Bioorganic & Medicinal Chemistry 20 (2012) 5085–5093

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Stochastic searches and NMR experiments on four Lewis A analogues: NMR experiments support some flexibility around the fucosidic bond

France-Isabelle Auzanneau [a,*], Trudy A. Jackson [b], Liang Liao [c]

[a] Department of Chemistry, University of Guelph, Guelph, Ontario, Canada N1G 2W1
[b] Present address: Skerrit's Pasture (gov't Works), St. John's, Antigua, Antigua and Barbuda
[c] Present address: Onyx Pharmaceuticals, 249 East Grand Avenue, South San Francisco, CA 94080, USA

ARTICLE INFO

Article history:
Received 11 June 2012
Revised 4 July 2012
Accepted 10 July 2012
Available online 20 July 2012

This paper is dedicated to Malcolm B. Perry (April 1930–June 2012) with much respect and admiration

Keywords:
Carbohydrates
Conformation analysis
NMR spectroscopy
Lewis A

ABSTRACT

We have compared the conformational behavior of three Le[a] analogues with that of Le[a] using stochastic searches (MOE2005) and selective ROESY experiments. In the analogues either or both the β-D-Gal and α-L-Fuc residues were replaced by β-D-Glc and α-L-Rha units, respectively. All compounds showed similar behavior and even though four conformational families were identified, the calculations and NMR experiments support that the 'stacked conformation' known for Le[a] is predominant for all analogues. Interestingly, ROESY showed a correlation between H-1 Fuc/Rha and H-3 GlcNAc which, although small, could be seen in all analogues. For two compounds, the corresponding distance was measured and found to be shorter (~3.7 Å) than that found in the global minimum (4.5 Å). While one published study suggests some motion around the fucosidic bond, this constitutes the first experimental evidence supporting such flexibility. Our MD simulation (Amber10/Glycam06) on Le[a] was in full agreement with previous studies which described a rigid conformation for this branched trisaccharide. Thus, NMR seems to indicate that these dynamic studies are underestimating flexibility around the fucosidic bond.

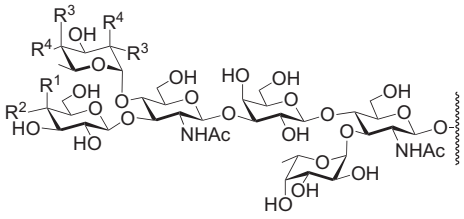
1. Introduction

The Le[a]Le[x] (**1**) hexasaccharide has long been established as a tumor associated carbohydrate antigen (TACA) particularly prevalent in squamous lung carcinoma and immunostaining experiments using an anti-Le[a]Le[x] mAb (43-9F) showed that it was largely associated with tumorigenicity of human small lung carcinoma (SLC) cells.[1] Thus, we are interested in using this hexasaccharide to develop a carbohydrate-based anti-cancer vaccine that would specifically trigger the destruction of these SLC cells by the immune system. MAb 43-9F has been shown to recognize internal epitopes displayed by Le[a]Le[x] while it only weakly binds to the Le[a] trisaccharide.[1a] However, immunization with the Le[a]Le[x] antigen will also likely trigger the production of anti-Le[a] antibodies and these antibodies will, in turn, trigger the destruction of the many normal cells that naturally display Le[a] at their surface.[1a,2] To avoid this potential autoimmune reaction, we are investigating the possibility of replacing the non-reducing end Le[a] trisaccharide with an analogue in which the galactosyl and/or fucosyl residues are replaced by a glucosyl or rhamnosyl unit, respectively, as shown in compounds **2**–**4** (Chart 1). By modifying the non-reducing end trisaccharide, we are attempting to identify an analogue of the TACA that will trigger the production of 43-9F-like antibodies but not that of anti-Le[a] antibodies. In this study we examine how such substitutions may affect the conformational behavior of the non-reducing end trisaccharide using analogues **5**–**8** previously synthesized in our laboratory[3] and running a combination of stochastic searches (MOE2005/AMBER94) and NMR experiments. While analogue **5** is the natural Le[a] trisaccharide, analogue **6** carries a β-D-glucosyl residue rather than a β-D-galactosyl unit at O-3 of the GlcNAc residue, compound **7** is α-L-rhamnosylated rather than α-L-fucosylated at O-4 of the GlcNAc residue and in compound **8** both the β-D-galactosyl and α-L-fucosyl units are replaced by β-D-glucose and α-L-rhamnose, respectively (Chart 1).

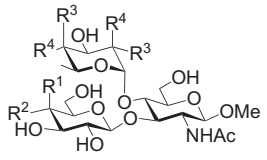
Since Lemieux's early work[4a] on the conformation of the Le[a] antigen, numerous studies using molecular mechanics calculations and/or NMR experiments have been published.[4b–k] These studies conclude that the Le[a] trisaccharide is likely a fairly rigid branched trisaccharide which adopts a rather well-defined 'stacked' conformation. In 1995 however, Imberty et al. published a thorough investigation of the conformational space that can be populated by the Le[a] antigen using the CICADA (channels in conformational space analyzed by driver approach) search approach interfaced with the MM3 program.[5] Interestingly, they came to the conclusion that the Le[a] trisaccharide was possibly less rigid than thought and that it could potentially adopt a conformation different from the stacked global minimum even if this second conformational

* Corresponding author.
E-mail address: fauzanne@uoguelph.ca (F.-I. Auzanneau).

0968-0896/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmc.2012.07.017



1 R^1, R^3 = OH, R^2, R^4 = H (Le^aLe^x)
2 R^1, R^4 = H, R^2, R^3 = OH
3 R^1, R^4 = OH, R^2, R^3 = H
4 R^1, R^3 = H, R^2, R^4 = OH



5 R^1, R^3 = OH, R^2, R^4 = H (Le^a)
6 R^1, R^4 = H, R^2, R^3 = OH
7 R^1, R^4 = OH, R^2, R^3 = H
8 R^1, R^3 = H, R^2, R^4 = OH

Chart 1.

family was not energetically favored. In addition, conformations found within the global minimum family displayed variations of about 20° around the glycosidic torsion angles, supporting some limited flexibility around the glycosidic bonds.[5] In that study however, no experimental evidence was provided to support the calculations. With the development of high field NMR spectrometers and enhanced sensitivity brought about by the use of cryoprobes, we are here able to produce experimental evidence supporting that the Le^a trisaccharide antigen is more flexible than originally thought. Indeed, the results described below support, through calculations and NMR experiments, that analogues **5**–**8** adopt preferentially the known stacked conformation but present some flexibility around the α-L-Fuc/Rha-(1→4)-β-D-GlcNAc glycosidic bond.

2. Results and discussion

Following the approach that we have previously applied successfully to the modeling of Le^x analogues,[6] we performed stochastic searches on trisaccharides **5**–**8** using MOE2005 (molecular operating environment software suite, Ref. 7) and the AMBER94 force field.[8] We focused our attention on the conformational features associated with the glycosidic linkages since these have the most substantial impact on the overall conformations of the analogues.[9] The orientations adopted around these linkages are described by the dihedral angles: Φ = O5-C1-O1-Cx and Ψ = C1-O1-Cx-Cx+1 and the signs of the torsions are in agreement with the recommendations of the IUPAC-IUB Commission of Biochemical Nomenclature.[10] In these stochastic searches, minima in the potential energy surface are randomly sampled through the rotation of all bonds (including ring bonds) to random dihedral angles as well as through a random 0.4 Å Cartesian perturbation of atom positions. For each analogues **5**–**8**, two searches (150,000 iterations) were carried out to generate new conformers that were each minimized with implicit solvation using the GB/SA continuum model.[11] In the first search a RMS tolerance of 0.01 Å on heavy atoms was used to determine whether generated conformations were considered as new. In the second search a RMS tolerance of 0.1 Å was applied on all atoms including all hydrogens atoms to account for possible hydrogen bonding. For each compound, the two databases were combined and those conformations found more than 10 kcal mol^{-1} above the global minima were eliminated. As shown in Table 1, the searches yielded between 1167 and 1270 conformations within 10 kcal mol^{-1} from the global minima but the numbers of conformations found within 3 kcal mol^{-1} of the global minima were considerably smaller from 35 conformations for analogue **6** to 78 for analogue **5**. Since those conformations more than 3 kcal mol^{-1} above the global minima are energetically less relevant, we focused our analysis on those found within 3 kcal mol^{-1}of the global minima identified for each analogue **5**–**8**. We define Φ^1, ψ^1 and Φ^2, ψ^2 as the torsion angles measured, respectively, for the β-D-Gal/Glc-(1→3)-β-D-GlcNAc and α-L-Fuc/Rha-(1→4)-β-D-GlcNAc glycosidic bonds. These conformations clustered around well-defined values of the glycosidic torsions that fell into 4 distinct families (**A**–**D**). The overlaid lowest energy conformations found in each family **A**–**D** for each analogue **5**–**8** are shown on Figure 1a. Table 2 (entries 1–4) gives the glycosidic torsions and standard deviations found for the global minimum **A** and lowest energy representatives of each family (**B**–**D**) averaged over the four analogues (**5**–**8**). As can be seen on Figure 1a and in Table 2, the same global minimum (**A**) was found for all analogues. This conformation corresponds to the global minimum identified as the preferred rigid conformation described in the reported studies of the Le^a trisaccharide.[4,5] Two local minima (Fig. 1, **B** and **C**) were also identified within 3 kcal mol^{-1} of the global minimum for Le^a (**5**) and analogues **6**–**8**. The third local minimum (Fig. 1, **D**) was found 3 kcal mol^{-1} from the global minimum for analogue **7** but at slightly higher energy: 3.8, 4.4 and 4.5 kcal mol^{-1} from the global minima for **5**, **6** and **8**, respectively. As can be seen from these results there is little variation within the analogues from the calculated average values, and we thus conclude that all analogues have a very similar conformational behavior regardless of the substitutions β-D-Gal for β-D-Glc (in analogues **6** and **8**), or α-L-Fuc for α-L-Rha (in analogues **7** and **8**).

Table 1
Number of conformations found within 10 and 3 kcal mol^{-1} from the global minima

Analogue	10 kcal mol^{-1}	3 kcal mol^{-1}
5	1181	78
6	1167	35
7	1270	41
8	1210	50

Interestingly, while the global minimum described here matched that identified by Imberty et al.[5] in the CICADA search (compare Table 2, entries 1 and 5), none of the conformational families **B**–**D** identified in our searches matched the conformational **Fam II** described in their study (Table 2, entry 6). This conformational family differs from the global minimum in the orientation of the fucosyl residue and mostly the value of the ψ^2 angle. Our results also support higher flexibility of the α-L-Fuc/Rha-(1→4)-β-D-GlcNAc linkage than of the β-D-Gal/Glc-(1→3)-β-D-GlcNAc glycosidic bond. Indeed, while the fucose/rhamnose residue adopted different orientations in conformational families **B**–**D** than in the global minimum family **A**, the β-D-Gal/Glc-(1→3)-β-D-GlcNAc retained the orientation found in the global minimum in two out of the three local minima families (**B** and **D**). Even though these results provide evidence for flexibility around the α-L-Fuc/Rha-(1→4)-β-D-GlcNAc linkage in these trisaccharides, Figure 1b shows that the global minimum **A** is largely the predominant conformation in the databases for all analogues (**5**–**8**). Indeed, the Boltzman-averaged torsion angles calculated on each database for the four analogues **5**–**8**, (Table 2, entries 7–10) clearly support that this global minimum is energetically the dominant conformation in all cases. Thus, as previously described for the Le^a antigen[4e–h], this favored conformation should be easily supported for all our

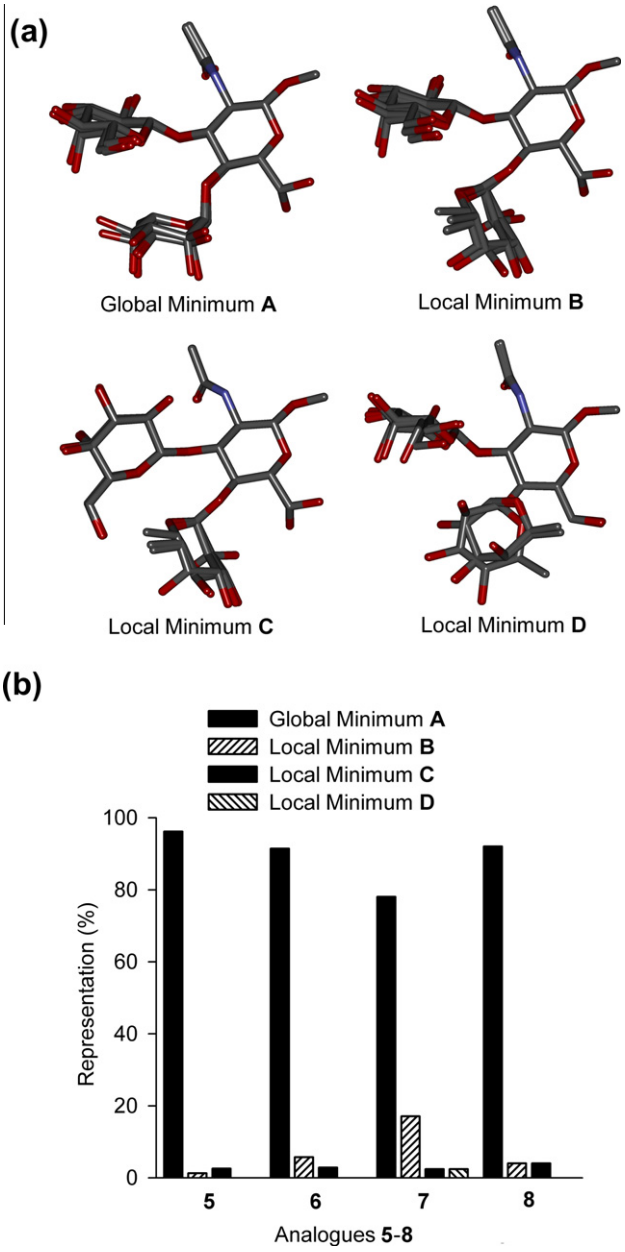


Figure 1. (a) Overlaid of global (**A**) and local mimima (**B**–**D**) identified for analogue **5**–**8** in stochastic searches; (b) representation (%) of each conformation **A**–**D** in the databases (3 kcal mol^{-1}) calculated for each analogue **5**–**8**.

Table 2
Calculated torsion angles[a]

Entry		Φ^1	ψ^1	Φ^2	ψ^2
1	**A**[b]	−73 ± 4	133 ± 3	−71 ± 2	−97 ± 2
2	**B**[b]	−70 ± 3	132 ± 4	−161 ± 2	−157 ± 5
3	**C**[b]	−150 ± 1	97 ± 1	−150 ± 5	−150 ± 1
4	**D**[b]	−67 ± 6	148 ± 5	−84 ± 11	69 ± 10
CICADA search by Imberty et al.[c]					
5	**Fam I**	−75	144	−79	−98
6	**Fam II**	−63	178	−96	−171
Boltzman average for each analogue[d]					
7	**5**	−70	133	−72	−96
8	**6**	−71	135	−72	−96
9	**7**	−73	137	−70	−98
10	**8**	−70	133	−68	−97
Molecular dynamic simulations					
11	**5**	−69 ± 8	133 ± 8	−68 ± 8	−101 ± 7
12	**9**	—[e]	—[e]	−84 ± 16	−131 ± 20

[a] Φ^1, ψ^1 = β-D-Gal/Glc-(1→3)-GlcNAc; Φ^2, ψ^2 = α-L-Fuc/Rha-(1→4)-GlcNAc.
[b] Torsions calculated for the global minima (**A**) and lowest energy representatives of each local minima (**B**–**D**) averaged on all analogues **5**–**8** given below is the standard deviation within the analogues.
[c] See Ref. 5.
[d] Calculated on the 3 kcal mol^{-1} databases.
[e] Not applicable, see Ref. 16.

Table 3
Calculated and measured inter-proton distances

Entry		H-1′/H-3 (Å)	H-1″/H-4 (Å)	H-5″/H-2′ (Å)
Lowest energy conformers[a]				
1	**A**	2.3	2.5	2.6
Boltzman average[b]				
2	**5**	2.4	2.5	2.6
3	**6**	2.4	2.5	2.6
4	**7**	2.4	2.5	2.7
5	**8**	2.4	2.6	2.5
NMR measurements				
7	**5**	2.6[c]	2.4	2.7
8	**6**	2.4	2.4	2.7
9	**7**	2.7[c]	2.4[c]	2.7[c]
10	**8**	2.3	—[d]	2.9

[a] Distances calculated for global minimum **A** averaged on all analogues.
[b] Boltzman-averaged distances on the 3 kcal mol^{-1} databases for each analogue.
[c] Slopes were corrected to account for the overlapping cross-relaxation signals issued from known short inter-proton distances.
[d] Cross relaxation signals were clearly observed (Fig. 2d) signal overlap precluded measurement.

analogues (**6**–**8**) through NMR measurements, particularly based on the inter-proton distances: H-1′ to H-3, H-1″ to H-4, H-2′ to H-5″ (Table 3, entries 1–5), as well as H-1″ to H-6a and H-6b (see Supplementary data for specific values).

The ^{1}H chemical shift assignments for trisaccharides **5**–**8** measured at 300 K in D_2O are reported in the Supplementary data. The vicinal coupling constants measured for the three sugar units have been reported and support an average $^4C_1$ conformation for the galactose, glucose and *N*-acetylglucosamine rings and a $^1C_4$ conformation of the fucose and rhamnose rings.[3] Inter-residue NOE interactions were evaluated using 1D ROESY experiments at 300 K selectively irradiating H-1′, H-1″ and H-5″ for each analogues. We show in Figure 2 the ^{1}H NMR spectrum and the ^{1}H, ^{1}H ROESY spectra (mixing time 200 ms) acquired at 600 MHz on a cryoprobe-equipped spectrometer for analogues **5**–**8**. For all analogues, the expected intra-residue cross-relaxation signals were evident. Cross-relaxation signals were observed for the glycosidic linkages: H-1′ to H-3 and H-1″ to H-4; and inter-residue cross-relaxation signals were also found from H-1″ to H-6a/H-6b and from H-5′ to H-2″ in all spectra.[4e–h]

Cross-relaxation buildup curves were obtained over 9 mixing times ranging from 20 ms to 400 ms with 5 mixing times chosen between 20 and 100 ms to ensure accurate calculation of the slope at initial buildup. The normalized buildup curves (Fig. 3) were fitted to double exponential equations and the slopes at initial buildup were calculated and used to evaluate the inter-proton distances using the isolated spin pair approximation.[12] In all cases the R^2 values for the exponential fits were at least 0.99. As can be seen on Figure 2a, H-1 and H-1′ gave overlapped signal in analogue **5** and H-1, H-1′ and H-5″ gave overlapped signals for trisaccharide **7** (Fig. 2c). These signals were irradiated together (Fig. 3a and c) and the slopes at initial buildup were corrected for the known contributing distance before calculating the desired unknown inter-proton distances. The measured inter-proton distances for the glycosidic linkages (H-1′ to H-3, H-1″ to H-4) and between H-2′ (Gal/

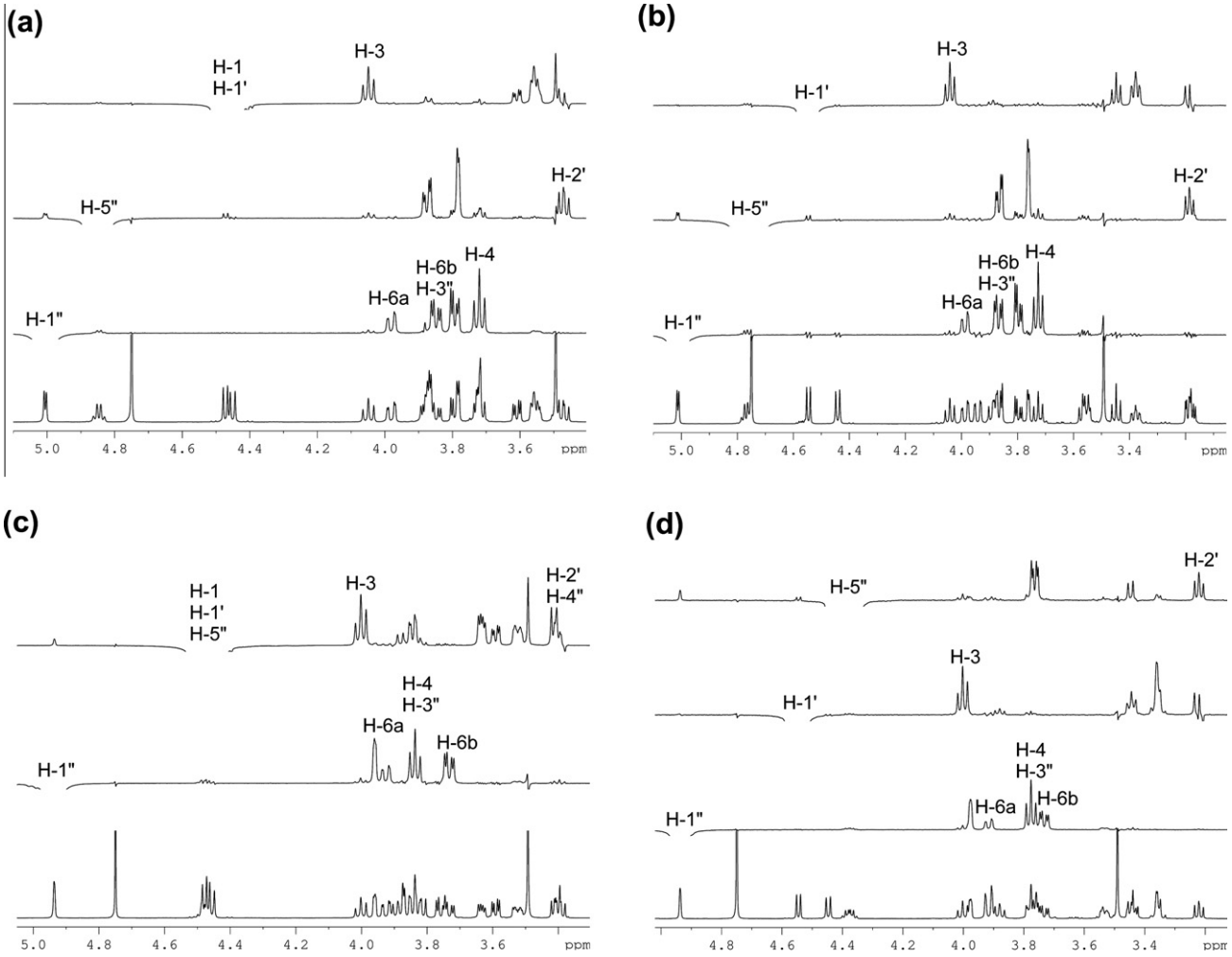


Figure 2. ^{1}H NMR spectra (600 MHz, 300 K) and ^{1}H, ^{1}H ROESY spectra (200 ms); (a) trisaccharide **5**: selective excitations of proton H-1″ and H-5″; combined excitations of H-1 and H-1′; (b) trisaccharide **6**: selective excitations of proton H-1″, H-5″ and H-1′; (c) trisaccharide **7**: selective excitation of proton H-1″ and combined excitations of H-1, H-1′, H-5″; (d) trisaccharide **8**: selective excitations of proton H-1″, H-1′ and H-5″.

Glc) to H-5″ (Fuc/Rha), which are most representative of the global minimum, are listed in Table 3 (entries 7–10). The distances measured from H-1″ to H-6a and H-6b are given in the Supplementary data. As can be seen from Table 3 these experimental distances are in very good agreement with the results obtained for all analogues in the stochastic searches. It is therefore clear that analogues **5**–**8** all adopt one major conformation in solution that corresponds to the stacked conformation known for the Le^{a} antigen.[4,5]

Upon close examination of the ^{1}H, ^{1}H ROESY spectra we noticed the presence of an unexpected cross-relaxation correlation between H-1″ (Fuc or Rha) and H-3 (GlcNAc). As can be seen on Figure 4a, this signal, albeit small, was clearly visible for all analogues. While this signal was well isolated for the Le^{a} trisaccharide **5** and analogue **6**, it could not be integrated accurately for analogues **7** and **8**. The normalized buildup curves for compounds **5** and **6** (Fig. 4b) were fitted to double exponential equations with acceptable R^2 values of 0.98 and 0.97 for **5** and **6**, respectively. The slopes at initial buildup were calculated and used to evaluate the distance between H-1″ and H-3 in these two analogues (Table 4, entries 1 and 2). Indeed, these distances were found to be shorter than those measured in the Boltzman averages for **5** and **6** (Table 4, entries 4 and 5) which are clearly dominated by the global minimum **A** (entry 6). Interestingly, local minima **B**–**D** which all involved rotation around the α-L-Fuc/Rha-(1→4)-β-D-GlcNAc glycosidic bond, led to H-1″ to H-3 distances shorter than that found in the global minimum **A** (entries 7–9). Similarly the conformational family II (**Fam II**) described in the CICADA study by Imberty et al.[5] also led to a shorter distance between H-1″ and H-3 (entry 10). Since NMR provides time-averaged measurements, it is likely that none of these higher energy conformations are heavily populated in solution at any given time but that rapid motion at 300 K leads to transient proximity of these two hydrogens. Thus, the distance measured here between H-1″ and H-3 provides experimental evidence that supports flexibility around the α-L-Fuc/Rha-(1→4)-β-D-GlcNAc glycosidic bond.

To further investigate the flexibility of the Le^{a} antigen, we ran a 20 ns molecular dynamic simulation in explicit water at 300 K using AMBER10[13] with the inclusion of the Glycam06 parameters for carbohydrates.[14] The average values of the Φ and Ψ dihedral angles for both glycosidic linkages as well as the associated calculated standard deviations are given in Table 2, entry 11. The average value and standard deviation for the distance between H-1″ and H-3 is given Table 4, entry 11. Since we are mostly interested in the fucosidic bond, the superimposition of the Φ-Ψ trajectory for the α-L-Fuc-(1→4)-β-D-GlcNAc linkage over the MM3 grid search energy maps obtained for the disaccharide linkages[5,15] is shown in Figure 5a. The individual trajectories of the Φ and Ψ angles for this glycosidic bond as well as the trajectory of the distance

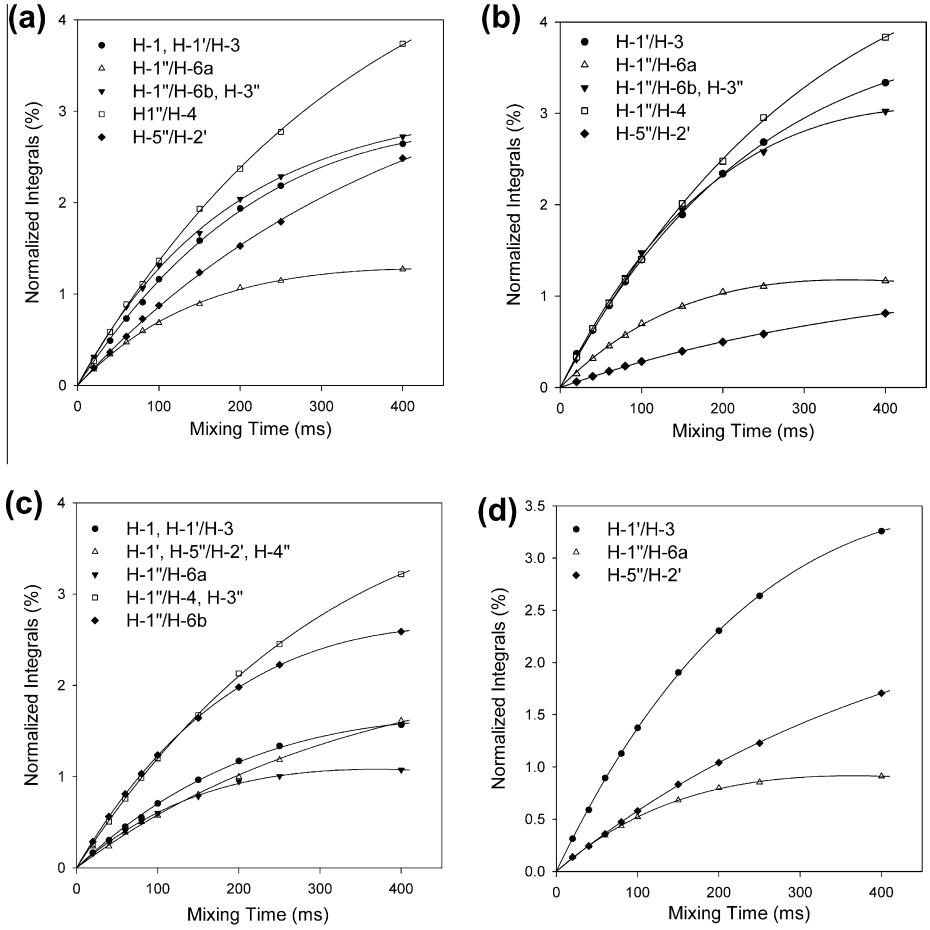


Figure 3. ^{1}H, ^{1}H ROESY buildup curves (a) trisaccharide **5**: combined excitations of H-1 and H-1′, selective excitations of H-1″ and H-5″; (b) trisaccharide **6**: selective excitations of H-1′, H-1″ and H-5″; (c) trisaccharide **7**: combined excitations of H-1, H-1′, H-5″, selective excitation of H-1″; (d) trisaccharide **8**: selective excitations of H-1′, H-1″ and H-5″.

between fucose H-1 and *N*-acetylglucosamine H-3 are also shown in Figure 5a. (similar map and Φ-Ψ trajectories for the β-D-Gal-(1→3)-β-D-GlcNAc linkage are given in the Supplementary data) In accordance with previous simulations on the same antigen[4d,g], this simulation supported very limited flexibility around either glycosidic bonds. In particular, the α-Fuc-(1→4)-β-D-GlcNAc linkage remained in the low-energy region surrounding the global minimum **A** with only limited deviation from the average value. As a result, the distance calculated between fucose H-1 and *N*-acetylglucosamine H-3 was identical to that measured for global minimum **A** or the Boltzman average obtained in the stochastic search (Table 4 entries 6 and 4) with limited deviation from this value (see Table 4, entry 11 and Fig. 5a(iii)). Thus, these results, while in full agreement with previous studies[4d,g], contrast with the results given by the NMR data which supports some flexibility around the α-Fuc-(1→4)-β-D-GlcNAc bond leading to a shorter measured distance between fucose H-1 and *N*-acetylglucosamine H-3 (Table 4, entry 1).

We compared the outcome of this simulation to that which we have described[16] for the linear trisaccharide **9**: α-L-Fuc-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal (Chart 2). The average values of the Φ and Ψ dihedral angles for the fucosidic linkage and standard deviations are given in Table 2, entry 12; the average value and standard deviation for the distance between H-1″ and H-3 is given Table 4, entry 12; and Figure 5b shows the trajectory for this glycosidic bond as well as the trajectory of the distance between H-1 Fuc and H-3 GlcNAc during this 8 ns simulation at 300 K. Not surprisingly, this linear trisaccharide displayed increased flexibility around the fucosidic bond. While it remained in the general proximity of the global minimum **A**/**Fam I**, the trajectory also visited the orientation of the fucose found in local minima **B** and **C** as well as **Fam II** identified for the Le^{a} trisaccharide. Interestingly, this increased flexibility led to average values for the Φ and Ψ angles (Table 2, entry 12) that were found to be in between the values measured by Imberty et al.[5] in the CICADA search for **Fam I** and **Fam II** (entries 5 and 6). In addition and even more remarkably, it also led to a shorter average distance between H-1 Fuc and H-3 GlcNAc and greater variation of this distance througout the simulation (see (iii) Fig. 5b and Table 4, entry 12). Looking back at the NMR data acquired[16] for trisaccharide **9** we indeed observed a correlation between H-1 Fuc and H-3 GlcNAc for which we were able to plot a buildup curve (see Supplementary data) that was fitted to a double exponential equation (R^2 = 0.99). The slope at initial buildup was used to evaluate the distance between these two hydrogens and the result is given Table 4, entry 3. Not surprisingly, this distance, shorter than that measured for the branched analogues **5** and **6**, supports an increased flexibility around the fucosidic linkage in the linear analogue than in the branched analogues. However, this measured distance is still much shorter than that given by the calculations (compare Table 4 entries 3 and 12) on analogue **9**. Thus in general, our experimental results suggest that, indeed as proposed by Imberty et al.,[5] the α-L-Fuc-(1→4)-β-D-GlcNAc glycosidic linkage

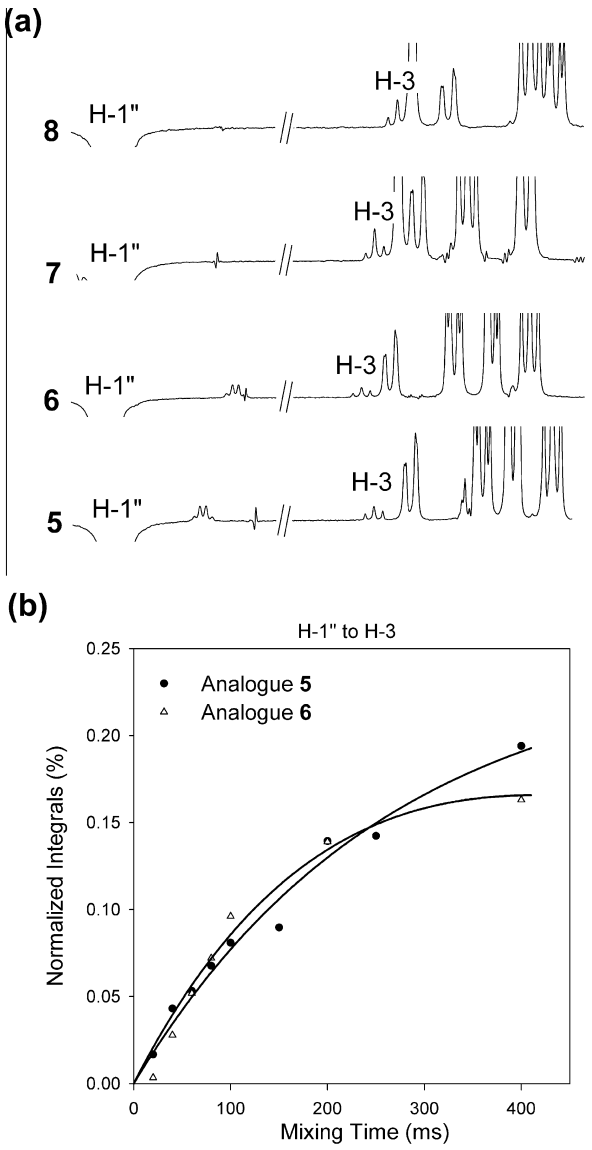


Figure 4. Evidence of NOE transfer between H-1″ and H-3 in analogues **5**–**8**. (a) ^{1}H, ^{1}H ROESY at 400 ms mixing time upon selective excitation of H-1″ in trisaccharides **5**–**8**; (b) build-up curves for analogues **5** and **6**.

Table 4
Calculated and measured distances between fucose H-1 and *N*-acetylglucosamine H-3

Entry		(Å)
NMR measurements		
1	**5**	3.8
2	**6**	3.7
3	**9**	3.4
Calculated distances		
4	Boltzman average for **5**[a]	4.5
5	Boltzman average for **6**[a]	4.5
6	Global minimum **A**[b]	4.5
7	Local minimum **B**[b]	3.3
8	Local minimum **C**[b]	3.6
9	Local minimum **D**[b]	2.0
10	CICADA **Fam II**[c]	3.6
11	Molecular dynamics **5**[d]	4.5 ± 0.1
12	Molecular dynamics **9**[d]	4.1 ± 0.3

[a] Boltzman-averaged distances on the 3 kcal/mol databases average on analogues **5**–**8**.
[b] Distances calculated for conformational families **A**–**D** averaged on all analogues.
[c] Family 2 identified by Imberty et al. in the CICADA search.[5]
[d] Average distance and standard deviation.

presents a higher degree of flexibility than molecular dynamic studies seemed to indicate.[4]

3. Conclusion

In conclusion, the stochastic searches and NMR experiments carried out on analogues **5**–**8** have clearly showed that all trisaccharides had identical conformational behavior regardless of the substitutions β-D-Gal for β-D-Glc (in analogues **6** and **8**), or α-L-Fuc for α-L-Rha (in analogues **7** and **8**). Even though these searches identified 3 local minima, both calculations and NMR data indicated that the stacked conformation, already well-established[4,5] for the Le^{a} antigen, was largely the predominant conformation for all analogues. Therefore, it is reasonable to assume that these substitutions in vaccines candidates **2**–**4** will not interfere with the overall conformation of the hexasaccharides. Thus, while these substitutions may abolish cross-reactivity with the Le^{a} trisaccharide we hope that analogues **2**–**4** will retain cross-reactivity with the natural Le^{a}Le^{x} TACA based on the internal epitope recognized by 43-9F-like antibodies. Competitive ELISA experiments using a Le^{a}-BSA immobilized antigen and analogues **5**–**8** as the soluble antigens competing for binding to the commercially available anti-Le^{a} mAb SPM522 (Abnova) will be performed in our laboratory in the near future. These studies will hopefully lead to the discovery that one of analogues **4**–**8** does not cross-react with the Le^{a} trisaccharide similarly to the analogue of Le^{x} in which the galactose residue is replaced by a glucose unit does not bind to the anti-Le^{x}mAb SH1.[6] If such analogue is discovered the corresponding hexasaccharide analogue of Le^{a}Le^{x} will be prepared and assessed for cross-reactivity with the natural antigen.

Despite a clear conformational preference for the global minimum, our NMR data offers evidence for some flexibility around the α-Fuc/Rha-(1→4)-β-D-GlcNAc. Flexibility around this glycosidic bond in Le^{a} was suggested by the CICADA search published by Imberty et al.[5] but experimental evidence had, thus far, been lacking to support these calculations. Indeed, molecular dynamic simulations performed on the Le^{a} trisaccharide here by us and reported by others[4d,g] come in sharp contrast with this suggestion as they support a very rigid conformation for this branched trisaccharide. However, the additional signal enhancement provided by a 600 MHz cryoprobe-equipped spectrometer permitted the observation in the ^{1}H, ^{1}H ROESY spectra of an unexpected cross-relaxation correlation between H-1″ (Fuc or Rha) and H-3 (GlcNAc) in all analogues. Indeed, for two of these analogues these signals could be translated into an inter-proton distance of ~3.8 Å. This correlation cannot be explained by the sole stacked minimum energy conformation **A** in which these two hydrogens are distant by 4.5 Å. Interestingly, looking back at NMR data acquired[16] for linear trisaccharide **9**, the distance between these hydrogens was measured to be shorter (3.4 Å) than in analogues **5** and **6**. Indeed, molecular dynamic simulations showed a greater flexibility around the fucosidic linkage in linear analogue **9** than in branched analogue **5** though this increased flexibility led to a calculated distance between H-1 Fuc and H-3 GlcNAc (4.1 Å) still greater that that measured in the NMR data. Thus, we conclude that molecular dynamics simulations on analogues **5** and **9** did not accurately account for the flexibility around the α-Fuc-(1→4)-β-D-GlcNAc glycosidic bond. Although, the NMR data presented here provides the first evidence of some limited flexibility around the α-Fuc-(1→4)-β-D-GlcNAc bond in Le^{a} analogues, it also clearly supports that the previously identified stacked conformation[4,5] is largely preponderant in solution. This limited flexibility, while not challenging that the biologically relevant conformation is indeed the known stacked conformation, may become of importance when considering the binding of the Le^{a} trisaccharide to lectins. Indeed, several examples have been described in the literature in which the bound

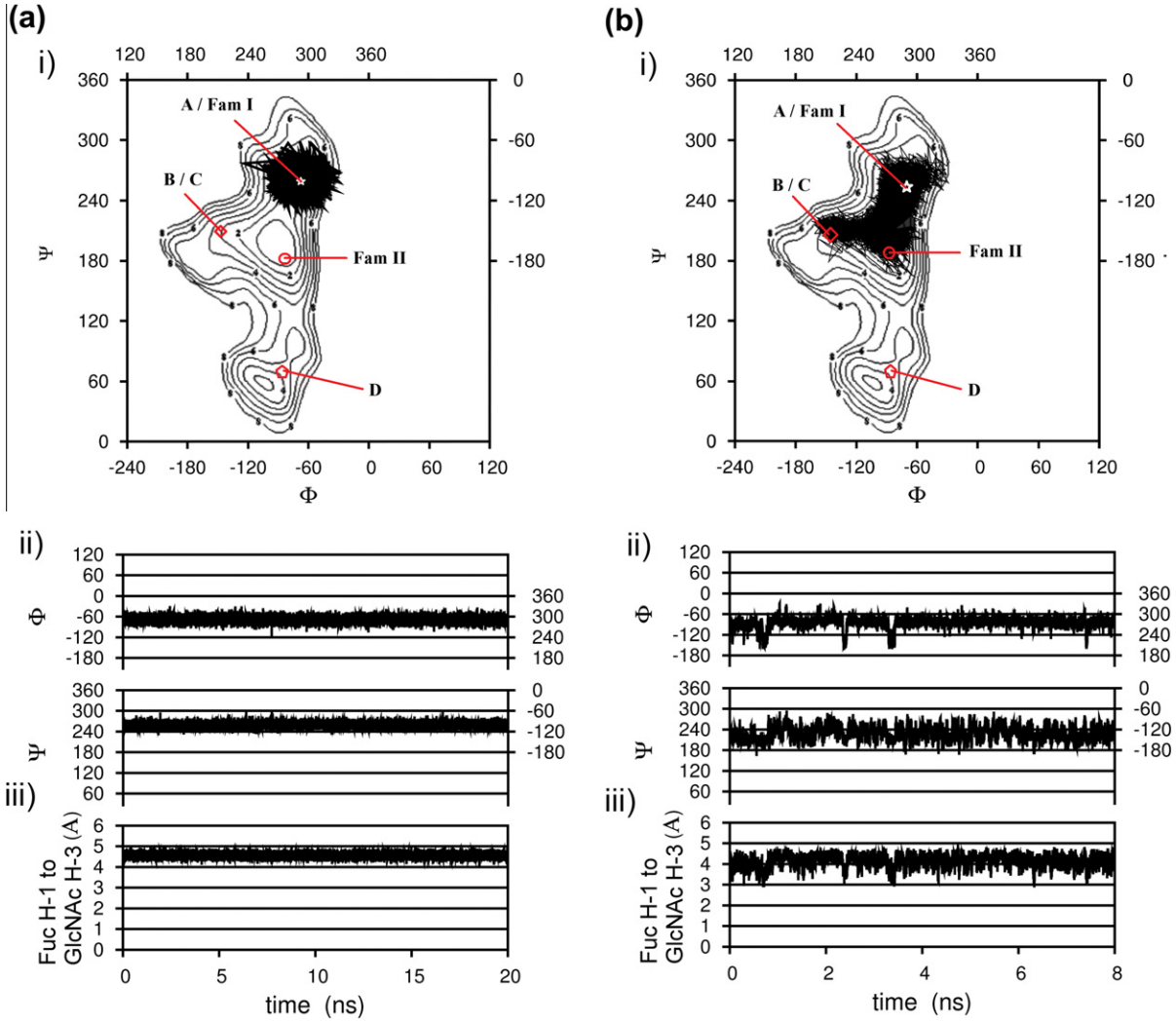
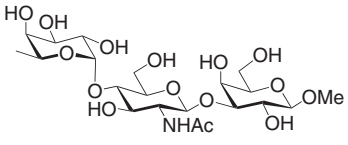


Figure 5. Flexibity of the α-L-Fuc-(1→4)-β-D-GlcNAc linkage in (a) trisaccharide **5** and (b) trisaccharide **9** as assessed by molecular dynamics (see text and Ref. 16): (i) Φ/ψ trajectories superimposed on the energy maps obtained for the disaccharide (black contours)[5,15], conformations **A**–**D**, **Fam I** and **Fam II** (see Table 2 entries 1–6) are indicated; (ii) Φ and ψ trajectories; (iii) trajectory of the distance between Fuc H-1 and GlcNAc H-3.



9

Chart 2.

conformation of Lewis antigens to lectins deviated slightly from the known global minimum for these ligands in solution.[17]

4. Experimental section

4.1. Stochastic searches

The trisaccharide was subjected to all atom and heavy atom stochastic conformational searches using the AMBER94 force field[8] in the Molecular Operating Environment (MOE2005, Ref. 7) program suite. In this implementation of AMBER94, any parameters missing for a class of compound are approximated by MOE based on parameter values for similar molecular fragments. The starting structure was built using the carbohydrate builder in MOE2005 and submitted to the stochastic search. This search is similar to the random incremental pulse search (RIPS) method[18] in which the coordinates of each atom are randomly perturbed, after which the entire molecule is minimized to attempt to generate a new conformer. In these stochastic searches, new conformations are generated by the rotation of all bonds (including ring bonds) to random dihedral angles. In our searches a bias of 30° was used. This means that dihedral angles were selected with a sum-of-Gaussians distribution with peaks at multiples of 30°. In addition to bond rotations, atom positions were randomly perturbed by 0.4 Å. These random perturbations were repeated 150,000 times per conformational search and each generated conformer was subjected to 500 steps of minimization with full degrees of freedom with implicit solvation by the Generalized Born/Surface Area (GB/SA) continuum solvation model[11] using a dielectric constant of 78. Two separate searches were performed and the databases were combined for analysis. Initially each new conformation was checked against the previously found conformations using a root-mean-square (RMS) tolerance of 0.1 Å on all atoms, including the hydrogen atoms to account for possible hydrogen bonding. The second search was carried out using a more stringent RMS value (0.01 Å) applied to the heavy atoms only. In both searches the cal-

culations terminated at the end of the 150,000 iterations as the number of failures criteria to find new conformations (10,000 iterations in a row) was never met. Structures that did not have the $^4C_1$ conformation for D-GlcNAc, D-Gal or D-Glc or those that did not maintain a $^1C_4$ conformation for L-Fuc or L-Rha were eliminated. The results of the two searches were combined for analysis, and conformations with energies over 10 kcal mol^{-1} above the global minima were rejected.

4.2. NMR Experiments

The trisaccharides **5**–**8** (4–7 mg) were lyophilized three times from D_2O (95%) and dissolved in 0.7 mL of D_2O (99.96%). The solutions were immediately transferred to dried 5 mm NMR tubes which were carefully sealed with parafilm. One-dimensional selective gradient ROESY experiments were acquired at 300 K on a 600 MHz spectrometer equipped with a cryoprobe following selective excitations of protons: H-1′, H-1″ and H-5″ and applying a 5 s relaxation delay between scans. The optimum irradiation range for each signal was used to determine the best length of the soft pulse applied. This value was automatically calculated by Bruker software (BUTSEL-NMR). 1D ROESY spectra were recorded at nine mixing times from 20 to 400 ms. Before Fourier transformation, the FIDs were zero-filled once and multiplied with a 1 Hz line broadening factor. Spectra were phase and baseline corrected and integrated. The integrals measured for the irradiated signals were plotted against mixing time and the obtained curves were fitted to a double exponential decaying function:

$$f(\tau_m) = -A[\exp(B\tau_m) + \exp(C\tau_m)]$$

where τ_m is the mixing time and A, B and C are adjustable parameters. The values of these integrals were extrapolated to 0 ms mixing time, and the integrals from cross-relaxation peaks were normalized through division by these extrapolated values. The normalized cross-relaxation integrals were plotted against the mixing times and the buildup curves were fitted to a double exponential equation of the form

$$f(\tau_m) = A[\exp(B\tau_m) - \exp(C\tau_m)]$$

the initial slopes at 0 ms mixing times were determined from the calculated first derivatives[12]

$$f'(0) = A(B - C)$$

Interproton distances were calculated based on the isolated spin pair approximation (ISPA):

$$r_{ij} = r_{ref}(S_{ref}/S_{ij})^{1/6}$$

where S is the initial slope at $\tau_m = 0$, and r is the proton-proton distance.

The intra-residue cross-peaks used as reference for distance determinations and known distances for the correction of slope when signals were overlapping were:

For analogue **5**: references H-1′/H-3′ (2.64 Å), H-1″/H-2″ (2.42 Å) and H-5″/H-3″ (2.50 Å); known distances corrected for H-1/H-3 (2.51 Å) and H-1″/H-3″ (3.76 Å).

For analogue **6**: references H-1′/H-5′ (2.39 Å), H-1″/H-2″ (2.42 Å) and H-5″/H-3″ (2.50 Å); known distances corrected for H-1″/H-3″ (3.76 Å).

For analogue **7**: references H-1′/H-5′ (2.42 Å) and H-1″/H-2″ (2.51 Å); known distances corrected for H-1/H-3 (2.51 Å), H-1″/H-3″ (3.77 Å), H-1′/H-2′ (3.05 Å), and H-5″/H-2″ (3.04 Å).

For analogue **8**: references H-1′/H-3′ (2.51 Å), H-1″/H-2″ (2.51 Å) and H-5″/H-3″ (2.59 Å); no known distances corrected for.

For analogue **9**[16]: reference H-1Fuc/H-2Fuc (2.42 Å) see Supplementary data.

4.3. Molecular dynamics simulations for analogue 5

The molecular dynamics simulations were carried out using the SANDER module of AMBER10[13] with the inclusion of Glycam06 parameters[14] for carbohydrates. The initial structure was built using the XLEAP module of AMBER. The molecule was solvated with 10 shells of TIPS3P water molecules[19] giving a total of 1070 solvent molecules in a 30 Å cubic box. Dynamics trajectories of 20 ns were calculated. Each simulation was carried out in five stages using the SANDER module of AMBER. First the water molecules were minimized at constant volume (NVT) whereas the trisaccharide was held fixed. A 1000 cycle limit was employed with 50 cycles of steepest descent followed by conjugate gradient minimization. A 1×10^{-3} kcal mol^{-1} Å gradient convergence criteria was used. A 10 Å cutoff was set, and non-bonded updates were made every 10 steps. The dielectric constant was set to 1. Then, the entire system was minimized without any restraints. A 2500 cycle limit was used with 500 cycles of steepest descent followed by conjugate gradient minimization. Afterward, the system was heated for 20 ps from 0 to 300 K with the saccharide weakly restrained. Time steps of 2 fs were used and translational momentum was removed every 1000 steps. Bonds involving hydrogens were constrained using the SHAKE algorithm with 0.00001 Å tolerance. Coordinates were output every 250 steps. The system was then equilibrated at a constant pressure (NPT) of 1 atm for 100 ps with no restraints. Finally, production runs of 20 ns were carried out at 300 K. The PTRAJ module of AMBER was used to analyze the results.

Acknowledgments

The authors thank the National Sciences and Engineering Research Council of Canada for financial support. This work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET: www.sharcnet.ca) and Compute/Calcul Canada.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmc.2012.07.017.

References and notes

1. (a) Pettijohn, D. E.; Stranahan, P. L.; Due, C.; Ronne, E.; Sorenson, H. R.; Olsson, L. *Cancer Res.* **1987**, *47*, 1161; (b) Stroud, M. R.; Levery, S. B.; Martensson, S.; Salyan, M. E. K.; Clausen, H.; Hakomori, S.-I. *Biochemistry* **1994**, *33*, 10672; (c) Battifora, H.; Sorenson, H. R.; Mehta, P.; Ahn, C.; Niland, J.; Hage, E.; Pettijohn, D. E.; Olsson, L. *Cancer* **1992**, *70*, 1867; (d) Stranahan, P. L.; LaRoe, J.; McCombs, R.; Goldsmith, A.; Rahim, I.; Overland, M.; Pettijohn, D. E. *Glycoconjugate J.* **1996**, *13*, 741; (e) Stranahan, P. L.; LaRoe, J.; McCombs, R.; Rahim, I.; Kuhn, C. W.; Pettijohn, D. E. *Oncol. Rep.* **1998**, *5*, 235; (f) Pettijohn, D. E.; Pfenninger, O.; Brown, J.; Duke, R.; Olsson, L. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 802; (g) Stranahan, P. L.; Howard, R. B.; Pfenninger, O.; Cowen, M. E.; Johnston, M. R.; Pettijohn, D. E. *Cancer Res.* **1992**, *52*, 2923.
2. Lemieux, R. U.; Baker, D. R.; Weinstein, W. M.; Switzer, C. M. *Biochemistry* **1981**, *20*, 199.
3. (a) Liao, L.; Auzanneau, F.-I. *Org. Lett.* **2003**, *5*, 2607; (b) Liao, L.; Auzanneau, F.-I. *J. Org. Chem.* **2005**, 6265; (c) Liao, L.; Auzanneau, F.-I. *Carbohydr. Res.* **2006**, 2426.
4. (a) Lemieux, R. U.; Bock, K.; Delbaere, L. T. J.; Koto, S.; Rao, V. S. *Can. J. Chem.* **1980**, *58*, 631; (b) Biswas, M.; Rao, V. S. R. *Carbohydr. Polym.* **1982**, *2*, 205; (c) Hounsell, E. F.; Jones, N. J.; Gooi, H. C.; Feizi, T.; Donald, A. S. R.; Feeney, J. *Carbohydr. Res.* **1988**, *178*, 67; (d) Mukhopadhyay, C.; Bush, C. A. *Biopolymers* **1991**, *31*, 1737; (e) Cagas, P.; Bush, C. A. *Biopolymers* **1990**, *30*, 1123; (f) Bechtel, B.; Wand, A. J.; Wroblewski, K.; Koprowski, H.; Thurin, J. *J. Biol. Chem.* **1990**, *265*, 2028; (g) Kogelberg, H.; Rutherford, T. J. *Glycobiology* **1994**, *4*, 49; (h) Martin-Pastor, M.; Bush, C. A. *Biochemistry* **2000**, *39*, 4674; (i) Martin-Pastor, M.; Bush, C. A. *Carbohydr. Res.* **2000**, 147.
5. Imberty, A.; Mikros, E.; Koča, J.; Mollicone, R.; Oriol, R.; Pérez, S. *Glycoconjugate J.* **1995**, *12*, 331.
6. Wang, J.-W.; Asnani, A.; Auzanneau, F.-I. *Bioorg. Med. Chem.* **2010**, *18*, 7174.

7. Chemical Computing Group Inc., Montreal, QC, Canada.
8. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
9. Imberty, A.; Pérez, S. *Chem. Rev.* **2000**, *100*, 4567.
10. IUPAC-I.U.B. Commission of Biochemical Nomenclature. *Eur. J. Biochem.* **1997**, *243*, 9.
11. (a) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127; (b) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005; (c) Schaefer, M.; Karplus, M. A. *J. Phys. Chem.* **1996**, *100*, 1578.
12. Maaheimo, H.; Kosma, P.; Brade, L.; Brade, H.; Peters, T. *Biochemistry* **2000**, *39*, 12778.
13. Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*; University of California: San Francisco, 2008.
14. Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. *J. Comput. Chem.* **2008**, *29*, 622.
15. Imberty, A.; Mazeau, K.; Pérez, S.; Rivet, A.; Sabin, C. Disaccharide Database. http://www.cermav.cnrs.fr/cgi-bin/di/di.cgi (accessed August 19, 2008).
16. Jackson, T. A.; Robertson, V.; Imberty, A.; Auzanneau, F.-I. *Bioorg. Med. Chem.* **2009**, *17*, 1514.
17. (a) Yuriev, E.; Farrugia, W.; Scott, A. M.; Ramsland, P. A. *Immunol. Cell Biol.* **2005**, *83*, 709; (b) Haselhorst, T.; Weimar, T.; Peters, T. *J. Am. Chem. Soc.* **2001**, *123*, 10705.
18. Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379.
19. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.