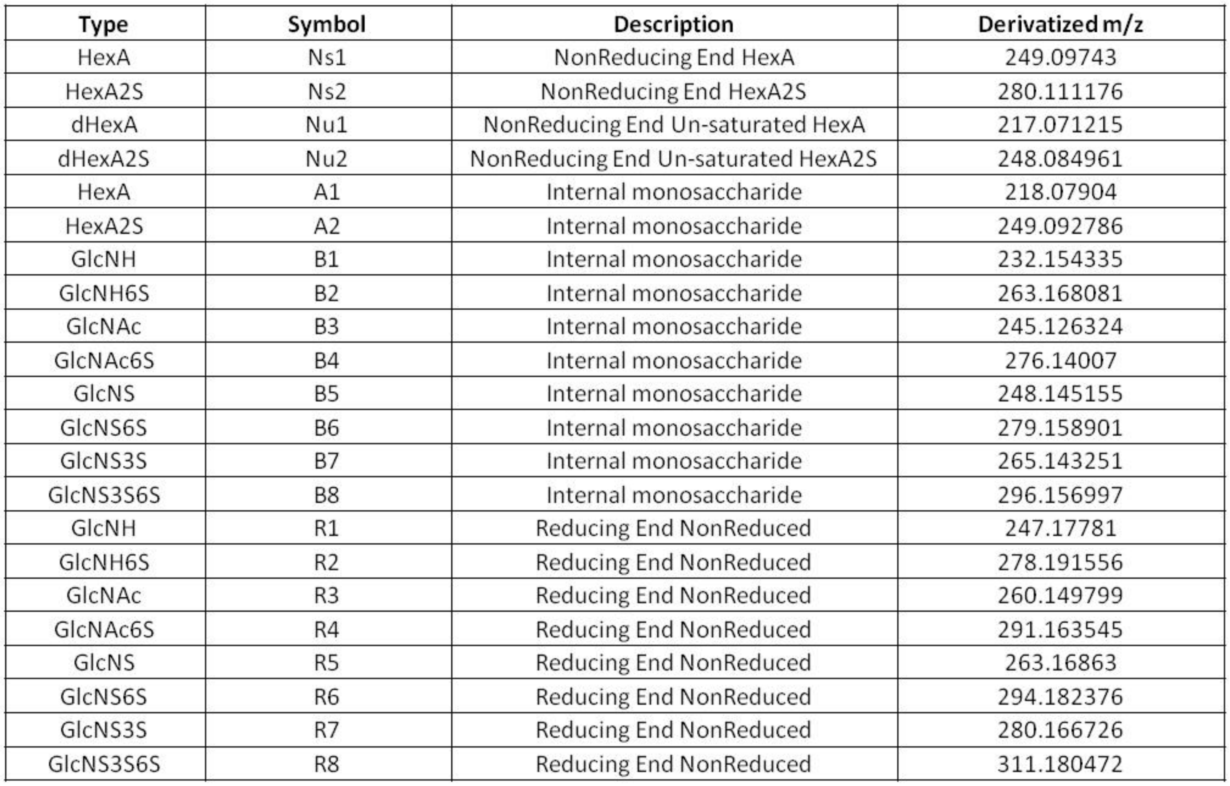
Supplementary Data

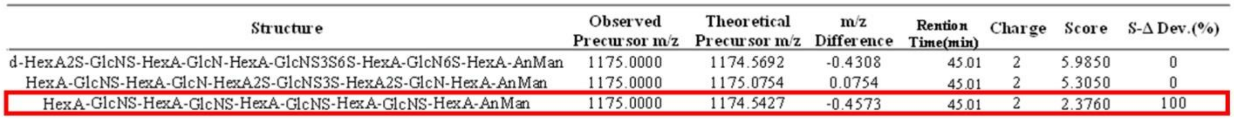
Calculated m/z for derivatized precursor M^+	Compound # and sequences
1117.596	**1a**: IdoA2S-GlcNAc-GlcA-GlcNAc-$(CH_2)_5NH_2$
	1b: IdoA-GlcNAc6S-GlcA-GlcNAc-$(CH_2)_5NH_2$
	1c: GlcA-GlcNAc-IdoA2S-GlcNAc-$(CH_2)_5NH_2$
	1d: GlcA-GlcNAc-IdoA-GlcNAc6S-$(CH_2)_5NH_2$
	1e: GlcA-GlcNAc-GlcA2S-GlcNAc-$(CH_2)_5NH_2$
1123.633	**2a**: IdoA2S-GlcNS-GlcA-GlcNS-$(CH_2)_5NH_2$
	2b: IdoA-GlcNS6S-GlcA-GlcNS-$(CH_2)_5NH_2$
	2c: GlcA-GlcNS-IdoA2S-GlcNS-$(CH_2)_5NH_2$
	2d: GlcA-GlcNS-IdoA-GlcNS6S-$(CH_2)_5NH_2$
	2e: GlcA-GlcNS-GlcA2S-GlcNS-$(CH_2)_5NH_2$
1148.609	**3a**: IdoA2S-GlcNAc6S-GlcA-GlcNAc-$(CH_2)_5NH_2$
	3b: IdoA-GlcNAc6S-GlcA-GlcNAc6S-$(CH_2)_5NH_2$
	3c: GlcA-GlcNAc-IdoA2S-GlcNAc6S-$(CH_2)_5NH_2$
	3d: GlcA-GlcNAc6S-IdoA-GlcNAc6S- $(CH_2)_5NH_2$
	3e: GlcA-GlcNAc-GlcA2S-GlcNAc6S-$(CH_2)_5NH_2$
	3f: GlcA-GlcNAc6S-GlcA-GlcNAc6S- $(CH_2)_5NH_2$
1154.648	**4a**: IdoA2S-GlcNS6S-GlcA-GlcNS-$(CH_2)_5NH_2$
	4b: IdoA-GlcNS6S-GlcA-GlcNS6S-$(CH_2)_5NH_2$
	4c: GlcA-GlcNS-IdoA2S-GlcNS6S-$(CH_2)_5NH_2$
	4d: GlcA-GlcNS-GlcA2S-GlcNS6S-$(CH_2)_5NH_2$
	4e: GlcA-GlcNS6S-GlcA-GlcNS6S- $(CH_2)_5NH_2$

Table S1. Sequences for four sets of synthetic tetrasaccharides with alkyl linker at the reducing end. GlcNAc, N-acetylated glucosamine; GlcNS, N-sulfated glucosamine; GlcA, glucuronic acid; IdoA, iduronic acid; 2S, 2-O-sulfate; 6S, 6-O-sulfate.



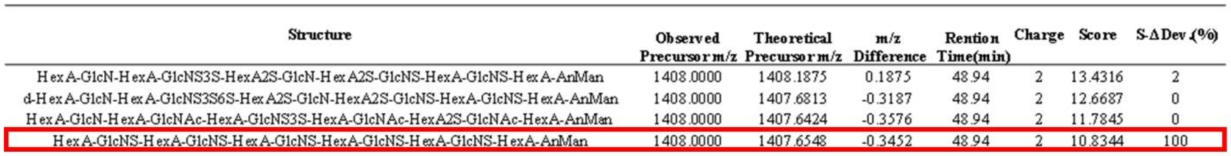
Type	Symbol	Description	Derivatized m/z
HexA	Ns1	NonReducing End HexA	249.09743
HexA2S	Ns2	NonReducing End HexA2S	280.111176
dHexA	Nu1	NonReducing End Un-saturated HexA	217.071215
dHexA2S	Nu2	NonReducing End Un-saturated HexA2S	248.084961
HexA	A1	Internal monosaccharide	218.07904
HexA2S	A2	Internal monosaccharide	249.092786
GlcNH	B1	Internal monosaccharide	232.154335
GlcNH6S	B2	Internal monosaccharide	263.168081
GlcNAc	B3	Internal monosaccharide	245.126324
GlcNAc6S	B4	Internal monosaccharide	276.14007
GlcNS	B5	Internal monosaccharide	248.145155
GlcNS6S	B6	Internal monosaccharide	279.158901
GlcNS3S	B7	Internal monosaccharide	265.143251
GlcNS3S6S	B8	Internal monosaccharide	296.156997
GlcNH	R1	Reducing End NonReduced	247.17781
GlcNH6S	R2	Reducing End NonReduced	278.191556
GlcNAc	R3	Reducing End NonReduced	260.149799
GlcNAc6S	R4	Reducing End NonReduced	291.163545
GlcNS	R5	Reducing End NonReduced	263.16863
GlcNS6S	R6	Reducing End NonReduced	294.182376
GlcNS3S	R7	Reducing End NonReduced	280.166726
GlcNS3S6S	R8	Reducing End NonReduced	311.180472

Table S2. The theoretical HS GAG sequence composition. Accurate mass is calculated for each monosaccharide unit after chemical derivatization. For even-numbered oligosaccharides, the uronic acid residue at the non-reducing end can be one of Ns1, Ns2, Nu1 or Nu2, where Ns1 and Ns2 indicates saturated uronic acids, and Nu1 and Nu2 indicate unsaturated uronic acids as would result from lyase cleavage (see Supplementary Figure S2). The A1-A2 group and the B1-B8 group represent for internal uronic acid residues and glucosamine residues respectively, while the R1-R8 group represents the glucosamine residue at the reducing end. Reduction of the reducing end and reducing end tags can also be input by the user, and will modify the mass of the glucosamine residue at the reducing end.



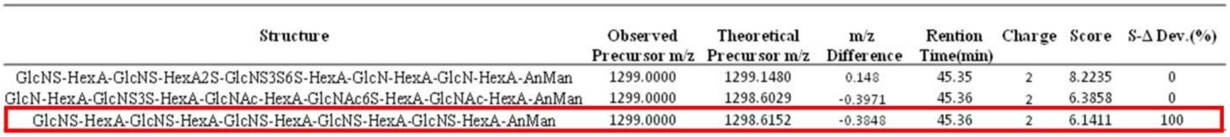
Structure	Observed Precursor m/z	Theoretical Precursor m/z	m/z Difference	Rention Time(min)	Charge	Score	S-Δ Dev.(%)
d-HexA2S-GlcNS-HexA-GlcN-HexA-GlcNS3S6S-HexA-GlcN6S-HexA-AnMan	1175.0000	1174.5692	-0.4308	45.01	2	5.9850	0
HexA-GlcNS-HexA-GlcN-HexA2S-GlcNS3S-HexA2S-GlcN-HexA-AnMan	1175.0000	1175.0754	0.0754	45.01	2	5.3050	0
HexA-GlcNS-HexA-GlcNS-HexA-GlcNS-HexA-GlcNS-HexA-AnMan	1175.0000	1174.5427	-0.4573	45.01	2	2.3760	100

Table S3. GAG-ID results for [NS-decamer+2Na], theoretical derivatized *m/z* 1174.5427^{2+}



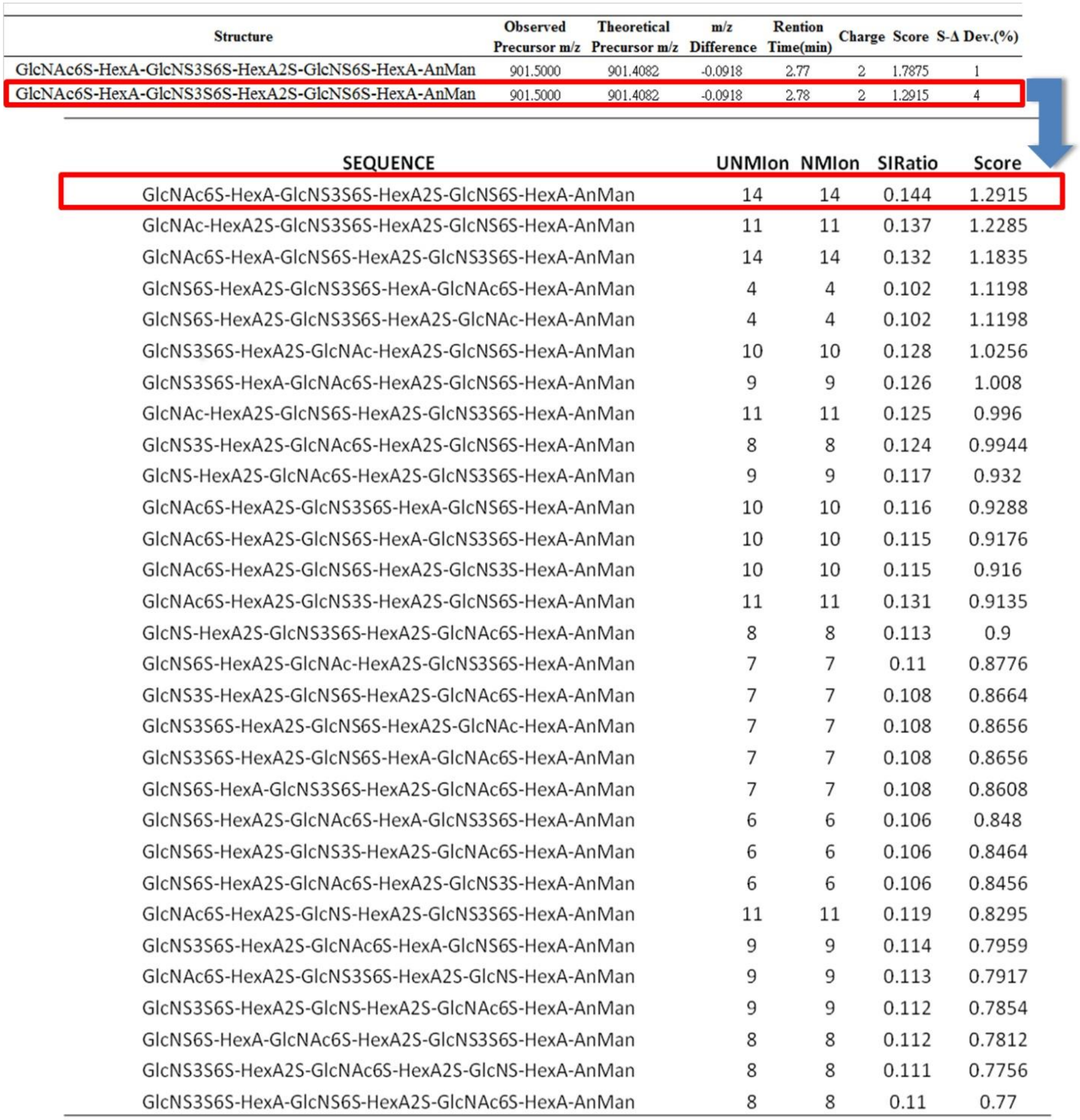
Structure	Observed Precursor m/z	Theoretical Precursor m/z	m/z Difference	Rention Time(min)	Charge	Score	S-ΔDev.(%)
HexA-GlcN-HexA-GlcNS3S-HexA2S-GlcN-HexA2S-GlcNS-HexA-GlcNS-HexA-AnMan	1408.0000	1408.1875	0.1875	48.94	2	13.4316	2
d-HexA-GlcN-HexA-GlcNS3S6S-HexA2S-GlcN-HexA2S-GlcNS-HexA-GlcNS-HexA-AnMan	1408.0000	1407.6813	-0.3187	48.94	2	12.6687	0
HexA-GlcN-HexA-GlcNAc-HexA-GlcNS3S-HexA-GlcNAc-HexA2S-GlcNAc-HexA-AnMan	1408.0000	1407.6424	-0.3576	48.94	2	11.7845	0
HexA-GlcNS-HexA-GlcNS-HexA-GlcNS-HexA-GlcNS-HexA-GlcNS-HexA-AnMan	1408.0000	1407.6548	-0.3452	48.94	2	10.8344	100

Table S4. GAG-ID results for [NS-dodecamer+2Na], theoretical derivatized *m/z* 1407.6548^{2+}



Structure	Observed Precursor m/z	Theoretical Precursor m/z	m/z Difference	Rention Time(min)	Charge	Score	S-Δ Dev.(%)
GlcNS-HexA-GlcNS-HexA2S-GlcNS3S6S-HexA-GlcN-HexA-GlcN-HexA-AnMan	1299.0000	1299.1480	0.148	45.35	2	8.2235	0
GlcN-HexA-GlcNS3S-HexA-GlcNAc-HexA-GlcNAc6S-HexA-GlcNAc-HexA-AnMan	1299.0000	1298.6029	-0.3971	45.36	2	6.3858	0
GlcNS-HexA-GlcNS-HexA-GlcNS-HexA-GlcNS-HexA-GlcNS-HexA-AnMan	1299.0000	1298.6152	-0.3848	45.36	2	6.1411	100

Table S5. GAG-ID results for [NS-undecamer+2Na], theoretical derivatized *m/z* 1298.6152^{2+}



Structure	Observed Precursor m/z	Theoretical Precursor m/z	m/z Difference	Rention Time(min)	Charge	Score	S-Δ Dev.(%)
GlcNAc6S-HexA-GlcNS3S6S-HexA2S-GlcNS6S-HexA-AnMan	901.5000	901.4082	-0.0918	2.77	2	1.7875	1
GlcNAc6S-HexA-GlcNS3S6S-HexA2S-GlcNS6S-HexA-AnMan	901.5000	901.4082	-0.0918	2.78	2	1.2915	4

SEQUENCE	UNMIon	NMIon	SIRatio	Score
GlcNAc6S-HexA-GlcNS3S6S-HexA2S-GlcNS6S-HexA-AnMan	14	14	0.144	1.2915
GlcNAc-HexA2S-GlcNS3S6S-HexA2S-GlcNS6S-HexA-AnMan	11	11	0.137	1.2285
GlcNAc6S-HexA-GlcNS6S-HexA2S-GlcNS3S6S-HexA-AnMan	14	14	0.132	1.1835
GlcNS6S-HexA2S-GlcNS3S6S-HexA-GlcNAc6S-HexA-AnMan	4	4	0.102	1.1198
GlcNS6S-HexA2S-GlcNS3S6S-HexA2S-GlcNAc-HexA-AnMan	4	4	0.102	1.1198
GlcNS3S6S-HexA2S-GlcNAc-HexA2S-GlcNS6S-HexA-AnMan	10	10	0.128	1.0256
GlcNS3S6S-HexA-GlcNAc6S-HexA2S-GlcNS6S-HexA-AnMan	9	9	0.126	1.008
GlcNAc-HexA2S-GlcNS6S-HexA2S-GlcNS3S6S-HexA-AnMan	11	11	0.125	0.996
GlcNS3S-HexA2S-GlcNAc6S-HexA2S-GlcNS6S-HexA-AnMan	8	8	0.124	0.9944
GlcNS-HexA2S-GlcNAc6S-HexA2S-GlcNS3S6S-HexA-AnMan	9	9	0.117	0.932
GlcNAc6S-HexA2S-GlcNS3S6S-HexA-GlcNS6S-HexA-AnMan	10	10	0.116	0.9288
GlcNAc6S-HexA2S-GlcNS6S-HexA-GlcNS3S6S-HexA-AnMan	10	10	0.115	0.9176
GlcNAc6S-HexA2S-GlcNS6S-HexA2S-GlcNS3S-HexA-AnMan	10	10	0.115	0.916
GlcNAc6S-HexA2S-GlcNS3S-HexA2S-GlcNS6S-HexA-AnMan	11	11	0.131	0.9135
GlcNS-HexA2S-GlcNS3S6S-HexA2S-GlcNAc6S-HexA-AnMan	8	8	0.113	0.9
GlcNS6S-HexA2S-GlcNAc-HexA2S-GlcNS3S6S-HexA-AnMan	7	7	0.11	0.8776
GlcNS3S-HexA2S-GlcNS6S-HexA2S-GlcNAc6S-HexA-AnMan	7	7	0.108	0.8664
GlcNS3S6S-HexA2S-GlcNS6S-HexA2S-GlcNAc-HexA-AnMan	7	7	0.108	0.8656
GlcNS3S6S-HexA2S-GlcNS6S-HexA-GlcNAc6S-HexA-AnMan	7	7	0.108	0.8656
GlcNS6S-HexA-GlcNS3S6S-HexA2S-GlcNAc6S-HexA-AnMan	7	7	0.108	0.8608
GlcNS6S-HexA2S-GlcNAc6S-HexA-GlcNS3S6S-HexA-AnMan	6	6	0.106	0.848
GlcNS6S-HexA2S-GlcNS3S-HexA2S-GlcNAc6S-HexA-AnMan	6	6	0.106	0.8464
GlcNS6S-HexA2S-GlcNAc6S-HexA2S-GlcNS3S-HexA-AnMan	6	6	0.106	0.8456
GlcNAc6S-HexA2S-GlcNS-HexA2S-GlcNS3S6S-HexA-AnMan	11	11	0.119	0.8295
GlcNS3S6S-HexA2S-GlcNAc6S-HexA-GlcNS6S-HexA-AnMan	9	9	0.114	0.7959
GlcNAc6S-HexA2S-GlcNS3S6S-HexA2S-GlcNS-HexA-AnMan	9	9	0.113	0.7917
GlcNS3S6S-HexA2S-GlcNS-HexA2S-GlcNAc6S-HexA-AnMan	9	9	0.112	0.7854
GlcNS6S-HexA-GlcNAc6S-HexA2S-GlcNS3S6S-HexA-AnMan	8	8	0.112	0.7812
GlcNS3S6S-HexA2S-GlcNAc6S-HexA2S-GlcNS-HexA-AnMan	8	8	0.111	0.7756
GlcNS3S6S-HexA-GlcNS6S-HexA2S-GlcNAc6S-HexA-AnMan	8	8	0.11	0.77

Table S6. GAG-ID results for [Arixtra-like heptamer+2Na], theoretical derivatized *m/z* 901.4082^{2+}

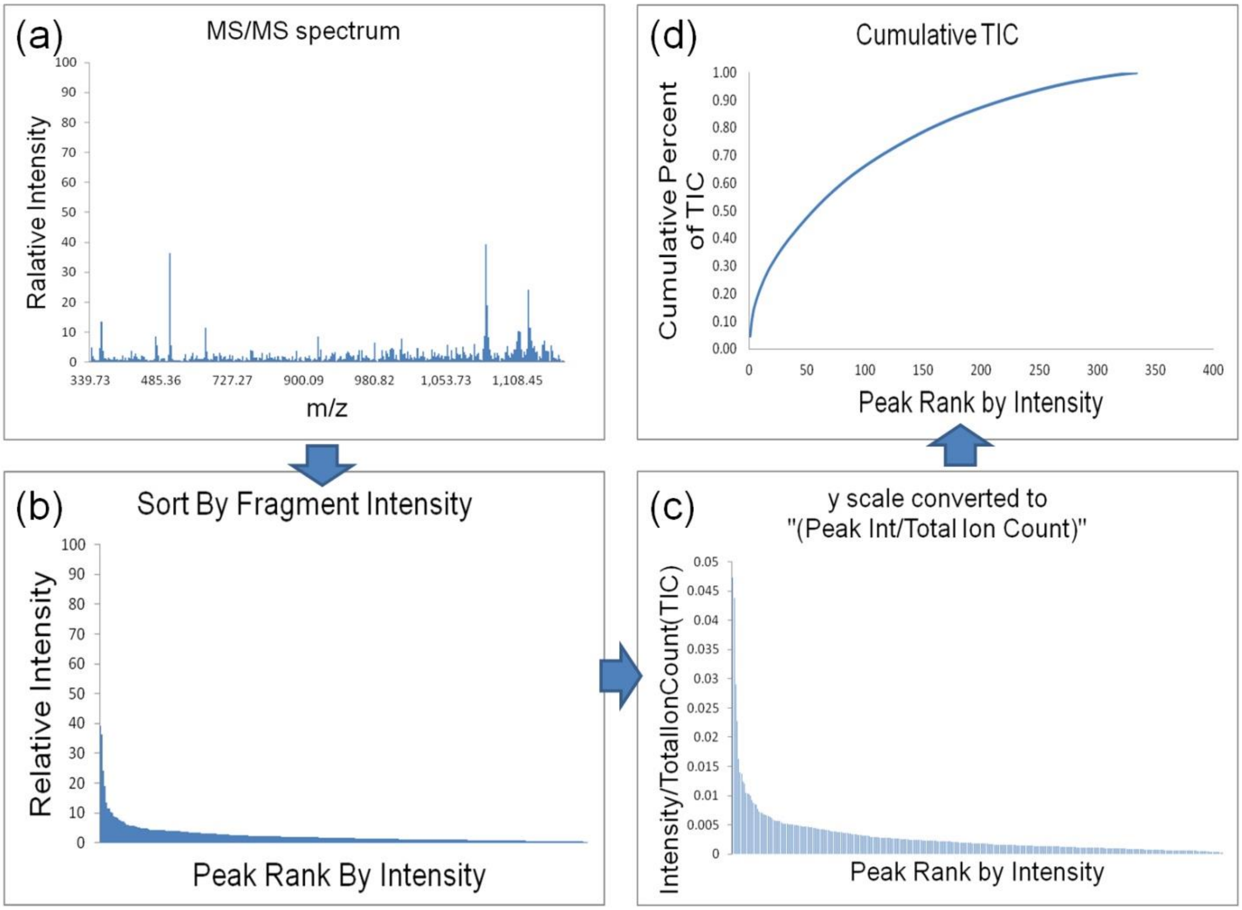


Figure S1 MS/MS spectrum preprocessing. (a) Taking an MS/MS spectrum as an example, all spectra acquired for scoring were preprocessed. (b) All peaks sorted by intensity in decreasing order were listed from the most intense to the least intense peak. (c) All peak intensities were divided by total ion count (TIC) to convert the y-axis to "peak intensity/TIC". (d) The cumulative intensity from the most intense peak to all other peaks was computed (represented here by the curve).

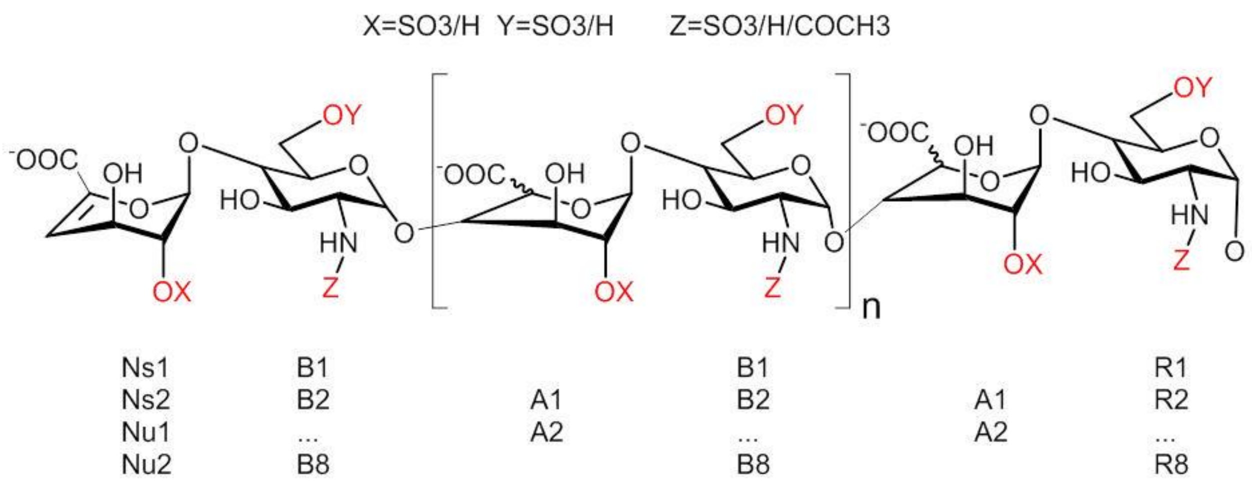


Figure S2 Monosaccharide unit included in HS sequences The *m/z* for each unit block with different modification was shown in Table S2. All possible sequences from tetramer to dodecamer were iterated through all combinations and then stored in GAG-DB.

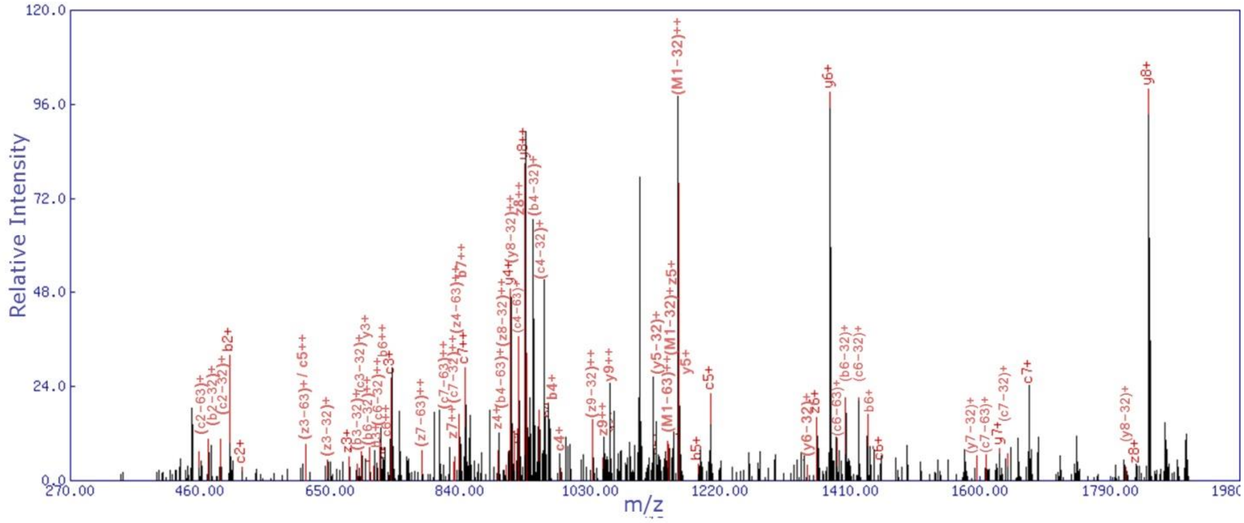


Figure S3 Automated spectrum annotation for NS-decamer.