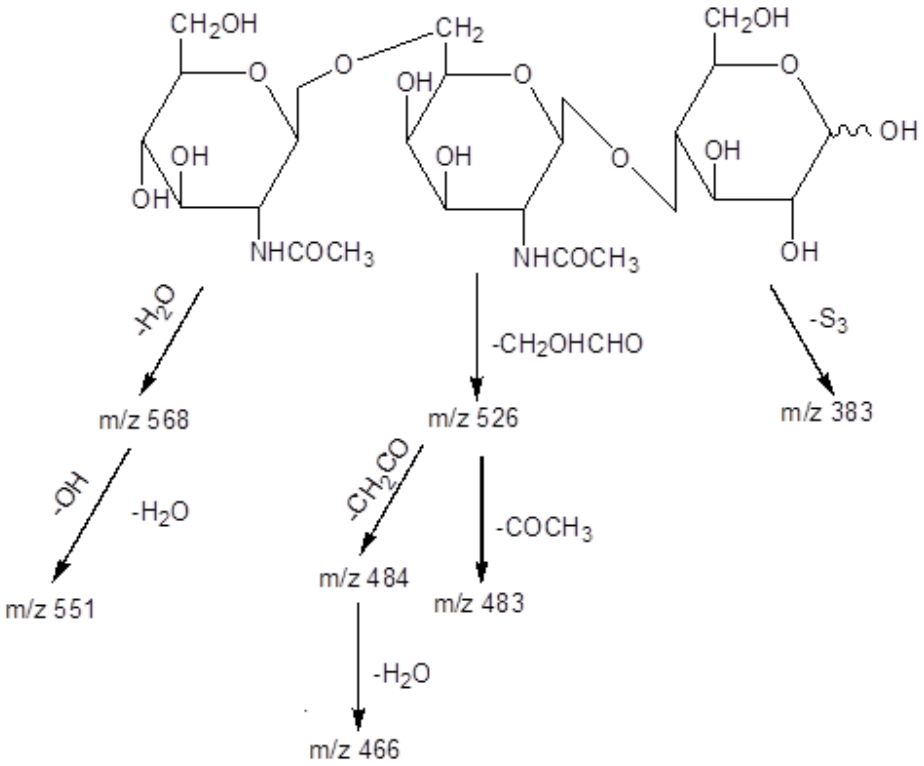
Scheme-A Mannich-Siewart hydrolysis of compound A **(Bosiose)**



Scheme-B Mannich-Siewart hydrolysis of compound B **(Bovisose**)

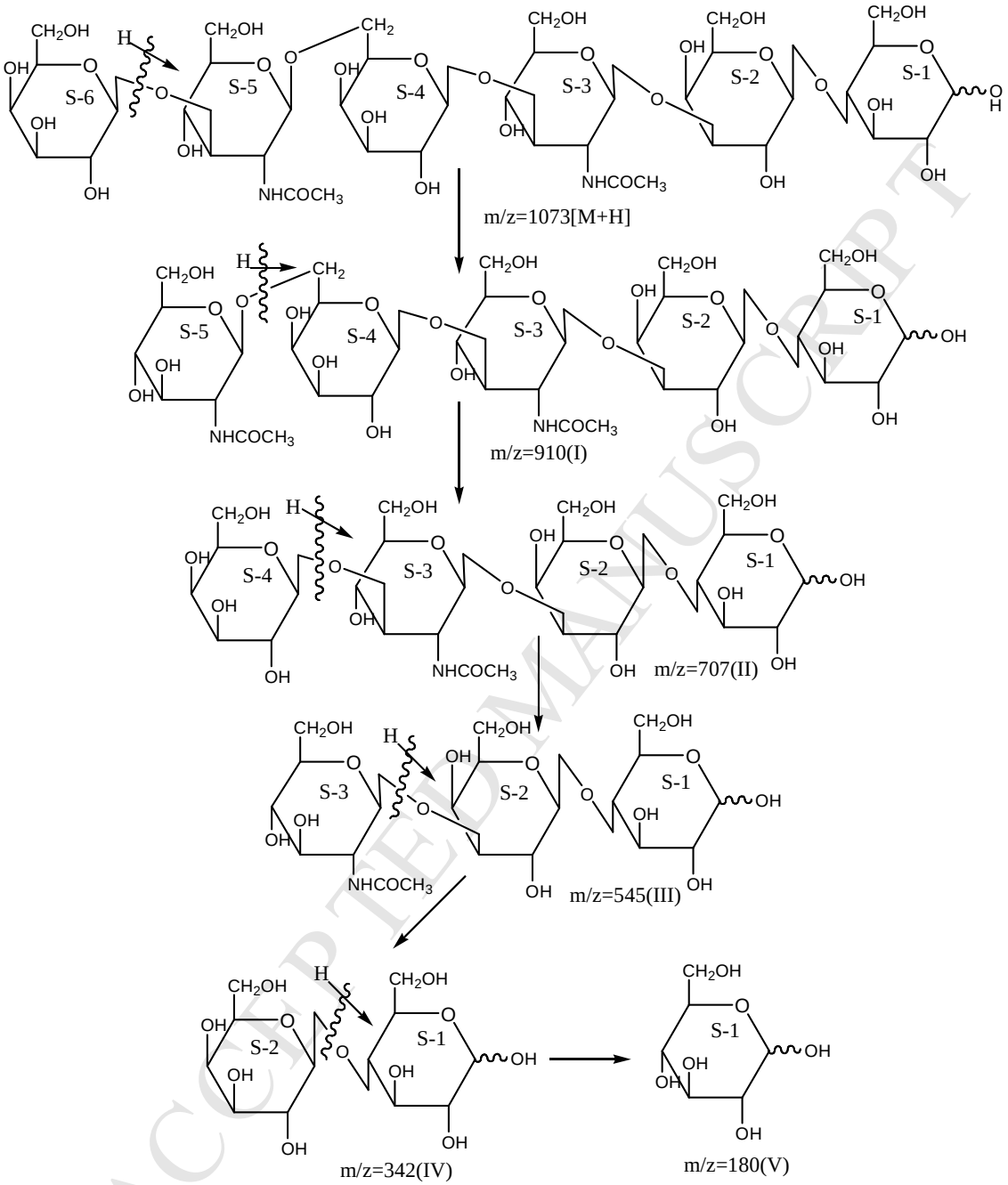


(a)

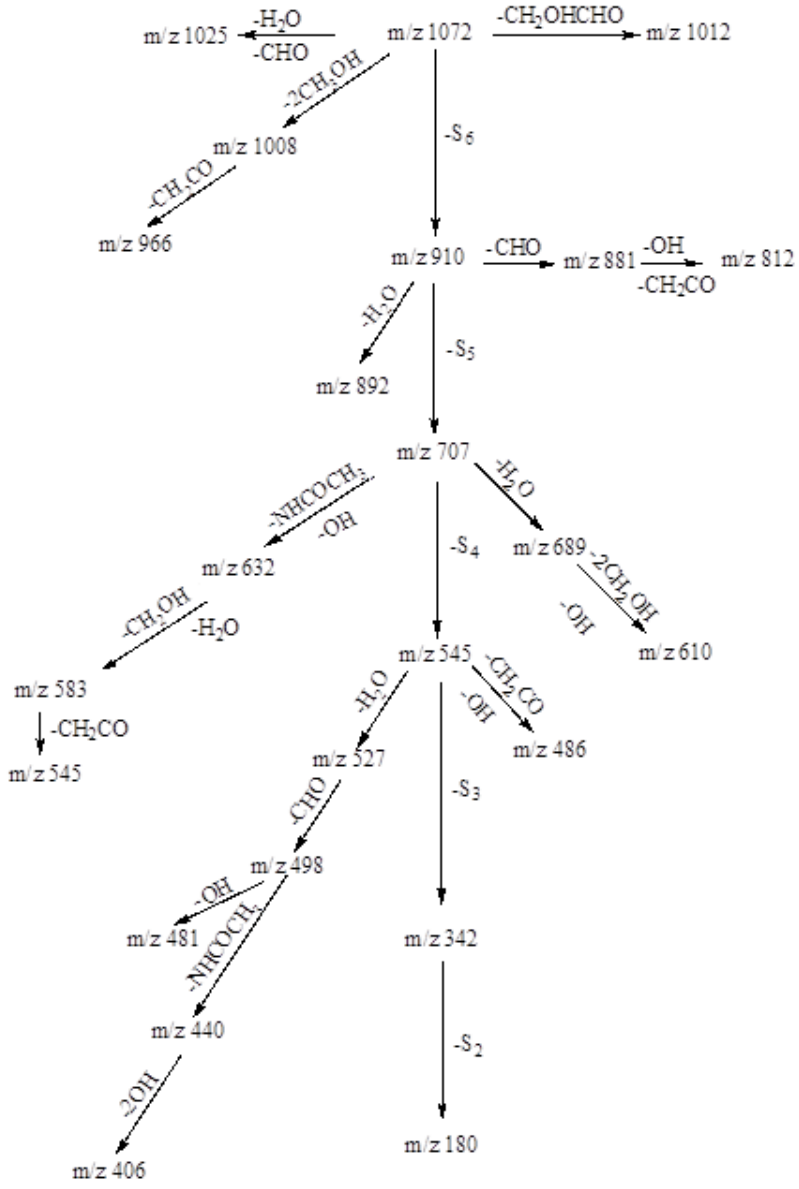


(b)

Figure 3. Mass fragmentation of oligosaccharide 'A' **(Bosiose)** (a) by repeated H transfer (b) ES-MS mass fragments

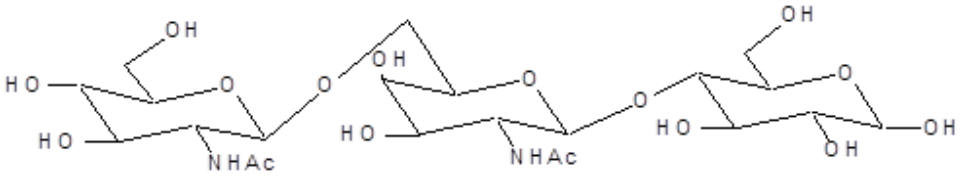


(a)



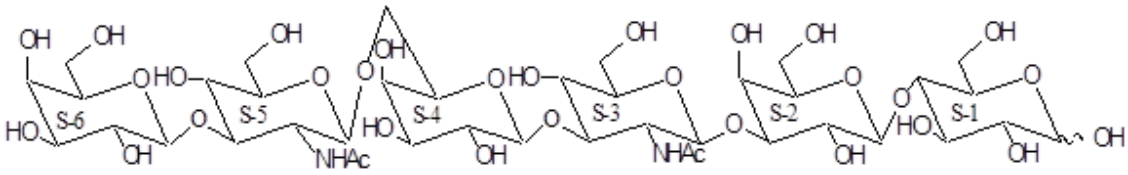
(b)

Figure 4. ES Mass fragmentation of oligosaccharide B **(Bovisose)** (a) by repeated H transfer (b) ES-MS mass fragments



GlcNAc-(1→6)GalNAc(1→4)Glc

Figure 5. The elucidate structure of isolated compound A **(Bosiose)**



Gal-β(l→3)GlcNAc-β(1→6)Gal-β(1→3)GlcNAc-β(1→3)Gal-β (1→4)Glc

Figure 6. The elucidate structure of isolated compound B **(Bovisose**)

Highlights

► Two novel oligosaccharides were isolated from Yak milk.

►The structure of novel oligosaccharides were elucidate with the help of spectroscopy.

► Geometry of all novel oligosaccharides were study with the help of DFT.