Isolation and Structure Elucidation of an Undefined Hexasaccharide 'Thisose' from Milk of Rathi Cow

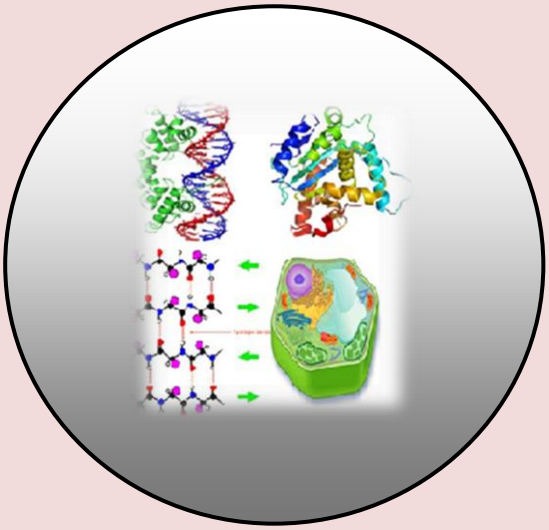
By

Desh Deepak A.P. Singh Chauhan, Sarita Chauhan, Manisha Shukla and Desh Deepak

ISSN 2319-3077 Online/Electronic
ISSN 0970-4973 Print

Index Copernicus International Value
IC Value of Journal 82.43 Poland, Europe (2016)
Journal Impact Factor: 4.275
Global Impact factor of Journal: 0.876
Scientific Journals Impact Factor: 3.285
InfoBase Impact Factor: 3.66

J. Biol. Chem. Research
Volume 41 (1), 2024 Pages No. 73-88



Journal of Biological and Chemical Research

An International Peer Reviewed / Referred Journal of Life Sciences and Chemistry

Indexed, Abstracted and Cited in various International and National Scientific Databases

Published by Society for Advancement of Sciences®

J. Biol. Chem. Research. Vol. 41, No. 1, 73-88, 2024
Ms 41/01/09/2024

ISSN 2319-3077 (Online/Electronic)
ISSN 0970-4973 (Print)

D.D.A.P.S. Chauhan
http:// www.sasjournals.com
http:// www.jbcr.co.in
jbiolchemres@gmail.com

RESEARCH PAPER
Received: 28/02/2024 Revised: 20/06/2024 Accepted: 21/06/2024

Isolation and Structure Elucidation of an Undefined Hexasaccharide 'Thisose' from Milk of Rathi Cow

Desh Deepak A. P. Singh Chauhan, Sarita Chauhan,
*Manisha Shukla and *Desh Deepak
Department of Chemistry, JNM P.G. College, Lucknow, U.P. India 226004
*Department of Chemistry, University of Lucknow, Lucknow, U.P. India 226007

ABSTRACT

Oligosaccharides present in various animal species have shown anti-complementary, anti-inflammatory, anti-cancer, anti-oxidant, immunostimulant and brain development activities. Amongst the various animal species the cow is the largest producer of milk which fulfills the requirement of mankind. In India number of cow species like Kankrej, Rathi, Tharparkar, Sahiwal, Jarsi etc. are available for their milk. Since these species belong to various parts of India and due to their different geographical availability and different food habits they constitute varied variety of milk oligosaccharides. Cow milk oligosaccharides are generally made up of Glucose, Galactose, GlcNHAc and GalNHAc with different conformation, configuration and glycosidic linkages present therein. Keeping in mind the above facts we have selected the Rathi cow milk for its Oligosaccharide content. For this purpose the Rathi cow milk was collected in bulk and was processed by modified method of Kobata and Ginsberg. The milk oligosaccharide mixture so obtained was acetylated and purified on silica column chromatography which resulted in the isolation of a novel hexasaccharide Thisose. The stereoscopic structure of this oligosaccharide was elucidated by the data generated from ^{1}H, ^{13}C, COSY, TOCSY, HSQC and Mass spectrometry along with chemical degradation and chemical transformations. In the light of above results the structure of Thisose was elucidated which is as under,

Gal-β-(1→2)-Gal-α-(1→4)-Glc-α-(1→3)-GalNHAc-β-(1→3)-GalNHAc-β-(1→4)-Glc

Keywords: Rathi Cow, Milk, Oligosaccharides, Thisose, NMR and Mass Spectrometry.

INTRODUCTION

The cow milk is extremely nutrient for the mankind and is full of protein, fat and carbohydrate. It is a good source of calcium, phosphorous, vitamins, Iodine and magnesium. It plays a definite role in lowering the blood pressure, muscle function, strengthening of immune system, repair of damage cells. The literature revealed from ancient medicinal system that it is described as amrata, and is used for the deceases like epilepsy, jaundice, spleen disorder and piles (Ling et al, 1961). With the recent advancement in science particularly in glycochemistry, it was confirmed that all these activities in the milk of cow is due to its oligosaccharide content. These oligosaccharides are long and branched chain of monosaccharides like Glucose, Galactose, GlcNHAc and GalNHAc which are linked together with O-glycosidic linkages having α and β configurations. The variability in the structure of oligosaccharides is also due to position of linkages in the monosaccharide units. The variations in the oligosaccharides also depends on the species of cow i.e. Friesian, Gir, Rathi, A2 etc. and the geographical flora and fauna around the animal and fodder they eat. In the present communication we have selected the Rathi cow milk from Panchmukhi district of Rajasthan state of India for its oligosaccharide content. The milk was processed by Modified method of Kobata and Ginsberg (Kumar et al, 2018) following the methodology described in our previous communications (Chauhan et al, 2024) and the mixture of oligosaccharides so obtained were purified by combining the various chromatographic techniques like gel filtration, HPLC (Churms, 1996), Column chromatography (Das M et al, 1998) and Thin Layer Chromatography which resulted into the isolation of novel hexasaccharide Thisose. The structure of this novel oligosaccharide was established by recent physicochemical experiments of NMR involving ^{1}H, ^{13}C, COSY (Davis, 1985), TOCSY (Davis, 1985), HSQC (Kay L et al, 1992) and Mass spectrometry (Barr et al, 1991). The traditional technique of chemical degradation and chemical transformations has also been incorporated. The acetylated compound was explained as ‘b’ while natural compound was described as ‘B’.

EXPERIMENTAL

General Procedure

General procedures were the same as described in our earlier communication (Chauhan et al, 2024).

Isolation of Rathi Cow Milk Oligosaccharides by Modified Method of Kobata and Ginsburg (Kumar et al 2016)

10 litre Rathi cow milk (1-5 day) was collected from a domestic cow of Panchmukhi district of Rajasthan state of India. The oligosaccharide was isolated by the process described in our previous communication (Chauhan et al, 2024) affording crude oligosaccharides mixture (283 gm).

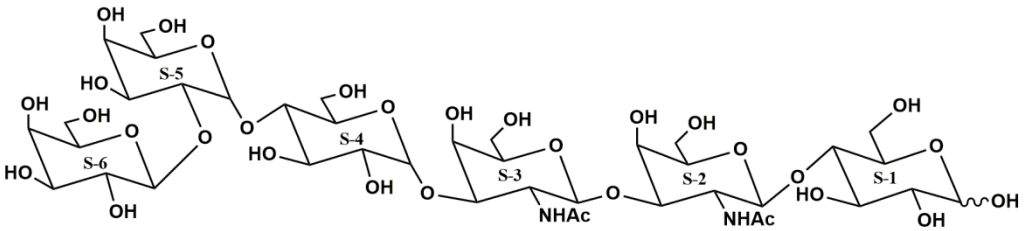
Acetylation of Oligosaccharide Mixture

10.2 gm oligosaccharide mixture was acetylated with pyridine (10.2 ml) and acetic anhydride (10. 2ml) at 60°C and was processed by earlier method (Chauhan et al, 2024) yielding the acetylated oligosaccharide mixture (10.6 gm). The TLC showed five spots i.e. a, b, c, d and e in the acetylated oligosaccharide mixture of Rathi cow milk.

Deacetylation of Compound ‘b’, Thisose Acetate (Chauhan et al, 2024)

Compound **‘b’** (45 mg) was obtained from column chromatography of acetylated oligosaccharide mixture.

Further 40 mg of compound 'b' was dissolved in acetone (4 ml) and 4 ml of NH_4OH was added in for obtaining natural oligosaccharide which gave the deacetylated natural oligosaccharide **'B'**, Thisose (31 mg).



Compound B, THISOSE

Methyl Glycosidation/Acid Hydrolysis of Compound B: (Shukla et al; 2023)
Compound '**B'** (10 mg) was refluxed with absolute MeOH (2 ml) at 70ºC for 18 hr in the presence of cation exchange IR-l20 (H) resin. The reaction mixture was filtered while hot and filtrate was concentrated. To this reaction mixture of methylglycoside 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 completed in 24 hrs, and gave α-and *β*-methylglucosides along with the Glc, Gal and GalNHAc. Identification of monosaccharides in compound 'B' was confirmed by comparison with authentic samples (TLC, PC) of *α*-and *β*-methylglucosides along with the Glc, Gal and GalNHAc.

Killiani Hydrolysis of Compound B, Thisose: (Khan M et al; 2019)
Compound 'B' (5 mg) was dissolved in 2 ml Killiani mixture (AcOH-H_2O-HCl, 7:11:2) and heated at 100°C for 1 h followed by evaporation under reduced pressure which afforded Glc, Gal and GalNHAc which were identified by their authentic samples of Glc, Gal and GalNHAc.

Description of Compound B, Thisose
Compound 'b' (43mg) obtained from column chromatography. On deacetylation of 35 mg of substance 'b' with NH_4OH/acetone, it afforded substance B (29 mg) $[\alpha]_D^{25}$ = -18º (*c* 1% H_2O). For experimental analysis, this compound was dried over P_2O_5 at 100°C and 0.1 mm pressure for 8 hr. It gave positive Phenol-sulphuric acid test (Partridge et al 1949), Feigl test (Feigl F, 1975) and Morgan-Elson test (Warren L, 1960).

$C_{40}H_{68}O_{31}N_2$	%C	%H	%N
Calculated	**44.77**	**6.34**	**2.61**
Found	**44.78**	**6.36**	**2.61**

^{1}H NMR of Compound-b, Thisose Acetate in $CDCl_3$ at 300 MHz.
δ6.16[d, 1H, J=3.6Hz, *α*-Glc(S-1) H-1], δ5.62[d, 1H, J=8.1Hz, *β*-Glc(S-1) H-1], δ5.35[d, 1H, J=3.6Hz, *α*-Gal(S-5) H-1], δ5.34[d, 1H, J=3.3Hz, *α*-Glc(S-4) H-1], δ4.60[d, 1H, J=8.1Hz, *β*-Gal(S-6) H-1], δ4.52[d, 2H, J=8.1Hz, *β*-GalNHAc(S-3) H-1], δ4.48[d, 2H, J=8.1Hz, *β*-GalNHAc(S-2) H-1], δ3.90[m, 1H, *α*-Gal(S-5) H-2], δ3.84[m, *β*-GalNHAc(S-2) H-3], 3.83[m,*α*-Glc(S-4) H-4], δ3.82[m, *β*-GalNHAc(S-3) H-3] and δ3.82 [m, *β*-Glc(S-1) H-4].

^{13}C NMR of Compound-b, Thisose Acetate in $CDCl_3$ at 300 MHz.
δ88.90[1C, *α*-Glc(S-1) C-1], δ89.12[2C, *α*-Glc(S-4) & *α*-Gal(S-5) C-1], δ90.36[1C, *β*-Glc(S-1) C-1], δ100.83[2C, *β*-GalNHAc(S-2) & *β*-GalNHAc(S-3) C-1], δ101.20[1C, *β*-Gal(S-6) C-1].

^{1}H NMR of Compound-B, Thisose in D_2O at 300 MHz.

δ5.74[d, 1H, J=3.9Hz, *α*-Glc(S-1) H-1], δ5.22[d, 2H, J=3.6Hz, *α*-Glc(S-4) & *α*-Gal(S-5) H-1], δ4.56[d, 1H, J=8.1Hz, *β*-Glc(S-1) H-1], δ4.52[d, 1H, J=7.8Hz, *β*-Gal(S-6) H-1], δ4.45[d, 1H, J=7.5Hz, *β*-GalNHAc(S-3) H-1], δ4.38[d, 1H, J=7.8Hz, *β*-GalNHAc(S-2) H-1], δ3.27[t, *β*-Glc(S-1) H-2], δ2.00[s, 3H, ($NHCOCH_3$), *β*-GalNHAc(S-3)], δ1.98[s, 3H, ($NHCOCH_3$), *β*-GalNHAc(S-2)].

ES MASS ion fragments of compound B, Thisose

1095 $[M+Na]^+$, 1072$[M]^+$,1054, 1020, 960, 910, 899, 875, 846, 841, 778, 748, 743, 713, 690, 677, 654, 633, 623, 586, 577, 527, 461, 460, 459, 427, 415, 410, 383, 334, 304, 180, 162.

Result and Discussion of Compound B, Thisose

Compound B, Thisose, $C_{40}H_{68}O_{31}N_2$ $[\alpha]_D^{25}$= -18º gave positive Phenol-sulphuric acid test (Partridge et al 1949, Dubois M et al, 1956), Feigl test (Feigl F, 1975) and Morgan-Elson test (Warren L, 1960) showing the presence of normal and amino sugars moietie(s) in the Compound Thisose. The compound was purified and isolated into its acetylated derivative from Rathi cow milk by silica column chromatography. The acetylated Thisose was described as compound 'b'. In this paper while the deacetylated natural oligosaccharide Thisose was designated as 'B'. The HSQC spectrum of acetylated Thisose 'b' showed the presence of six cross peaks of seven anomeric protons and carbons in their respective region at δ6.16 × 89.86, δ5.62 × 90.36, δ5.34 × 89.12, δ5.35 × 89.12, δ4.60 × 101.20 and δ4.48 × 100.83 suggested the presence of seven anomeric protons and carbon in Compound 'b' in its reducing form.

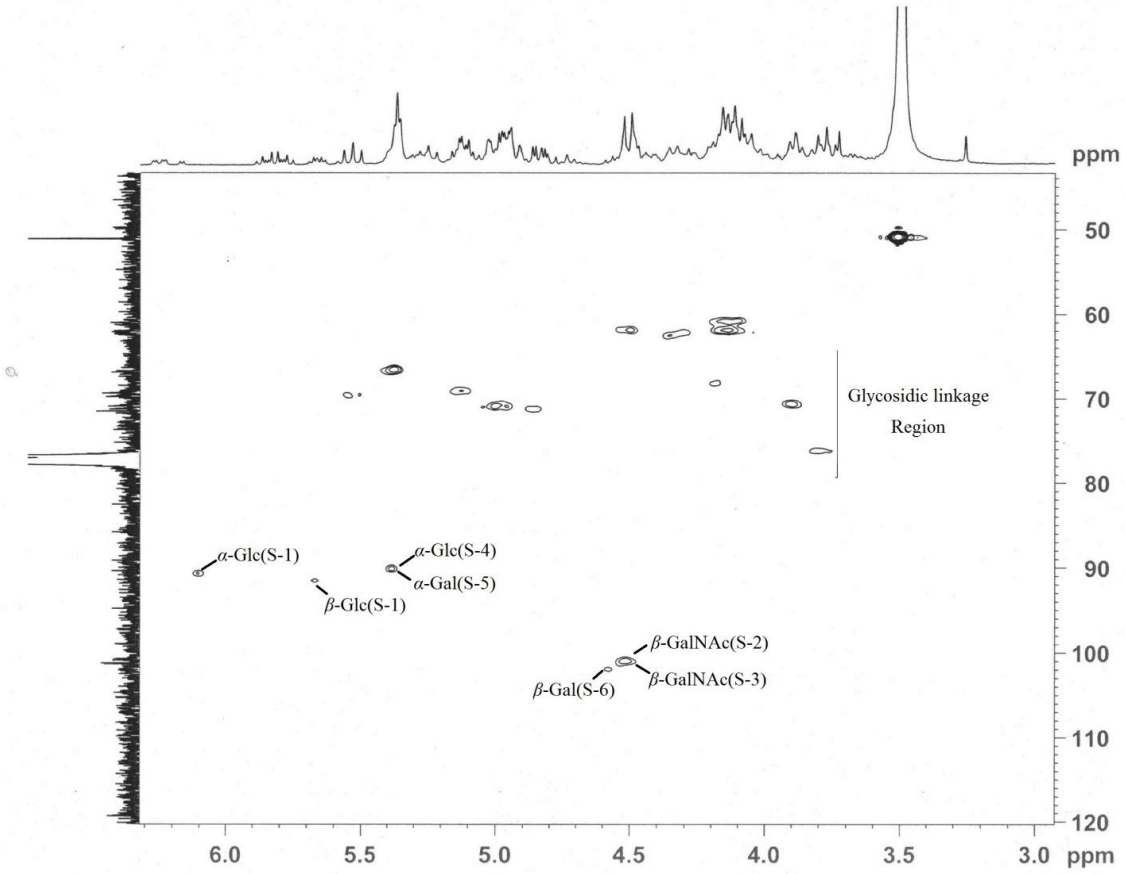
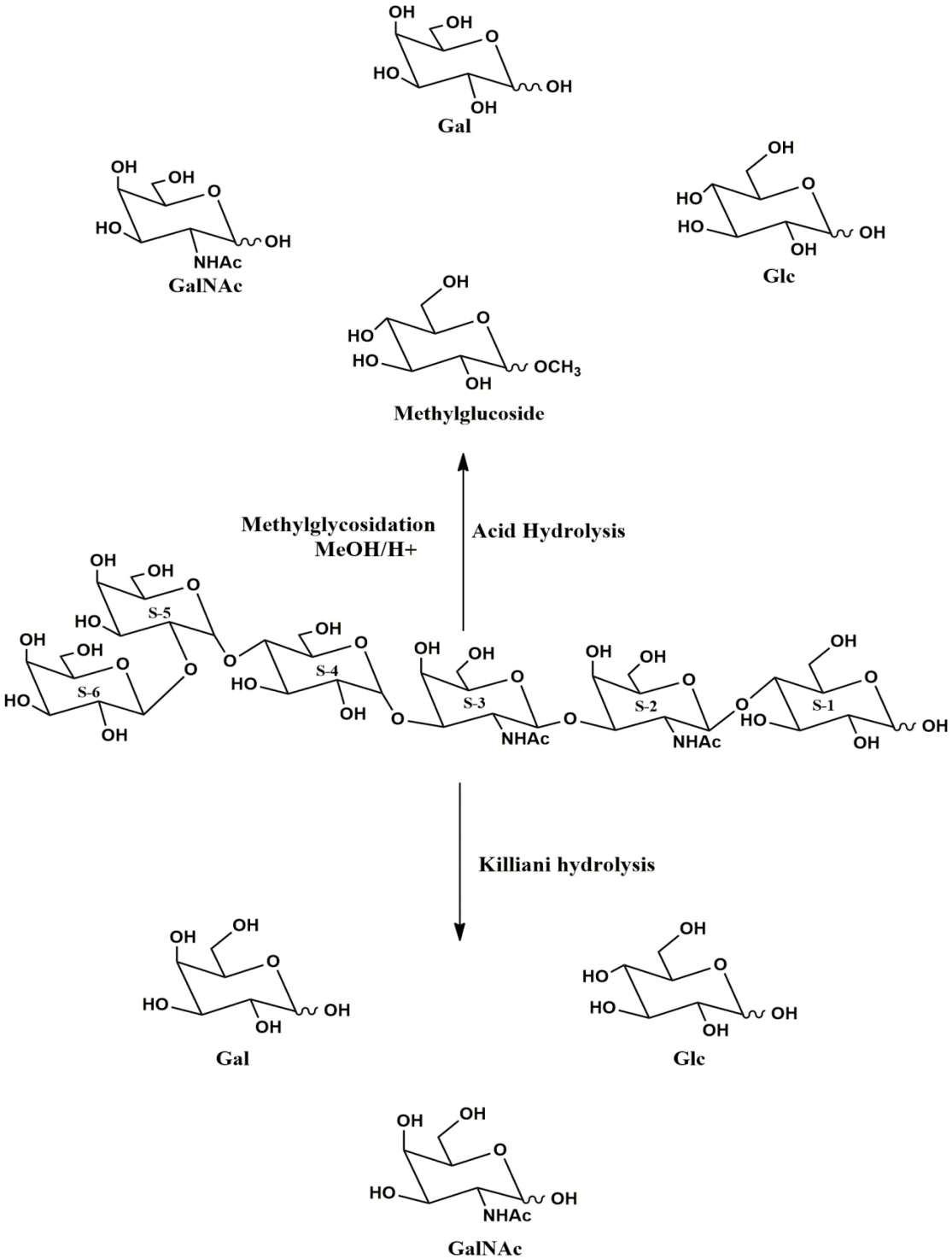


Figure 1. HSQC Spectrum of Thisose acetate in $CDCl_3$ at 300 MHz.



Scheme 1. Methylglycosidation/Acid Hydrolysis of Compound B, Thisose (Shukla et al; 2023).

The reducing nature of Thisose was further confirmed by its methyl glycosylation MeOH/H^+ followed by its acid hydrolysis, which led to the isolation of *α* and *β*-methylglucosides along with Glc, Gal and GalNHAc suggested the presence of glucose at the reducing end, for convenience all six monosaccharides were denoted as S-1, S-2, S-3, S-4, S-5 and S-6.

The monosaccharide constituents in Thisose were confirmed by its Killiani hydrolysis (Kiliani et al, 1930) under strong acidic condition, followed by paper chromatography and TLC. In this hydrolysis three spots were found identical with the authentic samples of Glc, Gal and GalNHAc by co-chromatography. Thus the Hexasaccharide contained three types of monosaccharide units i.e. Glc, Gal and GalNHAc.

Presence of six doublets for seven anomeric protons at δ6.16(1H), δ5.62(1H), δ5.34(1H), δ5.35(1H), δ4.60(1H) and δ4.48(2H) in the ^{1}H NMR of Thisose Acetate in $CDCl_3$ at 300 MHz. also confirmed in the presence of a Hexasaccharide in its reducing form.

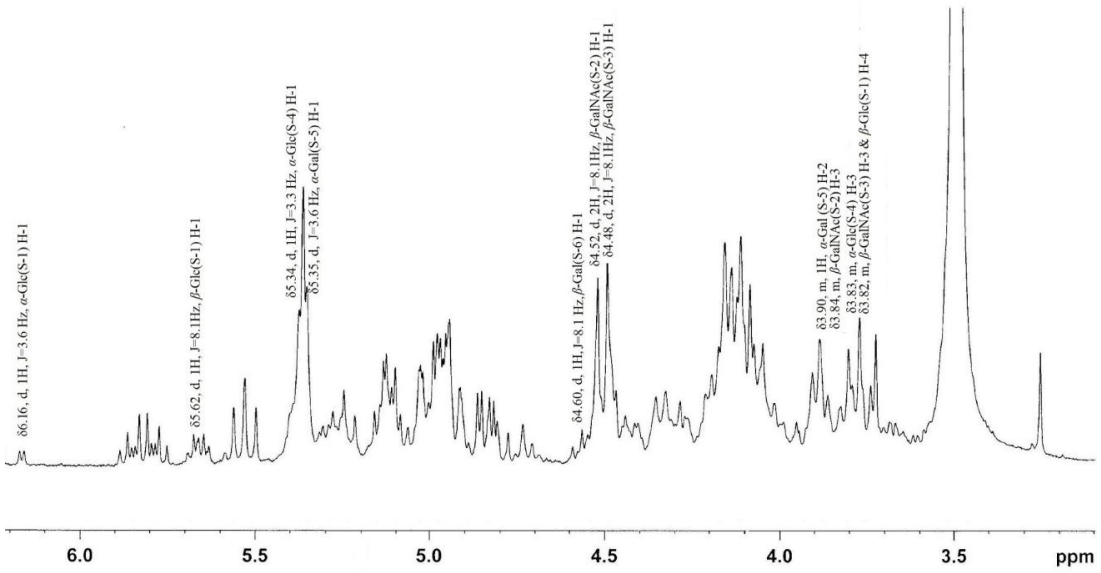


Figure 2. ^{1}H NMR Spectrum of Thisose acetate in $CDCl_3$ at 300 MHz.

Further the presence of five anomeric peaks for seven anomeric carbons at δ89.12 (2C), δ89.90 (1C), δ90.36 (1C), δ100.83 (1C) and δ101.20 (2C) in the ^{13}C NMR spectrum of acetylated Thisose in $CDCl_3$ at 300 MHz reconfirmed that Thisose was a Hexasaccharide in its reducing form.

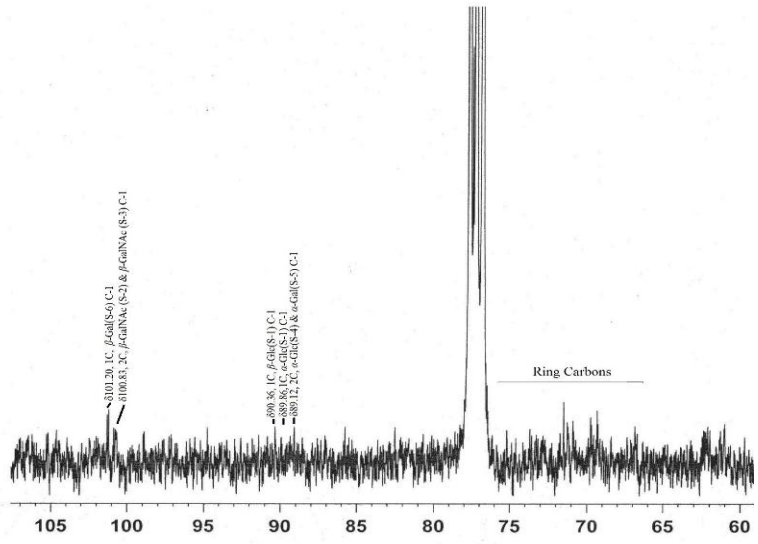


Figure 3. ^{13}C NMR Spectrum of Thisose acetate in $CDCl_3$ at 300 MHz.

The Hexasaccharide nature of Thisose was further supported by the presence of six anomeric proton doublets for seven anomeric protons at δ5.74(1H), δ5.22(2H), δ4.56 (1H), δ4.52(1H), δ4.46(1H) and δ4.38(1H) in ^{1}H NMR spectrum of Thisose in D_2O at 300 MHz.

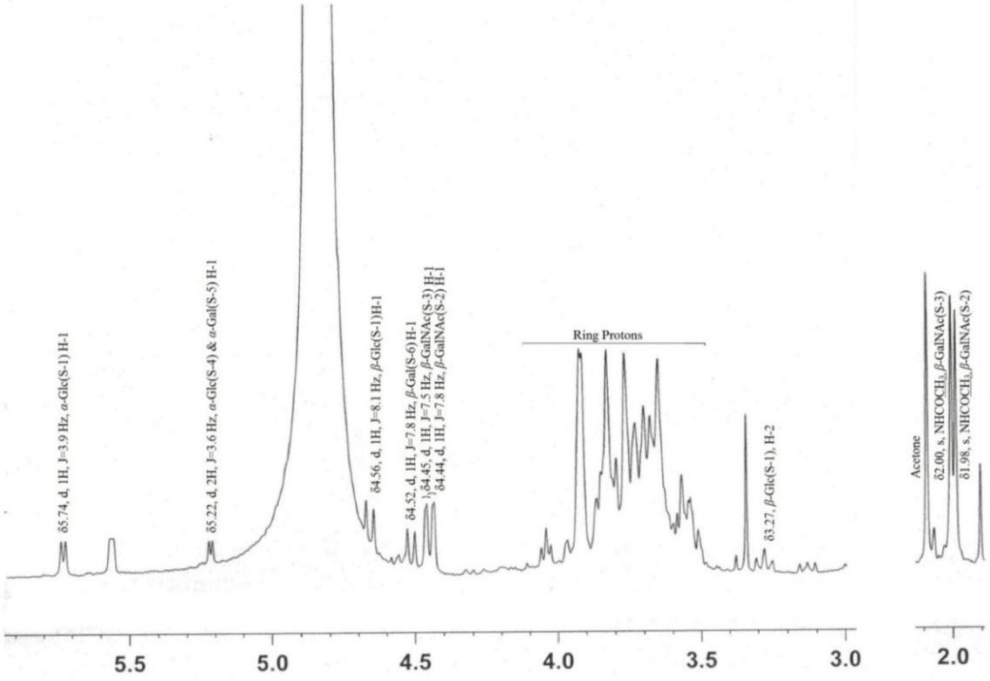


Figure 4. ^{1}H NMR Spectrum of Thisose in D_2O at 300 MHz.

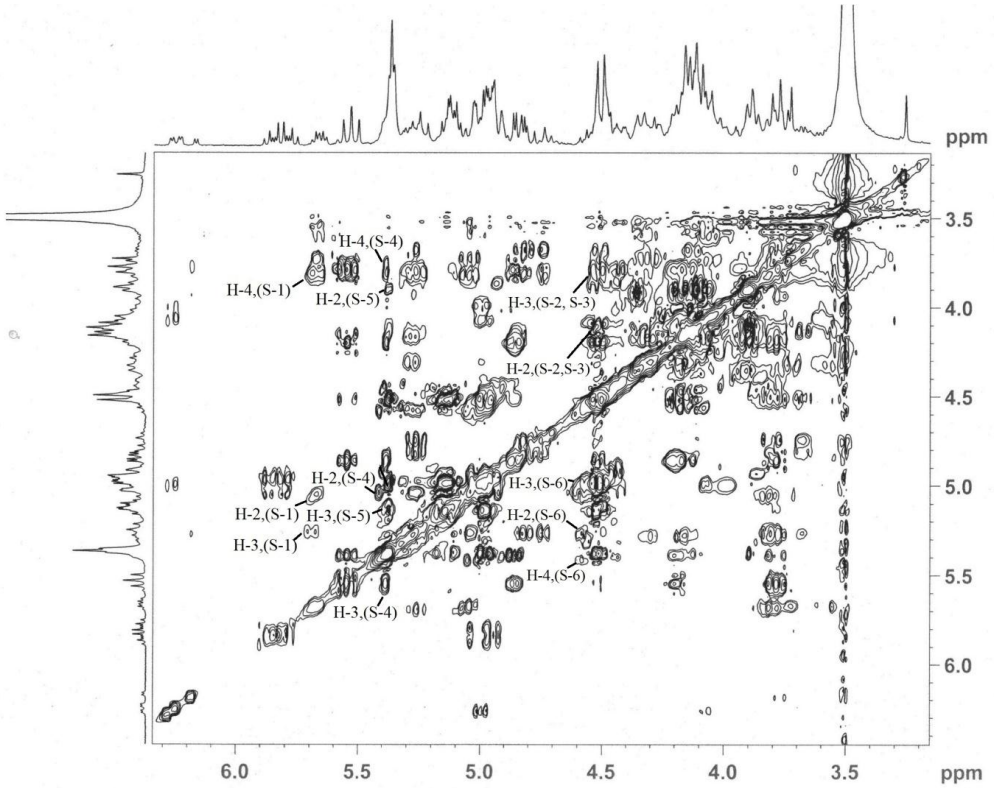


Figure 5. TOCSY Spectrum of Thisose acetate in $CDCl_3$ at 300 MHz.

^{1}H and ^{13}C NMR spectra justifying the seven anomeric signals for Hexasaccharide with total integral intensity of six anomeric proton/carbon. The molecular formula $C_{40}H_{68}O_{31}N_2$ was in agreement with mass ion peak obtained from ES-MS spectrum of Thisose which showed the highest mass ion peak at m/z 1095 and m/z 1072 due to $[M + Na]^+$ and $[M]^+$ respectively for a Hexasaccharide. The ^{1}H NMR spectrum of Thisose in D_2O at 300 MHz contain two anomeric doublets at δ5.74 (J=3.9Hz) and δ4.56 (J=8.1Hz) for *α* and *β* anomers of reducing monosaccharides (S-1) i.e. Glc.

Simultaneously ^{1}H NMR and ^{13}C NMR spectrum of Thisose acetate also showed downfield shifted *α* and *β* anomeric proton and carbon of reducing monosaccharides (S-1) i.e. Glc (S-1) at δ6.16, δ5.62 and δ89.86, δ90.36 respectively. The anomeric protons signal present at δ5.62 assigned for *β*-Glc (S-1) in TOCSY Spectrum of Thisose acetate contains three cross peaks at δ5.62 × 3.82, δ5.62 × 5.00 and δ5.62 × 5.22.

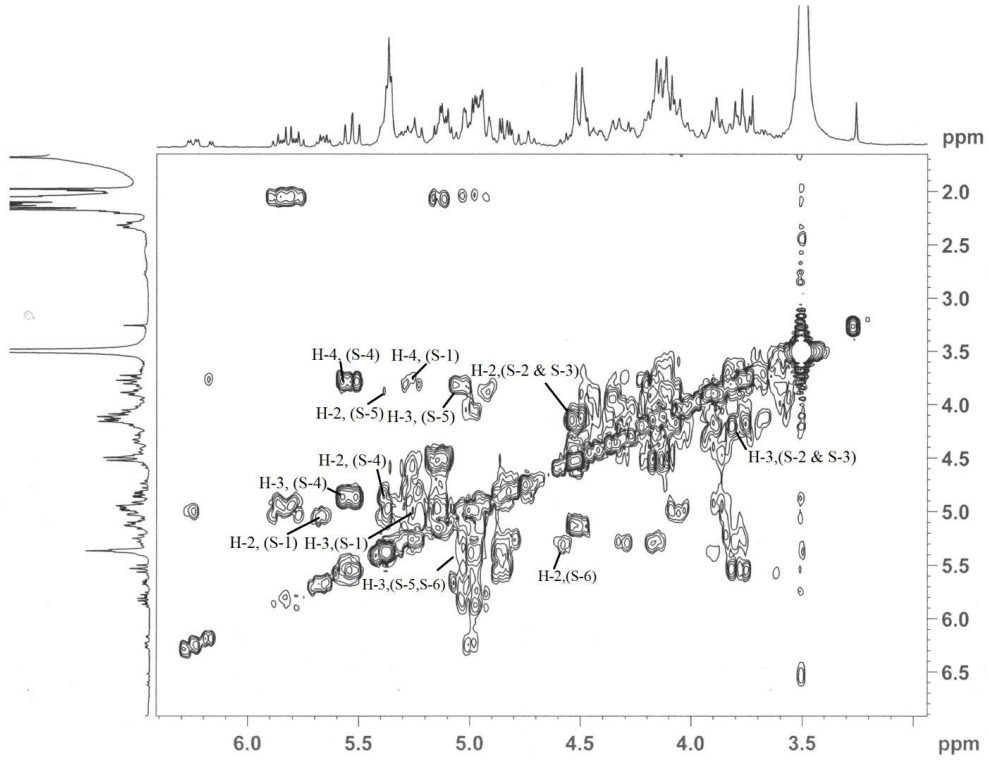


Figure 6. COSY Spectrum of Thisose acetate in $CDCl_3$ at 300 MHz.

Table 1. ^{1}H NMR values of Compound B, Thisose in D_2O and $CDCl_3$ at 300MHz.

Moieties	In D_2O		In $CDCl_3$	
	^{1}H NMR(δ)	Coupling constant(J)	^{1}H NMR(δ)	Coupling constant(J)
α-Glc(S-1)	5.74	3.9Hz	6.16	3.6Hz
β-Glc(S-1)	4.56	8.1Hz	5.62	8.1Hz
β-GalNHAc(S-2)	4.38	7.8Hz	4.48	8.1Hz
β-GalNHAc(S-3)	4.45	7.5Hz	4.52	8.1Hz
α-Glc(S-4)	5.22	3.6Hz	5.34	3.3Hz
α-Gal(S-5)	5.22	3.6Hz	5.35	3.6Hz
β-Gal(S-6)	4.52	7.8Hz	4.60	8.1Hz

The chemical shift of the cross peak at δ5.62 × 3.82 suggested that in sugar S-1 only one position was available for glycosidic linkage by the next monosaccharide unit, which was later identified as H-4 of S-1 by COSY spectrum of Thisose acetate. Further the presence of another anomeric proton doublet at δ4.38 (J=7.8Hz) along with singlet of three proton of amide methyl at δ1.98 in ^{1}H NMR spectrum of Thisose in D_2O and anomeric proton doublet at δ4.48 (J=8.1Hz) in ^{1}H NMR spectrum of Thisose acetate showed the presence of *β*-GalNHAc (S-2) in Thisose. The presence of anomeric proton signals of Glc and GalNHAc along with a triplet at δ3.27 (Bush C A, 1988) for H-2 of *β*-Glc suggested the presence of Lactose type of structure in which Gal was replaced with GalNHAc i.e. GalNHAc-(1→4)-Glc at the reducing end of Thisose hence confirming the (1→4) linkage between S_2 and S_1. The appearance of H-4 signal of *β*-Glc at δ3.82 in the ^{1}H NMR of Thisose acetate also confirmed the (1→4) linkage between S_2 and S_1 which was assigned by COSY and TOCSY spectrum of Thisose acetate. This linkage was also supported by the presence of cross peak at δ3.82 × 76.95 in glycosidic linkage region of HSQC spectrum of Thisose acetate in $CDCl_3$. Larger coupling constant (J), 7.8Hz of anomeric proton at δ4.48 confirmed a *β* glycosidic linkage between S-2 & S-1. The anomeric proton doublet at δ4.48 assigned for *β*-GalNHAc (S-2) in ^{1}H NMR spectrum of Thisose acetate contain three cross peaks at δ4.48×3.84, δ4.48×4.10 and 4.48×4.96 in its TOCSY spectrum. The chemical shift value of the cross peaks at δ4.48×3.84 and δ4.48×4.10 suggested that in sugar (S-2) two positions may be available for substitution which was later identified as H-2 and H-3 of S-2 by COSY spectrum of Thisose acetate.

Table 2. Anomeric proton/carbon assignments in Thisose Acetate by HSQC Spectrum.

Anomeric Proton × Anomeric Carbon	**Cross-peaks of Anomeric Proton × Anomeric Carbon**
H-1(S-1α) × C-1(S-1α)	δ6.16 × 89.86
H-1(S-1β) × C-1(S-1β)	δ5.62 × 90.36
H-1(S-2) × C-1(S-2)	δ4.48 × 100.83
H-1(S-3) × C-1(S-3)	δ4.52 × 100.83
H-1(S-4) × C-1(S-4)	δ5.34 × 89.12
H-1(S-5) × C-1(S-5)	δ5.35 × 89.12
H-1(S-6) × C-1(S-6)	δ4.60 × 101.20

Further it was confirmed by COSY spectrum of Thisose Acetate that H-2 position of *β*-GalNHAc (S-2) at δ4.10 was substituted by –NHAc group and the multiplet present at δ3.84 in ^{1}H NMR spectrum of Thisose acetate suggested that the H-3 of *β*-GalNHAc (S-2) was available for glycosidation by the next monosaccharide moiety (S-3). Further another anomeric proton doublet appeared at δ4.48 (d, J=8.1Hz) in the ^{1}H NMR spectrum of Thisose acetate gave its complimentary signal at δ100.83 in its HSQC spectrum in $CDCl_3$ at 300 MHz. The chemical shift value of anomeric carbon and anomeric proton were having resemblance with literature value of anomeric chemical shift value of *β*-GalNHAc(S-3) hence S-3 monosaccharide was confirmed as *β*-GalNHAc(S-3) (Bush C A, 1988). Further the presence of *β*-GalNHAc(S-3) as next monosaccharide in Thisose was supported by appearance of anomeric proton signal at δ4.46 (J = 7.5Hz) along with singlet of amide methyl at δ2.00 ^{1}H NMR spectrum of Thisose in D_2O at 300 MHz.

Since it was ascertained by the COSY and TOCSY spectrum of Thisose acetate that the *β*-GalNHAc (S-2) has two vacant position i.e. H-2 and H-3, and it was already confirmed that H-2 of S-2 was linked with -NHAc hence the left over H-3 position of *β*-GalNHAc (S-2) at δ3.84 must be linked to *β*-GalNHAc (S-3).

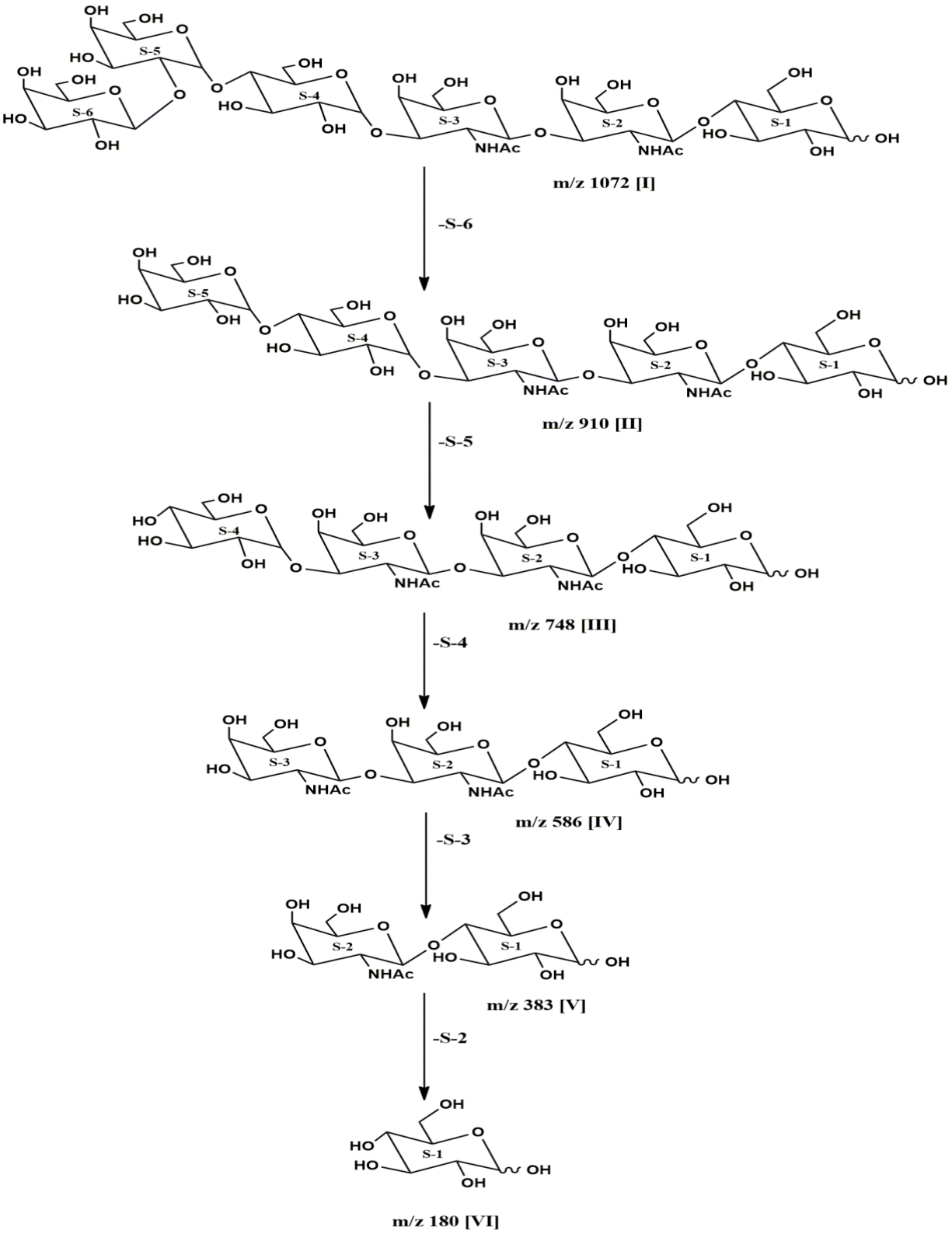
The (1→3) linkage between S-3 and S-2 was supported by the presence of H-3 signal of S-2 at δ3.84 in upfield region of ^{1}H NMR spectrum of Thisose acetate. This linkage was further supported by the presence of cross peak at δ3.84 × 76.95 in glycosidic region of HSQC spectrum of Thisose acetate 'b' in $CDCl_3$. The anomeric proton signal present at δ4.45 for GalNHAc (S-3) had a J value of 7.5Hz confirmed the *β*-(1→3) glycosidic linkage between S-2 and S-3. The anomeric proton doublet at δ4.48 assigned for *β*-GalNHAc (S-3) in ^{1}H NMR spectrum of Thisose acetate in $CDCl_3$ showed three cross peaks at δ4.48×3.82, δ4.48×4.10, and δ4.48×4.96 in its TOCSY spectrum, out of which proton signal arised at δ4.10 corresponded to H-2 position of *β*-GalNHAc (S-3) and another proton signal arised at δ3.82 represented the linkage region of *β*-GalNHAc (S-3) which was later identified as H-3 of *β*-GalNHAc (S-3) by COSY spectrum of Thisose acetate which was available for (1→3) glycosidic linkages by the next monosaccharide unit (S-4). Next anomeric proton doublet appeared at δ5.34 in ^{1}H NMR spectrum of Thisose acetate showed its complimentary signal at δ89.12 in its HSQC spectrum. The chemical shift value of anomeric carbon and anomeric proton were having resemblance with literature value of anomeric chemical shift value of *α*-Glc (S-4) hence S-4 monosaccharide was confirmed as *α*-Glc (S-4) (Bush, 1988). Presence of *α*-Glc (S-4) in Thisose was also confirmed by the appearance of anomeric proton doublet at δ5.22 in its ^{1}H NMR spectrum in D_2O. Since, H-3 position of S-3 was available for glycosidic linkage by the next monosaccharide unit (S-4), hence *α*-Glc (S-4) must be attached to H-3 of S-3. The (1→3) linkage between *α*-Glc (S-4) and *β*-GalNHAc (S-3) was further supported by the ^{1}H NMR spectrum of Thisose acetate in which the signal for H-3 of *β*-GalNHAc (S-3) appeared at δ3.82 which was later confirmed by COSY and TOCSY spectrum of Thisose acetate. This linkage was also supported by the presence of cross peak at δ3.82×76.95 in glycosidic region of HSQC spectrum of Thisose acetate in $CDCl_3$. The small coupling constant, J=3.6Hz of *α*-Glc (S-4) confirmed as *α*-(1→3) glycosidic linkage between *α*-Glc (S-4) and *β*-GalNHAc (S-3). The anomeric proton signal at δ5.34 assigned for *α*-Glc (S-4) in ^{1}H NMR spectrum of Thisose acetate gave four cross peaks at δ5.34×3.83, δ5.34 × 4.22, δ5.34 × 4.86 and δ5.34 × 5.56 in its TOCSY spectrum, out of which one cross peak arised at δ5.34 × 3.83 suggested that in sugar S-4, only one position was available for glycosidation by the next monosaccharide moiety (S-5), which was identified as H-4 of *α*-Glc (S-4) by COSY spectrum of Thisose acetate. The next anomeric proton signal which appeared at δ5.35 (J=3.6Hz) in the ^{1}H NMR spectrum of Thisose acetate had its complimentary signal at δ89.12 in HSQC spectrum of Thisose acetate. The chemical shift value of anomeric carbon and anomeric proton were having resemblance with literature value of anomeric chemical shift value of *α*-Gal hence S-5 monosaccharide was confirmed as *α*-Gal (S-5) (Bush C A, 1988). Further the presence of *α*-Gal (S-5) (J=3.6Hz) as the next monosaccharide in Thisose was supported by the appearance of anomeric proton doublet at δ5.22 in the ^{1}H NMR spectrum of Thisose in D_2O at 300 MHz. Since it was ascertained by the TOCSY and COSY spectrum of Thisose acetate that the H-4 position of S-4 was available for glycosidic linkage with next monosaccharide unit (S-5), thus H-4 position of *α*-Glc (S-4) must be linked with *α*-Glc (S-5). This linkage was further supported by ^{1}H NMR spectrum of acetylated Thisose in which the signal for H-4 of S-4

appeared at δ3.83 which was later confirmed by COSY, TOCSY and HSQC spectrum of Thisose acetate. The (1→4) linkage between S-5 and S-4 was further confirmed by the presence of cross peak at δ3.83 × 76.95 in glycosidic region of HSQC spectrum of Thisose acetate in $CDCl_3$. The small coupling constant (J), 3.6Hz of Glc (S-5) confirmed the *α*-glycosidic linkage between S-5 and S-4. The anomeric proton signal assigned for *α*-Glc (S-5) at δ5.35 gave three cross peaks at δ5.35 × 3.90, δ5.35 × 4.96 and δ5.35 × 5.13 in its TOCSY spectrum. Out of which one cross peak arised at δ5.35 × 3.90 suggested the position of linkage, which was later identified as H-2 of *α*-Gal (S-5) by COSY spectrum of Thisose acetate, suggesting that the H-2 of *α*-Gal (S-5) was available for glycosidation by the next monosaccharide (S-6). Further next anomeric proton doublet appeared at δ4.60 in the ^1H NMR spectrum of Thisose acetate gave its complementary signal at δ101.20 in its HSQC spectrum in $CDCl_3$. The chemical shift values of anomeric carbon and anomeric proton were having resemblance with literature value of anomeric chemical shift value of *β*-Gal hence S-6 monosaccharide was confirmed as *β*-Gal (S-6) (Bush C A, 1988). Further the presence of *β*-Gal (S-6) (J=7.8Hz) as the next monosaccharide in Thisose was supported by the appearance of anomeric proton doublet at δ4.52 in the ^1H NMR spectrum of Thisose in D_2O at 300 MHz. The appearance of H-2 signal of S-5 at δ3.90 in the ^1H NMR spectrum of Thisose acetate suggested that *β*-Gal (S-6) may be linked to *α*-Gal (S-5). This linkage was also supported by the presence of cross peak at δ3.90 **×** 71.21 in glycosidic region of HSQC spectrum of Thisose acetate in $CDCl_3$. The coupling constant of anomeric signal (S-6) at δ4.52 with J value of 7.8 Hz confirmed the *β*-(1→2) configuration of the glycosidic linkage between S-6 and S-5. The anomeric proton signal assigned for *β*-Gal (S-6) at δ4.60 in ^1H NMR spectrum of Thisose acetate gave three cross peaks at δ4.60 **×** 4.96, δ4.60 **×** 5.05 and δ4.60 **×** 5.26 in its TOCSY spectrum since this anomeric proton does not showed any cross peak in the linkage region i.e. δ3.5-4.2ppm, hence confirmed that *β*-Gal (S-6) was present at non-reducing end and none of its -OH group were involved in glycosidic linkage, which was confirmed by the TOCSY and COSY spectrum of Thisose acetate in $CDCl_3$ at 300 MHz. All the ^1H NMR assignments for ring protons of monosaccharide units of Thisose were confirmed by COSY and TOCSY experiments. The positions of glycosidation in the oligosaccharide were confirmed by position of anomeric signals, Structure Reporter Group (Vliegenthart J F, 1983) and comparing the signals in ^1H and ^{13}C NMR of acetylated and deacetylated oligosaccharide. The glycosidic linkages in Thisose were assigned by the cross peaks for glycosidically linked carbons with their protons in the HSQC spectrum of acetylated Thisose. All signals obtained in ^1H and ^{13}C NMR of compound Thisose were in confirmity with the assigned structure and their position were confirmed by 2D NMR viz. COSY, TOCSY, HSQC experiments. Thus based on the pattern of chemical shifts of ^1H NMR, ^{13}C NMR, COSY, TOCSY and HSQC experiments it was interpreted that the compound was a Hexasaccharide having following structure as:

Gal-β-(1→2)-Gal-α-(1→4)-Glc-α-(1→3)-GalNHAc-β-(1→3)-GalNHAc-β-(1→4)-Glc

THISOSE

The ESI Mass Spectrometry data of Thisose not only confirmed the derived structure but also supported the sequence of monosaccharide in Thisose. The highest mass ion peaks were recorded at m/z 1072 and 1095 which were due to $[M]^+$ and $[M+Na]^+$ respectively,



Scheme 2. Mass fragmentation of Compound B, Thisose.

Hence confirming the molecular weight of Thisose as m/z 1072 and which was in agreement with its molecular formula.$C_{40}H_{68}O_{31}N_2$ Further the fragmentation was formed by H transfer spectrum in the oligosaccharide accompanied by the elimination of terminal sugar less water. The Thisose at m/z 1072 (I) fragmented to give mass ion at m/z 910 (II) [1072-S_6], this fragment was arised due to the loss of terminal *β*-Gal (S-6) moiety from hexasaccharide indicating the presence of *β*-Gal (S-6) at the non-reducing end. It was further fragmented to give mass ion peak at m/z 748 (III) [910-S_5] which was due to loss of *α*-Gal (S-5) moiety from pentasaccharide. The fragment of m/z 748 further gave mass ion peak at m/z 586 (IV) [748-S_4] which was due to loss of *α*-Glc (S-4) moiety from the tetrasaccharide.

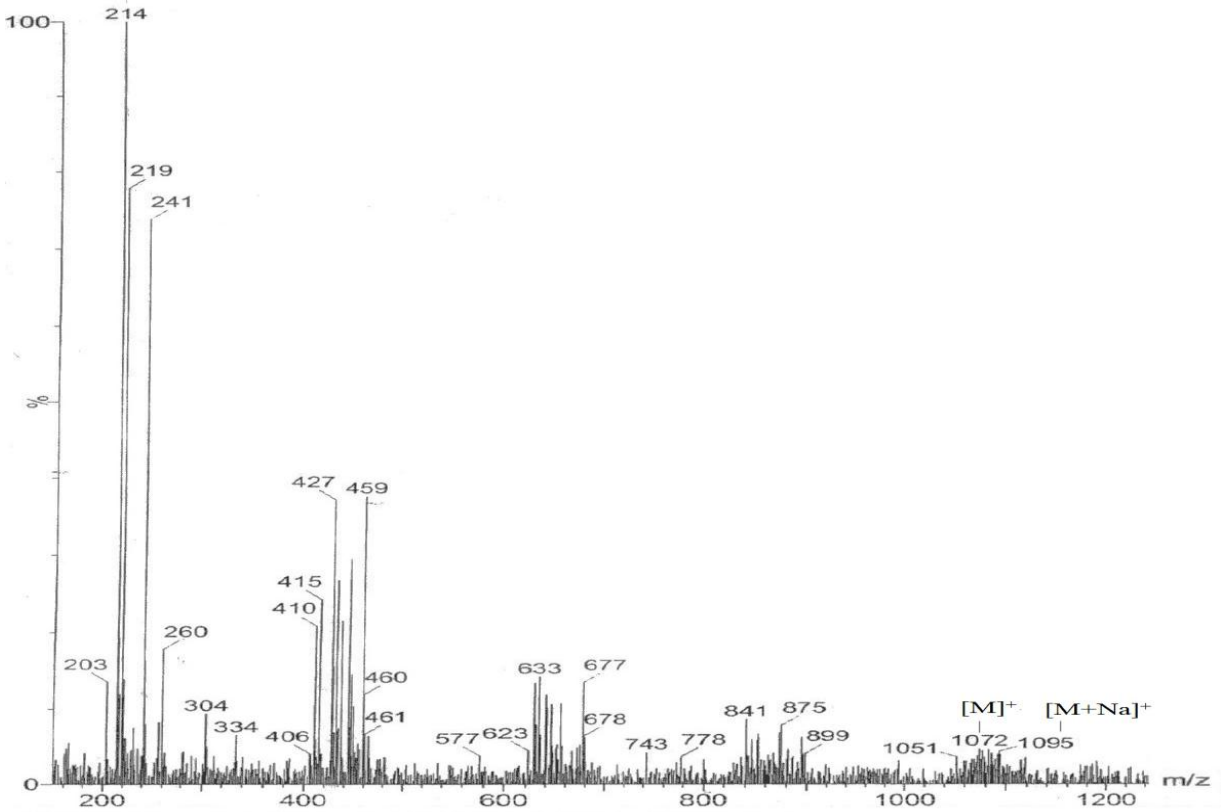
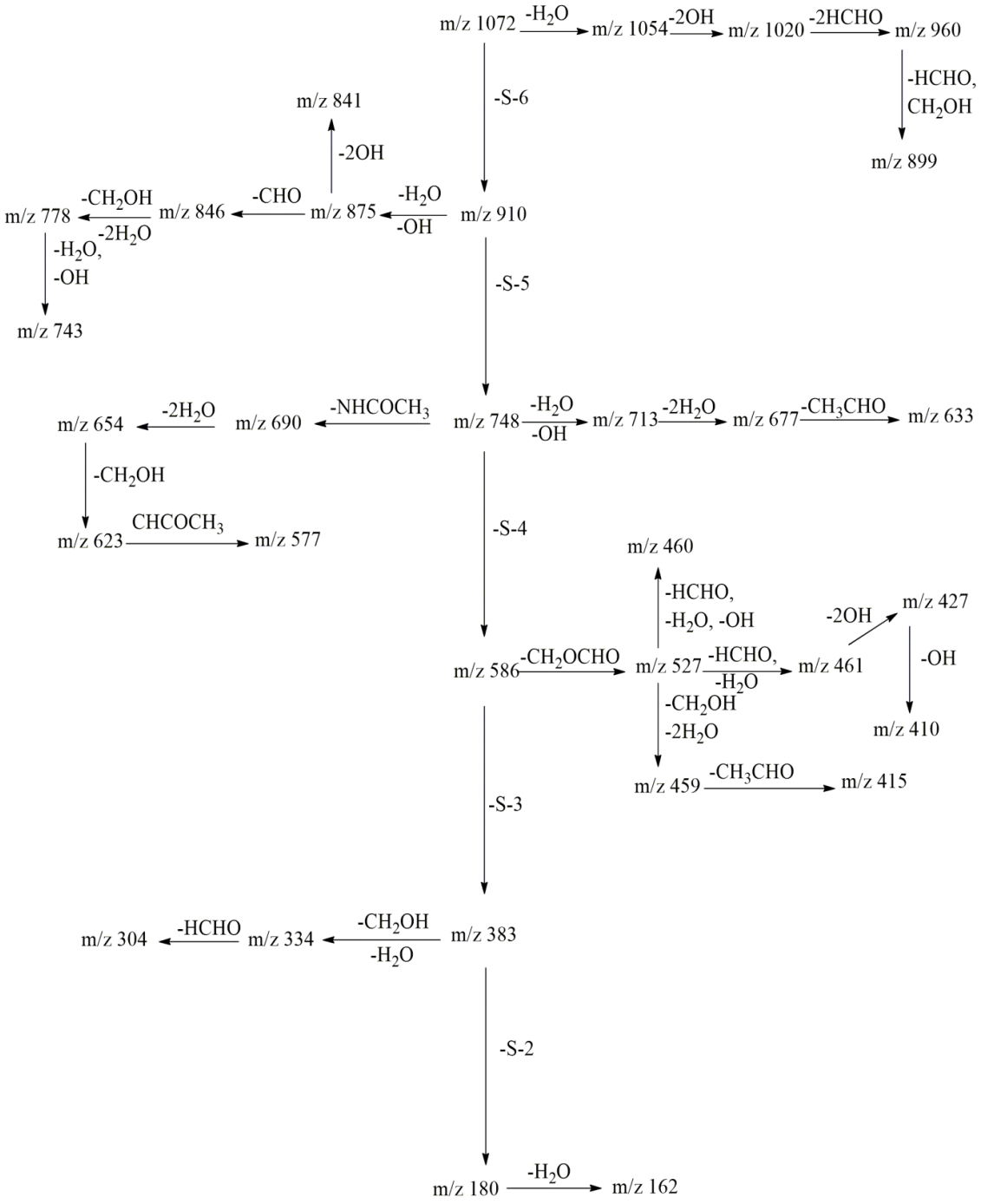


Figure 7. ES-MS Spectrum of compound B, Thisose.

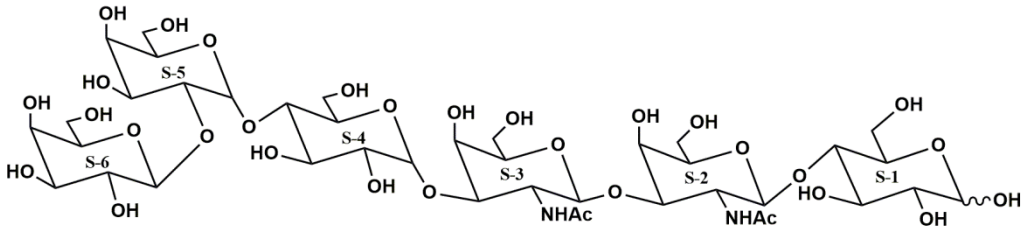
This fragment of 586 further fragmented to give mass ion peak at m/z 383 (V) [586-S_3] which was due to loss of *β*-GalNHAc (S-3) moiety from the trisaccharide. This disaccharide unit again fragmented to give mass ion peak at m/z 180 (VI) [383-S_2], which was due to loss of *β*-GalNHAc (S-2) moiety from disaccharide. The other fragmentation pathway in ES Mass spectrum of Compound B m/z 1072 shows the mass ion peak at 1054[1072-H_2O], 1020[1054-2OH], 960[1020-2HCHO], 910[1072-S-6], 899[960-HCHO, CH_2OH], 875[910-H_2O,-OH], 846[875-CHO], 841[875-2OH], 778[846-CH_2OH, $2H_2O$], 748[910-S-5], 743[778-H_2O, -OH], 713[748-H_2O,-OH], 690[748-$NHCOCH_3$], 677[713-$2H_2O$], 654[690-$2H_2O$], 633[677-CH_3CHO], 623[654-CH_2OH], 586[748-S-4], 577[623-$CHCOCH_3$], 527[586-CH_2OCHO**.**



Scheme 3. ES-MS fragments of Compound B, Thisose.

461[527-HCHO, -H_2O], 460[527-HCHO,-H_2O,-OH], 459[527-CH_2OH,-$2H_2O$], 427[461-2OH], 415[459-CH_3CHO], 410[427-OH], 383[586-S-3], 334[383-CH_2OH,-H_2O], 304[334-HCHO], 180[383-S-2] and 162[180-H_2O]. Based on result obtained from chemical degradation/acid hydrolysis, Chemical transformation, Electro spray mass spectrometry and 1D NMR viz. ^{1}H NMR, ^{13}C NMR and 2D NMR viz. COSY, TOCSY and HSQC spectra of Thisose acetate and Thisose, the structure and sequence of isolated Novel oligosaccharide Thisose structure was deduced as:

Gal-β-(1→2)-Gal-α-(1→4)-Glc-α-(1→3)-GalNHAc-β-(1→3)-GalNHAc-β-(1→4)-Glc



THISOSE

CONCLUSION

In the present study, the result obtained from chemical degradation, chemical transformation and data generated from 1H, 13C, COSY, TOCSY, HSQC and HMBC along with mass spectrometry were combined and it was concluded that Thisose was a novel hexasaccharide isolated from Rathi cow milk and was reported for first time from any of the natural or synthetic source. It was made up of Glc, Gal, GlcNHAc and GalNHAc with varied 1-4, 1-3 and 1-2 glycosidic linkage having α and β configurations.

ACKNOWLEDGEMENTS

Authors are thankful to Prof. Anil Mishra, Head Department of Chemistry for providing lab and library facilities for completing this work.

REFERENCES

Ling, E.R., Kon, S.K. and Porter, J.W. (1961). The Composition of Milk and the Nutritive Value of its Components. *The Mammary Gland and Its Secretion*. 2: 195-26.

Kumar, K., Singh, R. and Deepak, D. (2018). DFT Studies and Structure Elucidation of Novel Oligosaccharide from Camel Milk. *Chemistry & Biology Interface*. 8(2).

Churms S.C. (1996). Recent progress in carbohydrate separation by high-performance liquid chromatography based on hydrophilic interaction, *Journal of Chromatography* A, 720 (1–2).

Das, M., and Dasgupta, D. (1998). Pseudo-affinity column chromatography based rapid purification procedure for T7 RNA polymerase. *Preparative Biochemistry & Biotechnology*, 28(4), 339–348.

Davis, D.G. and Bax, A. (1985). Assignment of complex proton NMR spectra via two-dimensional homonuclear Hartmann-Hahn spectroscopy. *Journal of the American Chemical Society*, 107, 2820-2821.

Kay, L., Keifer, P.A., and Saarinen, T. (1992). Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *Journal of the American Chemical Society*, 114, 10663-10665.

Barr, J.R., Anumula, K.R., Vettese, M.B., Taylor, P.B., and Carr, S.A. (1991). Structural classification of carbohydrates in glycoproteins by mass spectrometry and high-performance anion-exchange chromatography. *Analytical Biochemistry*, 192 1, 181-92.

Chauhan, D.D.A.P.S., Shukla M., Chauhan, S. and Deepak, D. (2024). Isolation of Biologically Active Oligosaccharide 'Rathose' from Rathi Cow Milk and Their Structure Elucidation by 2D NMR. In press *J. Biol. Chem. Research.*

Kumar, K., Srivastava, K.A. and Deepak D. (2016). Isolation of a Novel Oligosaccharide from Goat Milk. *J. Biol. Chem. Research*. Vol. 33, No. 1: 381-387.

Khan, M., Mishra, A. and Deepak, D. (2019). Isolation and Structure elucidation of novel Dodecasaccharide Uruose from Lal-Muha Cow milk. *Trends in Carbohydrate Research*, 11(3): 1-19.

Dubois, M., Gilles, K.A., Hamilton, J.K., Rebers, P.A. and Smith, F. (1956). Colorimetric Method for Determination of Sugars and Related Substances. *Analytical Chemistry*, 28, 350-356.

Partridge, S.M. (1949). Aniline hydrogen phthalate as a spraying reagent for chromatography of sugars. *Nature*. 164: 443.

Feigl, F. (1975). Spot Tests in Organic Analysis. *Elsevier*. 7: 337.

Warren, L. (1960). Thiobarbituric Acid Spray Reagent for Deoxy Sugars and Sialic Acids. *Nature* 186, 237.

Shukla, M., Sharma, M., Singh, R. and Deepak, D. (2023). Structure Characterization of Novel Milk Hexasaccharide 'Omedose' from Camel Milk. *J. Biol. Chem. Research.* Vol. 40 No.1: 39-52.

Kiliani, H. (1930). Über Digitalinum verum. *European Journal of Inorganic Chemistry*, 63, 2866-2869.

Bush C.A. (1988). High Resolution NMR in the determination of structure in complex carbohydrates. *Bull Magn Reson.* 10(3/4):73-78.

Vliegenthart, J.F., Dorland, L. and van Halbeek, H. (1983). High-resolution, 1H-nuclear magnetic resonance spectroscopy as a tool in the structural analysis of carbohydrates related to glycoproteins. *In Advances in Carbohydrates Chemistry and Biochemistry* 1; 41: 209-374) Academic press.

Corresponding author: Prof. Sarita Chauhan, Department of Chemistry, JNM P.G. College, Lucknow, U.P. India 226004
Email: drsaritachauhan3@gmail.com