



HAL
open science

The hydration features of carbohydrate determinants of Lewis antigens

Michael Reynolds, Andreas Fuchs, Thisbe Kirsten Lindhorst, Serge Perez

► **To cite this version:**

Michael Reynolds, Andreas Fuchs, Thisbe Kirsten Lindhorst, Serge Perez. The hydration features of carbohydrate determinants of Lewis antigens. *Molecular Simulation*, 2008, 34 (04), pp.447-460. 10.1080/08927020701713878 . hal-00515020

HAL Id: hal-00515020

<https://hal.science/hal-00515020>

Submitted on 4 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The hydration features of carbohydrate determinants of Lewis antigens

Journal:	<i>Molecular Simulation</i> / <i>Journal of Experimental Nanoscience</i>
Manuscript ID:	GMOS-2007-0084.R1
Journal:	Molecular Simulation
Date Submitted by the Author:	27-Sep-2007
Complete List of Authors:	Reynolds, Michael; CERMAV-CNRS, Glycobiologie Moleculaire Fuchs, Andreas; Christiana Albertina University of Kiel, Otto Diels Institute of Organic Chemistry Lindhorst, Thisbe; Christiana Albertina University of Kiel, Otto Diels Institute of Organic Chemistry Perez, Serge; CERMAV-CNRS, Glycobiologie Moleculaire
Keywords:	Lewis X trisaccharide, Lewis Y tetrasaccharide, Molecular dynamics

SCHOLARONE™
Manuscripts

The hydration features of carbohydrate determinants of Lewis antigens

M. REYNOLDS[†], A. FUCHS[‡], T. K. LINDHORST[‡] and S. PEREZ^{†*}

[†] CERMAV-CNRS, BP 53-38041 Grenoble, France;

[‡] Otto Diels Institute of Organic Chemistry, Christiana Albertina University of Kiel,
Otto-Hahn-Platz 4, 24098 Kiel, Germany.

The hydration behaviour of two Lewis carbohydrate determinants, Lewis X (Le^x) (**1a**) and Lewis Y (Le^y) (**1b**), have been investigated by a comparative 20 ns molecular dynamics (MD) study. The detailed hydration of the two oligosaccharides was described using the AMBER-based GLYCAM04 force field which is developed specifically for modelling carbohydrates.

In comparison with previous *in vacuo* MD simulations and high resolution NMR studies, the solvated structures have shown to be dynamically stable and equally rigid with regards to glycosidic linkage torsion angles. Comparisons have also been made between the solvated structures and crystal structures currently available, namely the Le^x dimer crystal, the crystal structure of Le^x complex with the 291-2G3-A-Fab antibody; and the crystal structures of Le^y complexed with the hu3S193 and CBR96 antibodies. When the solvated structures and respective crystal structures were overlaid for comparison the only significant difference was the Galp(3)-O5-C5-C6-O6 and the GlcpNAc(2)-O5-C5-O6-C6 orientations.

The hydration sphere was investigated using a radial pair distribution function. Several solute heteroatoms were identified for the occurrence of bridging events where a solute heteroatom is linked to (at least) one other via a solvent molecule.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Keywords: Lewis X trisaccharide, Lewis Y tetrasaccharide, molecular dynamics

1. Introduction

The sequential addition of L-fucose monosaccharides onto oligosaccharide precursor chains of glycolipids or glycoproteins results in the formation of the so-called Lewis (Le) carbohydrate determinants that can be classified as type 1 (Le^a and Le^b) and type 2 (Le^x and Le^y) antigens. The incorporation of Le^a and Le^b determinants into erythrocyte membranes creates the blood-group antigens. Other structurally related determinants, i.e. Le^x and Le^y (Fig. 1) are expressed in only a few cell types such as certain epithelial cells. Lewis antigens have been found to be expressed at high density in many carcinomas.[1] Le^x and Le^y antigens are considered to be tumour-associated antigens and they represent targets for immunotherapeutic strategies for carcinoma.[2] The three-dimensional structures of these carbohydrate determinants of free Lewis system antigens and their complexes with antibodies have been extensively studied.[2] The rigid nature of Lewis determinants is fairly well established. Consequently, their free conformation is likely to represent the bio-active conformation, i.e. the one recognised by a protein.[2] This is true for Lewis determinants bound to antibodies specific for type 2 determinants. The rigid conformation of Le^y , when presented to the immune system on tumour cells, has been shown to elicit a structurally conserved antibody response.[3] The availability of well resolved 3D structures of proteins interacting with Lewis antigens enlightens the role of ordered water molecules in antibody recognition. Although water molecules have a fundamental role in carbohydrate-protein interactions and carbohydrate recognition, the biological importance of this aspect of intermolecular recognition in aqueous phase as well as the complexity of its molecular dynamics is not fully understood.

1
2 The aim of our investigation was to study the solvated structures of the type 2 Lewis
3
4 determinants using computer modelling and comparing the results with the structures
5
6 found in crystal structures and complexes with antibodies. Molecular dynamics (MD)
7
8 studies were carried out in order to monitor the hydration sphere of the tri and
9
10 tetrasaccharide, and in particular, to study water molecule residence times and bridging
11
12 events. The bridging events, which temporarily link one saccharide heteroatom to (at
13
14 least) one other saccharide heteroatom via a connecting water molecule, are thought to
15
16 have a great importance in the 3D structure of the saccharides which in turn leads to its
17
18 biological activity in living organisms. To investigate the solvated structures further, we
19
20 have made comparisons between the MD results and crystal structures in order to find
21
22 “common” bridging water molecules i.e. bridging water molecules which are not
23
24 displaced upon interaction with a receptor. Other heteroatoms, for example those found
25
26 in amino acid residues, which may displace bridging water molecules from the solute
27
28 have also been identified. The loss of ordered water from the solvated systems would
29
30 contribute to an increase in entropy of the system as it is being expelled from an ordered
31
32 position to the disordered bulk. By contrast, water molecules which are found to remain
33
34 “complexed” to the solute upon interaction would not be contributing to the change in
35
36 entropy and so therefore must be contributing to the affinity between the solute and its
37
38 receptor.
39
40
41
42
43
44
45
46
47
48

49 **2. Materials and methods**

50 51 52 **2.1 Nomenclature**

53
54 Lewis X trisaccharide (Le^x) [β -D-Galp-(1-4)[α -L-Fucp-(1-3)]- β -D-GlcpNAc] (Fig. 1a)

55
56 Lewis Y tetrasaccharide (Le^y) [α -L-Fucp-(1-2)- β -D-Galp-(1-4)[α -L-Fucp-(1-3)]- β -D-
57
58 GlcpNAc] (Fig. 1b)
59
60

[Figure 1 here]

All glycosidic linkages are described by the standard nomenclature of Φ and Ψ where $\Phi = O5-C1-Ox-Cx$, outlined in red in figure 1, and $\Psi = C1-Ox-Cx-C(x+1)$, outlined in blue in figure 1.

Conformational flexibility around the glycosidic linkage of Le^x between $Fucp(1) - GlcpNAc(2)$ is described by the two torsion angles: $\Phi = O5-C1-O3'-C3'$ and $\Psi = C1-O3'-C3'-C4'$. In the case of the $Galp(3) - GlcpNAc(2)$ linkage, the torsion angles are defined as: $\Phi = O5''-C1''-O4'-C4'$ and $\Psi = C1''-O4'-C4'-C5'$.

For Le^y , the glycosidic linkages of $Fucp(1) - GlcpNAc(2)$ and $Galp(3) - GlcpNAc(2)$ are defined as above for Le^x . The torsion angle of the $Fucp(4) - Galp(3)$ linkage is defined as $\Phi = O5'''-C1'''-O2''-C2''$ and $\Psi = C1'''-O2''-C2''-C3''$.

The orientation of the hydroxy methyl groups is given by torsion angle ω (outlined in green in figure 1) ($O5''-C5''-C6''-O6''$) and the gauche/trans conformers are denoted by the standard notation **gg**, **gt**, **tg**, which corresponds to angles of -60° , 60° and 180° respectively. The sign of the torsion angles is defined in agreement with the IUPAC Commission of Biochemical Nomenclature.[4]

2.2 Molecular dynamics simulations

2.2.1 Constant nTP molecular dynamics using AMBER.

Simulations were conducted using the AMBER software package together with the GLYCAM04 force field parameters, which is specifically designed for <http://mc.manuscriptcentral.com/tandf/jenmol>

1
2 carbohydrates.[5] The solution behaviour of the trisaccharide Le^x and the
3
4 tetrasaccharide Le^y was investigated by calculating a 20 ns constant nTP MD trajectory
5
6 with explicit water molecules for each saccharide. To model the water molecules, the
7
8 TIP3 energy function was applied.[6] All hydrogen atoms were explicitly included in
9
10 the simulation however all bond lengths, including hydrogen, were constrained using
11
12 the SHAKE algorithm.[7] This allowed for an integration time step of 2 fs within the
13
14 MD simulation. The periodic box was prepared by solvation of the crystal structures
15
16 with 503 and 517 TIP3 water molecules for Le^x and Le^y respectively. This was done
17
18 using the Xleap module supplied with the AMBER package to give a concentration of
19
20 5.7 % (w/w) for Le^x and 6.9 % (w/w) for Le^y . Initial coordinates for Le^x and Le^y were
21
22 taken from a crystal structure,[8] and extraction from an antibody complex
23
24 respectively.[9] A non-bonded interactions cut-off value of 8 Å was applied during the
25
26 whole simulation. The solvated systems were energy minimised with 50 steps of
27
28 steepest descent followed by 10 000 iterations of conjugate gradient minimisation in
29
30 order to relax steric conflicts during the artificial solvation process. After 10 ps of
31
32 constant nVT MD simulation at 300 K, with initially assigned velocities from a
33
34 Boltzman distribution at the same temperature, the systems were equilibrated during a
35
36 total of 8.5 ns MD trajectories under constant nPT conditions also at 300 K. After
37
38 equilibration, ten 2 ns trajectories of production gave a total of 20 ns for both
39
40 saccharides. Coordinate sets were saved every 50 steps (0.1 ps) for subsequent analysis.
41
42
43
44
45
46
47
48
49
50

51 **3 Results and discussion**

52 **3.1 Solute Conformation from Solvated MD Simulations**

53 **3.1.1 Solute conformation.**

54
55
56
57
58
59
60
As mentioned in the introduction, due to the rigidity of the Lewis determinants, it is

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

proposed that their conformations calculated from *in vacuo* MD simulations is the most biologically active. If this is true, then we should find that the conformations of the Lewis structures, and their rigid natures, do not change when placed in solution. From our MD simulation in solution, we indeed find that the structures of both Le^x and Le^y remain rigid throughout the entire simulation time. The mean distance between Galp(3)-C1 and Fucp(1)-C1 in both Le^x and Le^y has been calculated as $4.7 \pm 0.1 \text{ \AA}$ and the distance between Galp(3)-C4 and Fucp(1)-C4 is calculated to be $4.9 \pm 0.3 \text{ \AA}$ and $4.8 \pm 0.3 \text{ \AA}$ for Le^x and Le^y respectively.

Figure 2 shows the potential energy surfaces for all of the glycosidic linkages of the solvated Le^x and Le^y molecules in terms of (Φ, Ψ) -space. If several minima in Figure 2 had been seen, this would be evidence suggesting several solute conformations (and thus several torsion angles for each linkage) were accessible under the simulation conditions. However, in our study this is not the case, only one minimum is seen for each glycosidic linkage thus showing only one stable and accessible conformation where only subtle changes in the glycosidic linkage torsion angles are observed.

Figure 3, which shows the variation of each torsion angle with respect to time throughout the simulation, also confirms this rigid nature. For all glycosidic torsion angles, only small angle variations about a mean value are observed. If several conformations were accessible, one would expect to see a larger variation in the relevant torsion angles. An example of this would be a more exaggerated case than that seen for Galp(3)-Fucp(4) glycosidic linkage, where a different angle for Φ and Ψ would be occupied for longer time durations.

[Figure 2 here]

1
2 [Figure 3 here]
3
4
5

6 7 3.1.2 Hydroxyl group orientation at Galp(3)-O6. 8

9 The orientation of the primary hydroxyl group of Galp(3) was investigated by
10 calculating the torsion angle defined by ω and monitoring the amount of time each
11 rotomer was adopted. Figure 4 shows the histograms of the MD trajectories for the 20
12 ns simulations of the tri and tetrasaccharides, and shows the distribution of the torsion
13 angles adopted.
14
15
16
17
18
19
20
21
22

23 [Figure 4 here]
24
25
26
27

28 In both Le^x and Le^y, all of the rotomers are seen, however, in both cases a **gt**
29 conformation is occupied most of the time by the primary hydroxy group which is in
30 accordance with the stereoelectronic 'gauche' effect. The ratio of **gg** : **gt** : **tg** for Le^x and
31 Le^y is observed to be 2.8 % : 85.4 % : 11.8 % and 7.3 % : 80.1 % : 12.6 % respectively.
32 A study by Serianni *et al* investigated the conformation of methylated ¹³C labelled β -D-
33 Galp in D₂O solution using ¹H NMR. They found that for the 1-methyl- β -D-[1-¹³C]Galp
34 monomer, the distribution of conformers was ~0 % : 73 % : ~25 % thus the monomer
35 adopting a **gt** conformation most of the time, thus confirming our own results.[10] The
36 **gt** conformation exhibited by galactopyranosides has long been thought to have arisen
37 from the effects of solvent interactions.[11] Comparing the distribution of
38 conformations in Le^x, Le^y and the NMR results for the β -D-Galp monomer, the presence
39 of the Fucp(1)-Glc pNAc(2) moiety has a slight effect on the conformation of the
40 primary hydroxyl group at Galp(3)-O6. This causes a reduction in the popularity of the
41 **tg** conformation and increases the population in the **gt** and **gg** conformations. The
42 Fucp(1)-Glc pNAc(2) moieties, having a rigid nature, could exhibit a directing effect on
43 the Galp(3) unit by either increasing the stability of the **gg** conformation or decreasing
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2 the stability of the *tg* conformation in an aqueous solution. This could occur via
3
4 *Fucp*(1)-*Glc**p*N*Ac*(2), providing steric bulk around *Galp*(3)-C6 or by inducing a
5
6 disturbance in the solvation sphere surrounding the area thus causing hydrogen bonds to
7
8 occur at different locations and / or geometries. In the *Galp* monomer, it is the 1-3
9
10 diaxial interaction between *Galp*-O4 and *Galp*-O6 that causes the destabilisation in the
11
12 *gg* conformation.[12] With the introduction of the *Fucp*(1)-*Glc**p*N*Ac*(2) moieties, the
13
14 oxygen at the *Galp*(3)-O4 position may also be interacting with the *Fucp*(1)-HC2
15
16 position thus providing a slight stabilising effect on the *gg* conformation. The small
17
18 increase in the population of the *gt* conformation could simply be due to a change in the
19
20 hydration sphere surrounding the *Galp*(3) ring. The presence of *Fucp*(4) in *Le*^y seems to
21
22 induce an effect of its own as the *gt* conformation is slightly less favoured in
23
24 comparison with *Le*^x. This, again, is most likely due to a difference in hydration spheres
25
26 as steric hindrance is unlikely to occur from the *Fucp*(4) moiety as it is situated at the
27
28 opposite side of the pyranose ring.
29
30
31
32
33
34
35
36
37
38
39
40
41
42

3.2 Behaviour in Solution

3.2.1 Influence of water

43
44 It has long been known that water molecules induce effects on the conformational
45
46 behaviour of biologically important molecules. In particular, they have been described
47
48 as “chaperone” molecules involved in organising, in our case the protein and the
49
50 saccharide molecules, into active conformations.[13] It is also well known that even
51
52 small changes in molecular composition can induce large effects on the hydration
53
54 sphere and hence lead to different conformations and even different biological
55
56 properties and functions. There are several examples in the literature where, upon
57
58 glycosylation, the properties of the acceptor are dramatically altered.[14-17] One
59
60 particular example of this can be seen in the work of Avenosa *et. al.* and their studies on
the glycosylation of Serine and Threonine.[18] They report that a small difference in
amino acid structure induces large changes in saccharide-amino acid torsion angle and
hydration sphere leading to rather different saccharide-amino acid conformations (and

1
2 thus oligosaccharide presentation), different water-structuring abilities and in turn
3 induce a large difference in biological function (antifreeze activity being the most
4 prominent feature).
5
6

7
8
9 It should also be pointed out that within the studies of Avenosa *et. al.* and Wong *et. al.*,
10 that although the saccharide-amino acid torsion angles changed dramatically, the
11 conformation of the saccharides did not.[15, 18, 19] As our studies were concerned only
12 with the saccharides, and the variety and complexity of the system increases
13 dramatically when considering glycosylated biomolecules, we have used saccharides
14 methylated at the anomeric position of the GlcpNAc residue. This is the position that
15 the saccharides would be conjugated to other biomolecules, and by using only a methyl
16 group we would be able to study a somewhat “generic” behaviour (if one exists!) of
17 cases where the saccharides are conjugated to other biomolecules to form the Lewis and
18 sialyl-Lewis determinants of glycolipids and glycoproteins.
19
20
21
22
23
24
25
26
27

28
29 Nevertheless, it should still be pointed out that in order for one to study more exactly
30 the true behaviour of a glycosylated biomolecule, one must consider at least some
31 portion of the biomolecule to which the saccharide is conjugated as well as the
32 behaviour and influence of the first hydration shell on structural reorganisation.
33
34
35
36
37

38 3.2.2 Diffusion behaviour.

39
40 Following the Einstein relation,[20, 21] the translational coefficient was estimated by
41 calculating the mean square displacement auto correlation function at the centre of mass
42 of the solute molecule as a function of Δt . This equation is valid for long times however
43 the function is only calculated up to $\Delta t = 20$ ps as the statistical significance decreases
44 rapidly. Figure 5 shows the translational diffusion coefficient for the solute molecules
45 and that asymptotic behaviour (i.e. random translational diffusion) is reached by the end
46 of 20 ps. Thus, the molecule has no memory of its initial translational motion.
47
48
49
50
51
52
53
54
55
56
57

58 3.2.3 Rotational diffusion.

59 The rotational diffusion time was estimated by monitoring the dipole moment vector as
60 a function of time.[21] The rotational diffusion can be roughly determined by the time

1
2 when the average angular displacement becomes uncorrelated (orthogonal).[22] Figure
3
4 5b shows the rotational diffusion of both Le^x and Le^y and that random rotational
5
6 behaviour can be observed after 1 ns.
7
8

9
10
11 [Figure 5 here]
12
13

14 15 16 3.2.4 Radial pair distributions. 17

18 Radial pair distribution functions, rdf, $g(r)$, were calculated between selected atoms of
19
20 the saccharides and all of the water molecules in order to characterise the overall
21
22 solvation of Le^x and Le^y . [22] The distribution function gives the probability of finding a
23
24 water molecule at a given distance r from the specified saccharide heteroatom, relative
25
26 to the probability expected for a random distribution at the same distance. This is to say,
27
28 the random distribution of water molecules in the bulk solution will have a constant
29
30 density where as the density of the water molecules close to the saccharide molecules
31
32 will vary depending on the solute-solvent interactions. Hydrophobic regions of the
33
34 solute molecule will induce a reduction in the solvent density where as the contrary will
35
36 be true for hydrophilic regions of the saccharide molecules. The first sphere of nearest
37
38 neighbours is indicated by a maximum in $g(r)$ at 2.8 Å and the limitations of the first
39
40 sphere are indicated by a minimum in $g(r)$ at 3.5 Å (Fig. 6). [22] The radial distributions
41
42 were obtained using the ptraj tool supplied with the AMBER program package. The
43
44 number of nearest neighbour water molecules, calculated as a function of time, for Le^x
45
46 was found to be 9.9 and 40.1 at 2.8 and 3.5 Å respectively. Hydration numbers for Le^y
47
48 were calculated as 12.1 and 44.8 respectively.
49
50
51
52
53
54
55
56
57

58 [Figure 6 here]
59
60

3.2.5 Anisotropic hydration

1
2 The next step was to investigate the anisotropic hydration of the solute. Since radial pair
3 distribution functions only carry hidden information on localised hydration sites,
4 normalised two-dimensional radial pair distributions[23] were calculated for all possible
5 shared water density sites (Os1...Ow...Os2) where Ow is the water oxygen and Os the
6 solute oxygen. In these calculations, the likelihood of finding a water molecule in a
7 well-defined local intersection area is compared and scaled to the likelihood of finding a
8 bulk water molecule in the same local intersection area. Such calculations lead to water
9 densities higher than bulk water density when structural or semi-structural water
10 molecules are present. Table 1 shows the peak density of the most important inter-
11 glycosidic water bridges found in the trajectories.
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

28 [Table 1 here]
29
30
31
32

33 In general, a “normal” profile for exposed hydroxy groups is observed. However,
34 reduced hydration for inner oxygen atoms is visible, especially for the glycosidic
35 linkages and in fact the GlcpNAc(2)-O3 in Le^y has no first hydration shell. In terms of
36 solvent density, reduced hydration is visible for Galp(3)-O4 in Le^x and Le^y due to
37 shielding from the fucose residue.
38
39
40
41
42
43
44
45
46

47 For the 1D rdf of Le^y, the oxygen atoms in the locations of Galp(3)-OH3 and Fucp(4)-
48 OH2 have small and negative $\Delta\delta$ of -0.1 ppm. The relative maximum density in the rdf
49 is not reduced in comparison with the solvent exposed hydroxyl groups. This could
50 indicate that there is no steric hindrance of solvation compared with the shielded oxygen
51 atoms such as Galp(3)-O4 and the glycosidic oxygen atoms.
52
53
54
55
56
57
58
59
60

3.2.6 Radial hydration.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The calculation of the 2D radial pair distribution functions is performed in an analogous method to the 1D distribution function,[23] but with respect to two specific solute atoms,[24] and were obtained using the Fortan programs previously published.[22] Figure 7 shows the 2D distribution functions for selected heteroatoms of Le^x and Le^y . The 2D rdfs are, in general, identical for both saccharides. However, there is an exception for the saccharide atoms close to the additional fucose residue in Le^y . The 2D cross peak for Galp(3)-O3 - Fucp(4)-O2 (Fig. 7f) shows a bridging event which is occupied for 19.2 % of the simulation time. The average residence time for a water molecule interacting with Galp(3)-O3 is 0.32 ps longer in Le^y than that in Le^x (Table 1). This indicates that the additional 0.32 ps residence per water molecule is due solely to the presence of the Fucp(4) residue.

[Figure 7 here]

The radial distribution of water molecules between Galp(3)-O2 and GlcpNAc(2)-O6 (Fig. 7a) shows that the heteroatom – water interactions are independent of each other i.e. the same water molecule is never interacting with both heteroatoms at the same time. The radial distribution between Fucp(1)-O3 and Fucp(1)-O4 shows a small increase in “shared” solvent density between the heteroatoms but this is not indicative of a dedicated shared solvent molecule interaction. It is possible that this region represents a transition state position where one solvent molecule resides between individual interactions with the heteroatoms ($Os1...Ow \rightarrow Os1...Ow.....Os2 \rightarrow Os1.....Ow...Os2 \rightarrow Ow...Os2$). The other heteroatom radial distribution functions (Galp(3)-O6 and Fucp(1)-O2 of Le^x and Le^y and also Fucp(4)-O2 and Galp(3)-O3 in Le^y) have symmetrical hydration patterns indicating a strong interaction with a shared water molecule between the two heteroatoms. The minimum corresponds to the equilibrium distance between the heteroatoms where the water molecule resides

(Os1...Ow...Os2).

3.2.7 Residential behaviour of water molecules and bridging significance.

Residence times for individual selected saccharide oxygen atoms were obtained by analysing their trajectory, on a frame by frame basis,[22] counting all water molecules which entered the first hydration shell (as defined by the minimum in the individual radial distribution function). Bridging times and significance for selected intramolecular hydrogen bonds mediated by single water molecules (structural water molecules) were calculated by identifying solvent molecules (represented by their oxygen atom) closer than 3.5 Å to the corresponding sugar atoms.

Maximum and average residence times are calculated for the corresponding (water) oxygen atoms at a distance of less than 3.5 Å (Table 1). It can be seen that in the case of Le^x, the oxygen atoms at the positions of Fucp(1)-O2 and Galp(3)-O2 have the longest maximum H₂O residence times. This would correspond to points on the molecular surface of the molecule which are the most hydrophilic and least sterically hindered and thus easily accessible by solvent molecules. Assuming that the stronger solute-solvent interactions have longer residence times, and by comparing the maximum and average residence times, these two sites can be shown to experience a multitude of interactions of varying strengths. For the Fucp(1)-O2 this could be explained by the occurrence of bridging interactions. Most Fucp(1)-O2 – OH₂ interactions are relatively weak, leading to short residence times. This could be due to the solute interacting solely with the hydroxyl group at Fucp(1)-O2, thus giving rise to a two-body interaction (Os...Ow). However, as shown in figure 7b, the solvent molecule is known to interact with at least one other solute oxygen atom (at the Galp(3)-O6 and potentially also at the Fucp(1)-O3) position. Obviously, the more interactions experienced by the solvent molecule at the

1
2 same time (i.e. one water molecule interacting with more than one oligosaccharide
3 heteroatom), the stronger the solute-solvent interaction will be and hence a longer
4 maximum residence time will be observed.
5
6
7
8
9

10
11 In most cases, the average residence time (T_{av}) for Le^y is almost identical to that of Le^x .
12 The value of T_{av} for Le^x is, on average 0.02 ps longer. However, the maximum
13 residence times (T_{max}) of Le^x and Le^y can be seen to have a greater variation in
14 comparison with each other. The T_{max} of Le^x being, on average, 0.4 ps longer. The
15 oxygen at the Galp(3)-O2 is excluded from this comparison as obviously, this
16 environment changes greatly due to the glycosidic bond to the additional Fucp(4)
17 residue in Le^y . Some heteroatoms do seem to have different properties in the two
18 molecules, in particular the Fucp(1)-O2, GlcpNAc(2)-O6 and GlcpNAc(2)-N positions.
19 It is clear that the addition of the Fucp(4) residue is the soul cause for the difference in
20 hydration pattern. As seen in Fig. 3, the glycosidic bond between Galp(3) and Fucp(4) is
21 more flexible than the other linkages of Le^x and Le^y and it is perhaps this increase in
22 flexibility which exerts steric hindrance and / or induces a disturbance on the hydration
23 shell.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

44 3.2.8 NMR results.

45
46 The chemical shifts of the hydroxy protons of oligosaccharides vary depending on
47 interactions with solvent molecules. It is for this reason that they can be used as probes
48 for conformational analysis. Reduced hydration due to steric hindrance, or interaction
49 with acetal oxygen atoms, leads to a negative value of $\Delta\delta$ ($\Delta\delta = \delta_{Oligosaccharide} -$
50 $\delta_{Monosaccharide}$) where as a positive value of $\Delta\delta$ indicates that the hydroxy proton is
51 deshielded due to being in close proximity to, and interacting with, other hydroxyl
52 groups.[25]
53
54
55
56
57
58
59
60

1
2 For the hydroxy group at Galp(3)-OH4 there is a large, negative $\Delta\delta$ of approximately -
3
4 0.5 ppm for both Le^x and Le^y. As mentioned previously, this is due to shielding
5
6 provided by the fucose residue. This shielding also reduces the degree of hydration at
7
8 the Galp(3)-OH4 position and is confirmed by the reduced maximum in the 1D rdf in
9
10 comparison with other selected oxygen atoms for Le^x and Le^y (Figure 7).
11
12
13

14
15
16 In the case of Le^x, there is a negative value of $\Delta\delta$ for Fucp(1)-OH2 and a positive value
17
18 for the $\Delta\delta$ of Fucp(1)-OH3. However, this is not shown in the MD simulation as the
19
20 simulation suggests that the $\Delta\delta$ for both positions are identical. The complex structural
21
22 hydration pattern of Le^x is not likely to be due to steric hindrance from the acetamido
23
24 group of GlcpNAc(2) as the mean distance between the CO and Fucp(1)-OH2 is 4.0 +/-
25
26 0.8 Å and during only 3.5 % of the simulation time is that distance equal to, or closer
27
28 than, the H-bonding distance of 1.9 Å. In the 2D cross peaks for Fucp(1)-O2 with
29
30 GlcpNAc(2)-CO (Fig. 7e) there is a strong bridging interaction which, for the duration
31
32 of the simulation, was populated for 34.2 % of the time. There is also a strong bridging
33
34 interaction between Fucp(1)-O2 and Galp(3)-O6 (Fig. 7b) which is occupied for 35.4 %
35
36 of the simulation time and a weaker bridging interaction between Fucp(1)-OH2 and
37
38 Fucp(1)-OH3 (Fig. 7c) which is occupied for 28.4 % of the simulation. The bridging
39
40 events only take place for short periods of time yet Fucp(1)-O2 is involved in
41
42 intramolecular water-bridging for 90 % of the total simulation time.
43
44
45
46
47
48
49
50

51 **3.3 Comparison with previous studies**

52 3.3.1 Comparison with NMR and *in vacuo* MD simulations.

53
54
55 To allow analysis of the information available, it is important to note that the
56
57 nomenclature used here is different to that used in other studies. We have used the
58
59 notation C1-O1-C_x-C_{x+1} as opposed to C1-O1-C_x-C_{x-1} for the Ψ torsion angles, which
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

can be found in many literature sources. However, in order to make fair structural comparisons, the corresponding torsion angles were measured from the final structures of our MD simulations.

In comparison with previous NMR and *in vacuo* MD studies, our structures are in agreement in that the solvated Le^x and Le^y molecules are very rigid in terms of glycosidic torsion angle flexibility.[26-28] Our structures have also been shown to exhibit a variation of $\pm 20^\circ$ from the average torsion angle, with the glycosidic linkage unique to Le^y being more flexible. NMR experiments of Le^x carried out by Bush *et. al.* [27] show that the glycosidic torsion angles fluctuate by $\pm 10-15^\circ$ around the average structure, showing less flexibility than our simulations. The published *in vacuo* simulations however show a $\pm 10-20^\circ$ variation,[28] a value comparable to that reported here.

On comparing the structures from our solvated MD studies to the average structures from NMR and average structures from *in vacuo* MD studies, the torsion angles of our structures differ by up to 20° for both Le^x and Le^y.

3.3.2 Comparison with crystal structures.

The crystal structure of Le^x, and structures of Le^x and Le^y in complex with antibodies were taken from the literature as mentioned earlier, however in the case of the Le^x dimer, the hydrogen atoms were added using the SYBYL7.0 program for visualisation purposes.

[Figure 8 here]

The hydrated crystal structure of the Le^x dimer (Fig. 8) shows that the conformation

1
2 adopted by the Galp(3)-O6 position is in the **tg** form, with ω torsion angles of -168° and
3
4 -173° , which does not agree with the results of the MD simulation reported here. In our
5
6 study, bridging events can be observed between Galp(3)-O6 – Fucp(1)-O2 and Fucp(1)-
7
8 O2 – GlcpNAc-CO, as well as shared water density between Fucp(1)-O4 – Fucp(1)-O3
9
10 as seen in the 2D rdfs of figure 7. However in the crystal structure these events are not
11
12 seen. The differences between our study and the results obtained from X-ray
13
14 crystallography are most likely due to the fact that the interactions present in the crystal
15
16 structure of the Le^x dimer are very different to those of free Le^x in aqueous solution.
17
18 The crystal structure of Le^x complexed with 291-2G3-A,[29] shows that the
19
20 trisaccharide adopts a **gt** conformation, which *is* in agreement with our MD simulation.
21
22 This makes sense as Le^x complexed with an antibody represents an example of the
23
24 trisaccharide in its biologically active form. However, we are not able to see if the
25
26 hydration sphere of the saccharide is directly responsible for providing this active
27
28 conformation or whether there are a series of directing / triggering events occurring on
29
30 complexation with the protein that cannot be seen in the crystal structure.
31
32
33
34
35
36
37
38
39

40 Finally, when comparing all of the available crystal structures together with the final
41
42 structure produced from the 20 ns MD simulation, one can see that all the structures do
43
44 not vary greatly, with the exception of the ω torsion angles (Fig. 9) and the torsion angle
45
46 of the GlcNAc-O6 position.
47
48
49
50

51
52 [Figure 9 here]
53
54
55

56
57 Crystal structures of Le^y complexed with hu3S193,[3] and CBR96,[30] show that the
58
59 tetrasaccharide adopts a **gg** conformation, but again it is not known if this conformation
60
is directed by the hydration of the tetrasaccharide or by the influence of the antibody
upon complexation. In both of the crystal structures, the Galp(3) residue appears to be

1
2 embedded in the protein surface. The oxygen atoms of the sugar molecules interact with
3
4 more water molecules than mentioned here as can be seen in the crystal structures,
5
6 however, only the water molecules which correspond to some (of the many) possible
7
8 interactions are considered.
9

10
11
12
13
14 When comparing all of the crystal structures mentioned above with the final
15
16 conformation derived from the MD simulations, one can again see that there is no great
17
18 variance (Fig. 10). As for Le^x , there is an inconsistency surrounding the “most
19
20 occupied” torsion angle adopted for GlcNAc-O6 throughout the simulation, the *gt*
21
22 rotomer, and that occupied when complexed with the protein, the *gg* rotomer.
23
24
25
26
27

28 One interesting point is that the final structure from the MD simulation shows that this
29
30 ω torsion angle is in agreement with the other crystal structures, adopting a *gg*
31
32 conformation.
33
34
35
36
37

38 [Figure 10 here]
39
40
41
42

43 3.3.3 Bridging comparison with crystal structures.

44 After studying the crystal structure of Le^x , one can see that there are many hydrogen
45
46 bond interactions taking place and the molecule has a complex hydration sphere with a
47
48 network of solute-solvent and solute-solute interactions. It appears that some of the
49
50 interactions are bridging events, for example, between the heteroatoms at the positions
51
52 GlcNAc(2)-O4 and Galp(3)-O2. Another event can be observed between Galp(3)-O2
53
54 and Galp(3)-O3. One can also see a more complex bridging event between
55
56 GlcNAc(2)-O6 and Galp(3)-O2 via 2 water molecules as shown in Figure 11 below.
57
58
59 However, one would expect the more complex bridging events (involving more than
60
one solvent molecule, Os1...Ow1...Ow2...Owx...Os2), to occur less frequently than a

1
2 simple bridging event. Simple bridging events would require that the heteroatoms and
3
4 the water molecule adopt a favourable conformation for the event to occur. By
5
6 increasing the number of “bodies” involved in the interaction, the complexity and the
7
8 specificity of the necessary geometries of the molecules would increase dramatically.
9
10 Hence, solvent molecules involved in bridging events would potentially be necessary to
11
12 provide a template and support network to induce the biologically active conformation
13
14 of the saccharide. The greater the number and complexity of the hydration sphere of the
15
16 saccharide, the more specific the biological interaction. Although not explored in our
17
18 study, it may be interesting to compare the ratios of all of the conformations of
19
20 complexed Le^x and Le^y to the ratios of the conformations of Le^x and Le^y in a water
21
22 solution. One could then see if the most commonly found active conformation of the
23
24 molecules is the most commonly occupied conformation in solution etc.
25
26
27
28
29
30
31
32

33 [Figure 11 here]
34
35
36
37

38 3.3.4 Hydration features and comparison with crystal structures.

39
40 By comparing the solute heteroatoms which exhibit bridging events using the 2D radial
41
42 distribution functions in figure 7 and studying the crystal structures already mentioned
43
44 (291-2G3-A, CBR96 and hu3s193) we are able to look for cases where a bridging event
45
46 in the MD simulation could correspond to a bridging event in the crystal structure
47
48 (either from a water molecule or a heteroatom from an amino acid residue).
49
50
51
52
53

54 On inspection of the Le^x crystal structure, we are not able to see any water molecules
55
56 which correspond to those found in the bridging events throughout the MD simulation.
57
58 This does not necessarily indicate that they do not exist, or that the simulation is invalid,
59
60 simply that they are not seen. For the crystal structure of Le^x and 291-2G3-A bridging
events between $Galp(3)-O6 - Fucp(1)-O2$ or $Fucp(1)-O4 - Fucp(1)-O3$ are not evident.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

However, a bridging event between *Fucp*(1)-O2 – *Glc**p*N*Ac*(2)-CO via the alcohol group of a tyrosine residue (Tyr 32, distances of 4.163 Å and 5.147 Å from *Fucp*(1)-O2 and *Glc**p*N*Ac*(2)-CO respectively) can be seen. For the crystal structure of *Le*^y and CBR96, only one of the bridging events shown in Figure 7 could occur. This is the event between *Galp*(3)-O6 – *Fucp*(1)-O2, again, via the alcohol group of a tyrosine residue (Tyr 35, distances of 2.684 Å and 4.528 Å from *Galp*(3)-O6 and *Fucp*(1)-O2 respectively. A distance of 2.501 Å from *Fucp*(1)-O3 is also observed). The crystal structure of *Le*^y complexed with hu3s193 seems to be more faithful to the events described by the MD simulation. Two water molecules from the crystal structure (molecule 72 distances of 2.758 Å and 3.099 Å from *Fucp*(4)-O3 and *Fucp*(4)-O2 respectively and molecule 278 distances of 4.332 Å and 2.957 Å from *Fucp*(4)-O3 and *Fucp*(4)-O2 respectively) are both in a position to be involved with bridging the heteroatoms *Fucp*(4)-O3 – *Fucp*(4)-O2. The carbonyl of a serine residue (Ser 104, distances of 3.110 Å and 2.815 Å from *Fucp*(1)-O4 and *Fucp*(1)-O3 respectively) can also be in a position to allow a bridging event to occur between *Fucp*(1)-O4 – *Fucp*(1)-O3. Water molecules can also be seen to allow bridging interactions between the heteroatoms *Fucp*(1)-O2 – *Glc**p*N*Ac*(2)-CO (molecules 19 distances of 2.966 Å and 4.019 Å from *Fucp*(1)-O2 – *Glc**p*N*Ac*(2)-CO respectively) and *Fucp*(4)-O2 – *Galp*(3)-O3 (molecule 72 distances of 2.957 Å and 4.438 Å from *Fucp*(4)-O2 and *Galp*(3)-O3 respectively). An interesting point to note is the involvement of water molecule 72 in both the *Fucp*(4)-O3 – *Fucp*(4)-O2 and *Fucp*(4)-O2 – *Galp*(3)-O3 events and could represent a four “bodied” system.

In the cases of the *Le*^x, 291-2G3-A and CBR96 it is clear that the bridging events seen in the MD simulation are not seen. As previously mentioned, this could be because they are not seen in the crystal structures or they do not exist. If the latter is the case, this would imply that the water molecules of the free solute molecule would contribute to

1
2 the change in entropy upon association of the two components of the system and that
3
4 the active conformation is more likely to be induced by the environment experience by
5
6 the solute molecule. In the case of hu3s193, most of the bridging events investigated in
7
8 the MD simulation can be seen to occur tentatively in the crystal structure. This could
9
10 imply that the active solute conformation is the conformation found in free solution and
11
12 that the water molecules involved in bridging events in the free solution will be
13
14 maintaining the correct molecular conformation for binding with the antibody. In the
15
16 case of the bridging event occurring between Fucp(1)-O4 – Fucp(1)-O3, the bridging
17
18 water molecules observed in the MD simulation are displaced by the serine residue.
19
20 Their displacement would imply that they are contributing to the increase in enthalpy of
21
22 the system upon association of the solute with the antibody complex. The water
23
24 molecules which are not displaced are either found to be interacting with areas of the
25
26 solute molecule which remain to be exposed to the bulk solvent upon binding with the
27
28 antibody, or they are directly involved with interacting with the antibody thus indicating
29
30 that they contribute to the affinity of the system.
31
32
33
34
35
36
37
38
39

40 **4. Conclusions**

41
42
43
44 From our solution MD simulations of Le^x and Le^y we have shown that the solvated
45
46 structures of the oligosaccharides remain rigid throughout the 20 ns simulation time and
47
48 are comparable to previous NMR and *in vacuo* MD simulations.
49
50
51
52
53

54 This finding implies that these important oligosaccharides remain as targets for
55
56 immunotherapeutic strategies for countering carcinoma as therapeutic agents would not
57
58 have to account for large conformational changes expressed by the Lewis antigen
59
60 targets and hence are able to be structurally specific. Structural comparisons have also
been made to structures from four crystal structures currently available. Water molecule

1
2 residence times and bridging events between oligosaccharide heteroatoms have also
3
4 been investigated, and in particular bridging events have been found to occur between
5
6 Fucp(1)-O2 - GlcpNAc(2)-CO of both Le^x and Le^y and Fucp(4)-O2 - Galp(3)-O3 of
7
8 Le^y. By studying the crystal structures further, we have also been able to identify
9
10 occasions where bridging water molecules may be present and thus potentially inducing
11
12 structural changes in the solute allowing it to adopt a more active conformation. We
13
14 have also identified occasions where bridging water molecules are not present in the
15
16 crystal structures and have therefore been displaced from the solute thus contributing to
17
18 the entropy of the interaction. Displacement of bridging water molecules by protein
19
20 heteroatoms may also be seen.
21
22
23
24
25
26
27

28 Our work contributes to the understanding of carbohydrate interactions with water,
29
30 which may be of importance for the biologically events taking place in the glycocalyx.
31
32
33
34

35 5. Acknowledgements

36
37
38
39 We wish to thank the GlycoGold research training network, a part of the 6th research
40
41 framework program of the European Union, contract number MRTN-CT-2004-005645
42
43 and to the Sialinsäuregesellschaft e. V., Kiel, for a travel stipend for A. Fuchs.
44
45
46
47
48

49 6. References

- 50
51
52
53 [1] J. Sakamoto, K. Furukawa, C. Cordon-Cardo, B.W. Yin, W.J. Rettig, H.F. Oettgen, L.J. Old and K.O.
54 Lloyd, *Expression of Lewis^a, Lewis^b, X, and Y blood group antigens in human colonic tumors*
55 *and normal tissue and in human tumor-derived cell lines*, *Cancer. Res.* 46 (1986), 1553-61.
56 [2] E. Yuriev, W. Farrugia, A.M. Scott and P.A. Ramsland, *Three-dimensional structures of carbohydrate*
57 *determinants of Lewis system antigens: implications for effective antibody targeting of cancer*,
58 *Immunol. Cell. Biol.* 83 (2005), 709-17.
59 [3] P.A. Ramsland, W. Farrugia, T.M. Bradford, M.P. Hogarth and A.M. Scott, *Structural convergence of*
60 *antibody binding of carbohydrate determinants in Lewis Y tumor antigens*, *J. Mol. Biol.* 340
(2004), 809-18.
[4] A.D. McNaught, *International Union of Pure and Applied Chemistry and International Union of*
Biochemistry and Molecular Biology. Joint Commission on Biochemical Nomenclature.
Nomenclature of carbohydrates, *Carbohydr. Res.* 297 (1997), 1-92.

- 1
2 [5] D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, K.M.
3 Merz, B. Wang, D.A. Pearlman, M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V.
4 Hornak, G. Cui, P. Beroza, C. Schafmeister, J.W. Caldwell, W.S. Ross and A. P.A. Kollman
5 (2004), University of California, San Francisco.
- 6 [6] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey and M.L. Klein, *Comparison of simple*
7 *potential functions for simulating liquid water*, J. Chem. Phys. 79 (1983), 926-935.
- 8 [7] W.F.v. Gunsteren and H.J.C. Berendsen, *Algorithms for macromolecular dynamics and constraint*
9 *dynamics*, Mol. Phys. 34 (1977), 1311-1327.
- 10 [8] S. Perez, N. Mouhous-Riou, N.E. Nifant'ev, Y.E. Tsvetkov, B. Bachet and A. Imberty, *Crystal and*
11 *molecular structure of a histo-blood group antigen involved in cell adhesion: the Lewis x*
12 *trisaccharide*, Glycobiology 6 (1996), 537-42.
- 13 [9] W.F. Paul A. Ramsland, Tessa M. Bradford, P. Mark Hogarth and Andrew M. Scott, *Structural*
14 *Convergence of Antibody Binding of Carbohydrate Determinants in Lewis Y Tumor Antigens*, J.
15 Mol. Biol. 340 (2004), 809-818.
- 16 [10] Q. Pan, T. Klepach, I. Carmichael, M. Reed and A.S. Serianni, *4J(COCCH) and 4J(CCCCH) as*
17 *probes of exocyclic hydroxymethyl group conformation in saccharides*, J. Org. Chem. 70 (2005),
18 7542-9.
- 19 [11] N.K. Devries and H.M. Buck, *Different Rotamer Populations around the C-5-C-6 Bond for Alpha-*
20 *D-Galactopyranosides and Beta-D-Galactopyranosides through the Combined Interaction of the*
21 *Gauche and Anomeric Effects - a 300-Mhz H-1-Nmr and Mndo Study*, Carbohydr. Res. 165
22 (1987), 1-16.
- 23 [12] Y. Nishida, H. Ohrui and H. Meguro, *H-1-Nmr Studies of (6r)-Deuterated and (6s)-Deuterated D-*
24 *Hexoses - Assignment of the Preferred Rotamers About C5-C6 Bond of D-Glucose and D-*
25 *Galactose Derivatives in Solutions*, Tetrahedron Lett. 25 (1984), 1575-1578.
- 26 [13] R.U. Lemieux, *How water provides the impetus for molecular recognition in aqueous solution*,
27 Accounts. Chem. Res. 29 (1996), 373-380.
- 28 [14] X. Huang, J.J. Barchi, Jr., F.D. Lung, P.P. Roller, P.L. Nara, J. Muschik and R.R. Garrity,
29 *Glycosylation affects both the three-dimensional structure and antibody binding properties of*
30 *the HIV-1IIIB GP120 peptide RP135*, Biochemistry 36 (1997), 10846-56.
- 31 [15] W.G. Wu, L. Pasternack, D.H. Huang, K.M. Koeller, C.C. Lin, O. Seitz and C.H. Wong, *Structural*
32 *study on O-glycopeptides: Glycosylation-induced conformational changes of O-GlcNAc, O-*
33 *LacNAc, O-sialyl-LacNAc, and O-sialyl-lewis-X peptides of the mucin domain of MadCAM-1*, J.
34 Am. Chem. Soc. 121 (1999), 2409-2417.
- 35 [16] D.M. Coltart, A.K. Royyuru, L.J. Williams, P.W. Glunz, D. Sames, S.D. Kuduk, J.B. Schwarz, X.T.
36 Chen, S.J. Danishefsky and D.H. Live, *Principles of mucin architecture: Structural studies on*
37 *synthetic glycopeptides bearing clustered mono-, di-, tri-, and hexasaccharide glycodomains*, J.
38 Am. Chem. Soc. 124 (2002), 9833-9844.
- 39 [17] M.R. Pratt and C.R. Bertozzi, *Synthetic glycopeptides and glycoproteins as tools for biology*, Chem.
40 Soc. Rev. 34 (2005), 58-68.
- 41 [18] F. Corzana, J.H. Busto, G. Jimenez-Oses, M. Garcia de Luis, J.L. Asensio, J. Jimenez-Barbero, J.M.
42 Peregrina and A. Avenoza, *Serine versus threonine glycosylation: the methyl group causes a*
43 *drastic alteration on the carbohydrate orientation and on the surrounding water shell*, J. Am.
44 Chem. Soc. 129 (2007), 9458-67.
- 45 [19] F. Corzana, J.H. Busto, G. Jimenez-Oses, J.L. Asensio, J. Jimenez-Barbero, J.M. Peregrina and A.
46 Avenoza, *New insights into alpha-GalNAc-Ser motif: influence of hydrogen bonding versus*
47 *solvent interactions on the preferred conformation*, J. Am. Chem. Soc. 128 (2006), 14640-8.
- 48 [20] M.P. Allen and D.J. Tildesley, *Computer Simulations of Liquids*, Clarendon Press, Oxford, U.K.,
49 1987.
- 50 [21] C.H.d.P. S Engelsen, Serge Perez, *Molecular relaxation of sucrose in aqueous solutions: how a*
51 *nanosecond molecular dynamics simulation helps to reconcile NMR data* J. Phys. Chem. 99
52 (1995), 13334-13351.
- 53 [22] S.B. Engelsen and S. Perez, *The hydration of sucrose*, Carbohydr. Res. 292 (1996), 21-38.
- 54 [23] C. Andersson and S.B. Engelsen, *The mean hydration of carbohydrates as studied by normalized*
55 *two-dimensional radial pair distributions*, J. Mol. Graph. Model. 17 (1999), 101-5, 131-3.
- 56 [24] S.B. Engelsen and S. Perez, *Unique Similarity of the Asymmetric Trehalose Solid-State Hydration*
57 *and the Diluted Aqueous-Solution Hydration*, J. Phys. Chem. B 104 (2000), 9301-9311.
- 58 [25] S. Bekiroglu, L. Kenne and C. Sandstrom, *NMR study on the hydroxy protons of the Lewis X and*
59 *Lewis Y oligosaccharides*, Carbohydr. Res. 339 (2004), 2465-8.
- 60 [26] H.F. Azurmendi, M. Martin-Pastor and C.A. Bush, *Conformational studies of Lewis X and Lewis A*
trisaccharides using NMR residual dipolar couplings, Biopolymers 63 (2002), 89-98.
- [27] K.E. Miller, C. Mukhopadhyay, P. Cagas and C.A. Bush, *Solution structure of the Lewis x*
oligosaccharide determined by NMR spectroscopy and molecular dynamics simulations,
Biochemistry 31 (1992), 6703-9.

- 1
2 [28] A. Imberty, E. Mikros, J. Koca, R. Mollicone, R. Oriol and S. Perez, *Computer simulation of histo-*
3 *blood group oligosaccharides: energy maps of all constituting disaccharides and potential*
4 *energy surfaces of 14 ABH and Lewis carbohydrate antigens*, Glycoconj. J. 12 (1995), 331-49.
5 [29] A.M. van Roon, N.S. Pannu, J.P. de Vrind, G.A. van der Marel, J.H. van Boom, C.H. Hokke, A.M.
6 Deelder and J.P. Abrahams, *Structure of an anti-Lewis X Fab fragment in complex with its Lewis*
7 *X antigen*, Structure 12 (2004), 1227-36.
8 [30] P.D. Jeffrey, J. Bajorath, C.Y. Chang, D. Yelton, I. Hellstrom, K.E. Hellstrom and S. Sheriff, *The x-*
9 *ray structure of an anti-tumour antibody in complex with antigen*, Nat. Struct. Biol. 2 (1995),
10 466-71.

11 **Figure 1. Structure of Lewis x (Le^x) trisaccharide (1a) and Lewis y (Le^y) tetrasaccharide (1b).**

12 **Figure 2. Potential energy surfaces for the glycosidic linkages of Le^x Fucp(1)-GlcPNAc(2) (a),**
13 **GlcPNAc(2)-Galp(3) (b) and Le^y Fucp(1)-GlcPNAc(2) (c), GlcPNAc(2)-Galp(3) (d), Galp(3)-Fucp(4)**
14 **(e).**

15 **Figure 3. variation with time for the glycosidic linkages of Le^x (left) and Le^y (right). (a) Φ Fucp(1)-**
16 **GlcPNAc(2), (b) Ψ Fucp(1)-GlcPNAc(2), (c) Φ GlcPNAc(2)-Galp(3), (d) Ψ GlcPNAc(2)-Galp(3), (e)**
17 **Φ Galp(3)-Fucp(4), (f) Ψ Galp(3)-Fucp(4).**

18 **Figure 4. Histograms of the distribution of the population of torsion angles occupied for the**
19 **glycosidic linkage of O5-C5-C6-O6, ω , of the Galp(3) residue of Le^x (a) and Le^y (b).**

20 **Figure 5. Rotational (a) and translational (b) diffusion coefficients for Le^x and Le^y.**

21 **Figure 6. 1D Radial distribution functions (rdf) for selected oxygen atoms of Le^x (a) and Le^y (b).**

22 **Figure 7. 2D radial distribution functions for the selected hetero atoms (a) Galp(3)-O2 and**
23 **GlcPNAc(2)-O6 (b) Galp(3)-O6 and Fucp(1)-O2 (c) Fucp(4)-O2 and Fucp(4)-O3 (d) Fucp(1)-O3 and**
24 **Fucp(1)-O4 (e) Fucp(1)-O2 and GlcNAc(2)-CO of Le^x and Le^y and (f) Fucp(4)-O2 and Galp(3)-O3**
25 **of Le^y.**

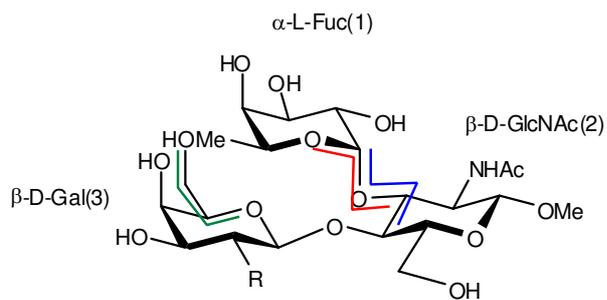
26 **Figure 8. Crystal structure of the hydrated Le^x dimer.**

27 **Figure 9. Comparison of the final structure from the 20 ns Le^x MD simulation (carbons coloured in**
28 **dark blue) with the Le^x crystal structure (carbons coloured in green), [8] and the structure**
29 **extracted from the 291-2G3-A complex (carbons coloured in light blue). [29]**

30 **Figure 10. Comparison of the final structure from the 20 ns Le^x MD simulation (carbons coloured**
31 **in dark blue) with the complexes hu3S193 (carbons coloured in light blue) and CBR96 (carbons**
32 **coloured orange).**

33 **Figure 11. Complex bridging event seen in the crystal structure of Lex. The hydrogen bonds are**
34 **represented by the black dashed lines.**

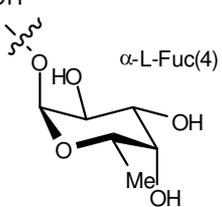
35 **Table 1. Maximum and average residence times for water molecules approaching within 3.5 Å of**
36 **the oligosaccharide heteroatom listed.**



13
14

1a Le^x R = OH

1b Le^y R =

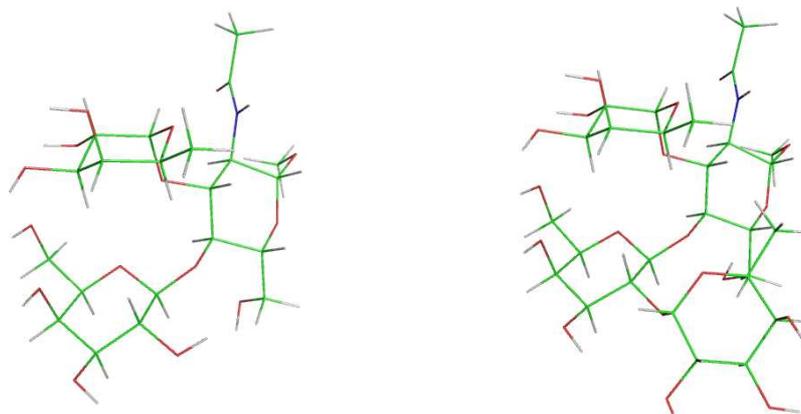


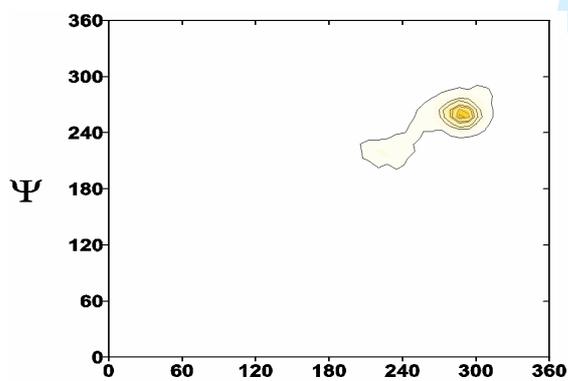
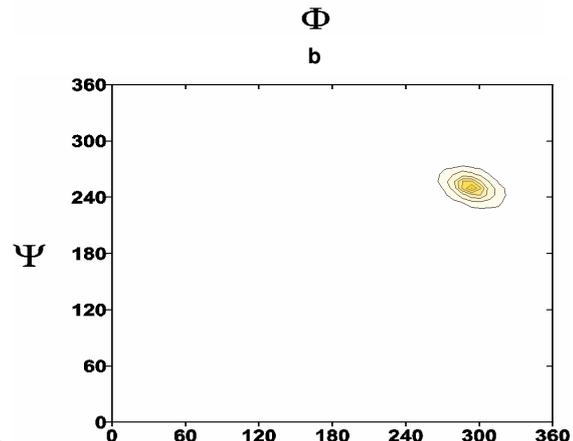
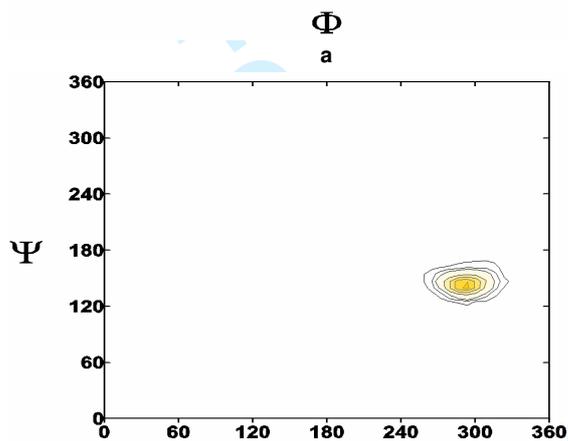
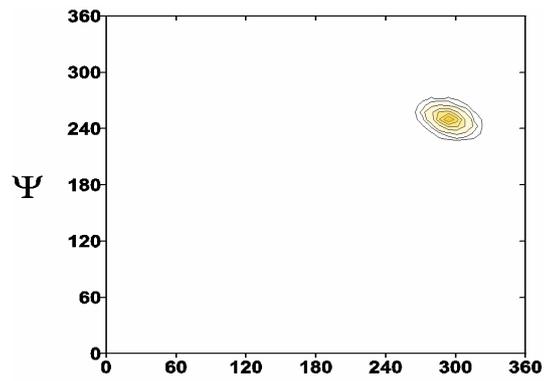
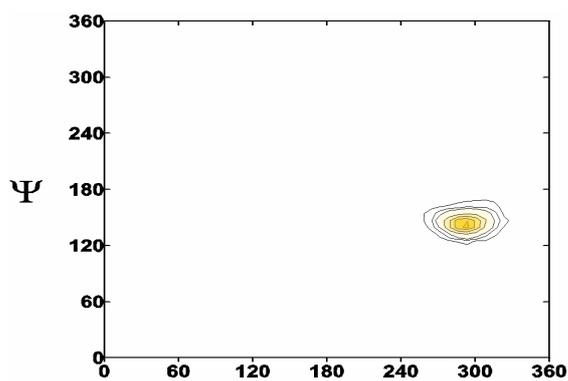
22
23

Lewis X trisaccharide (Le^x) [β -D-Galp-(1-4)[α -L-Fucp-(1-3)]- β -D-GlcpNAc]

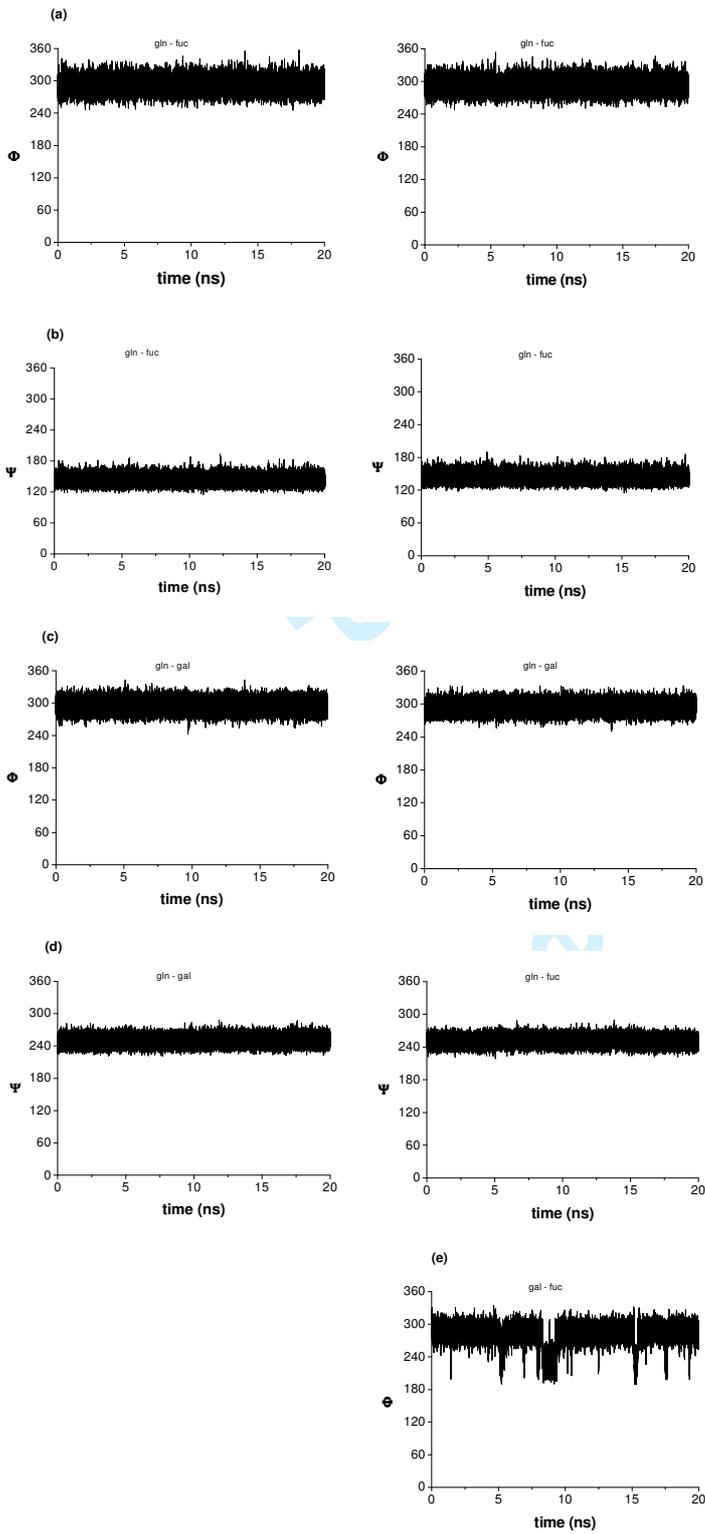
24
25

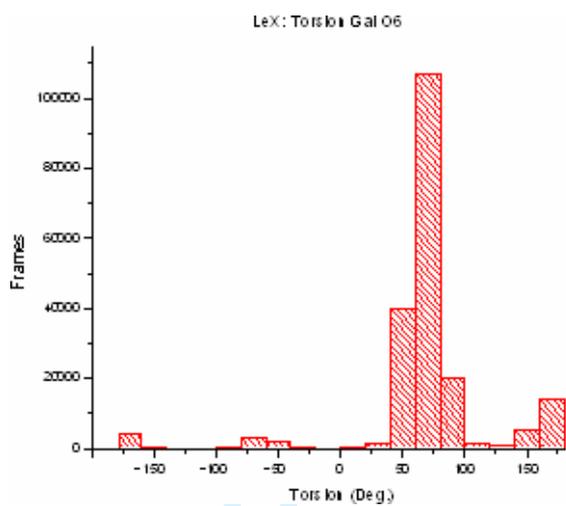
Lewis Y trisaccharide (Le^y) [α -L-Fucp-(1-2)- β -D-Galp-(1-4)[α -L-Fucp-(1-3)]- β -D-GlcpNAc]



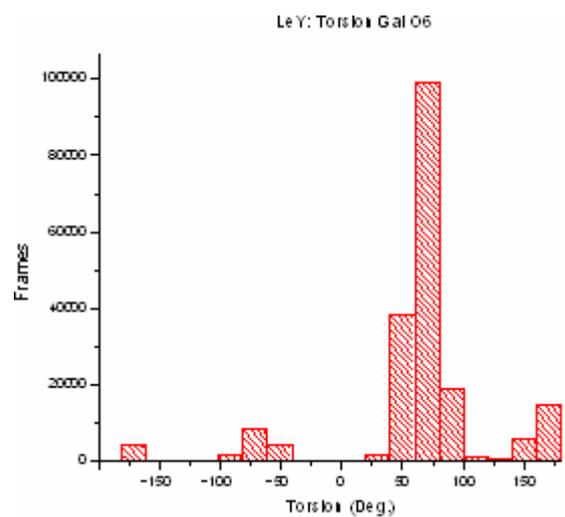


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

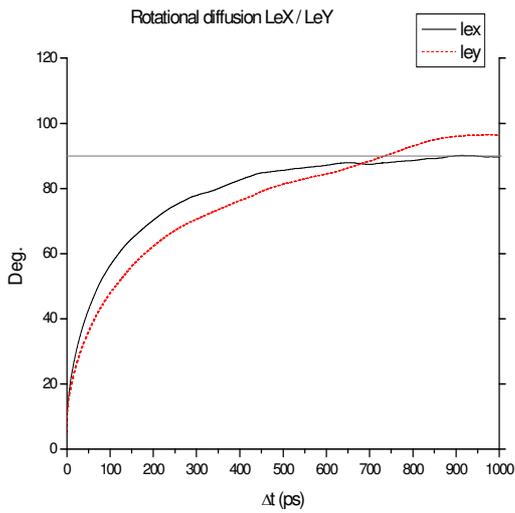




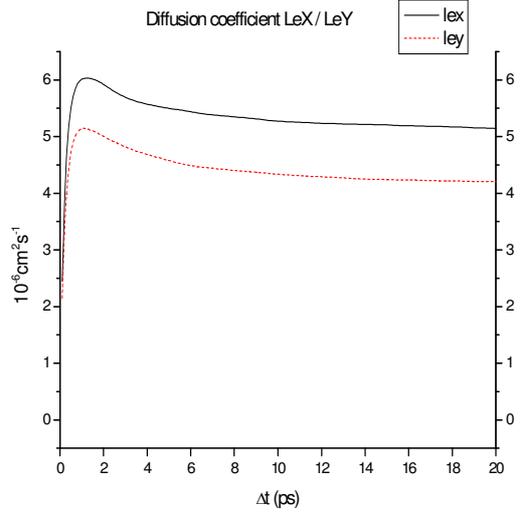
(a)



(b)



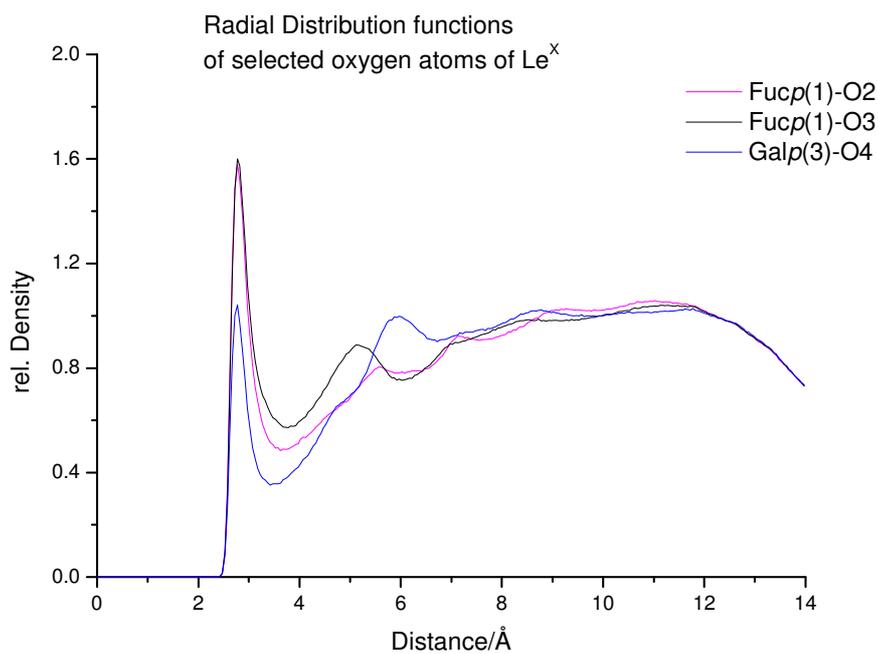
(a)



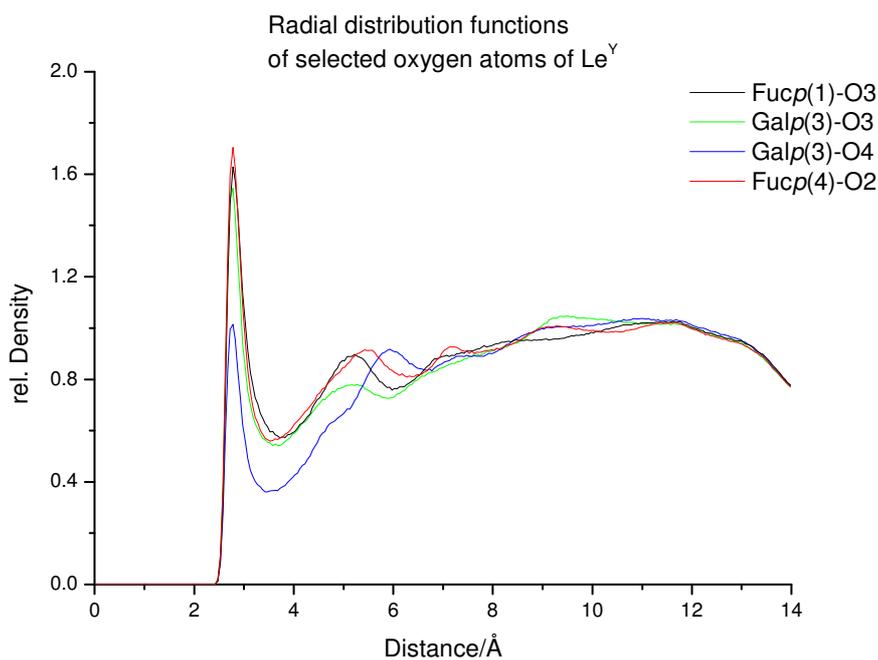
(b)

Peer Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

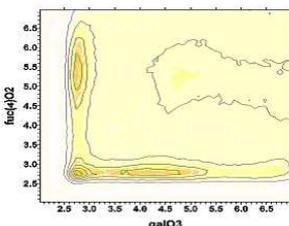
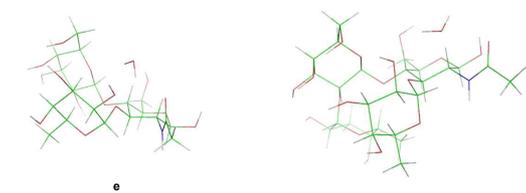
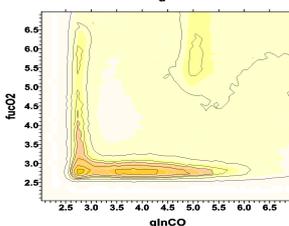
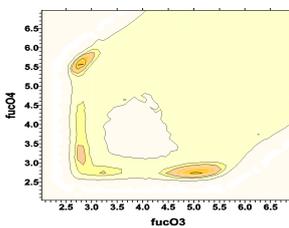
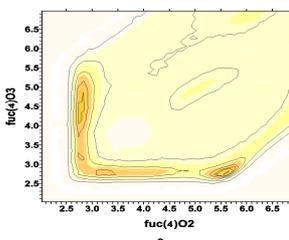
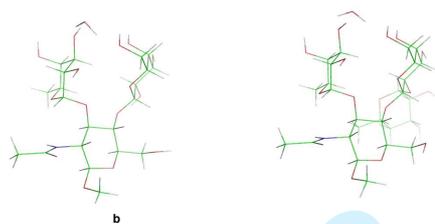
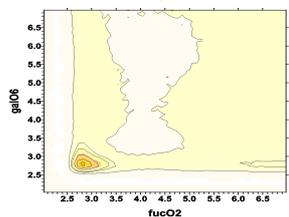
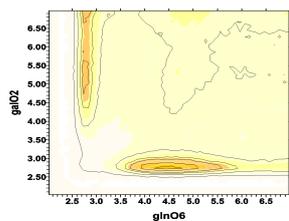


a

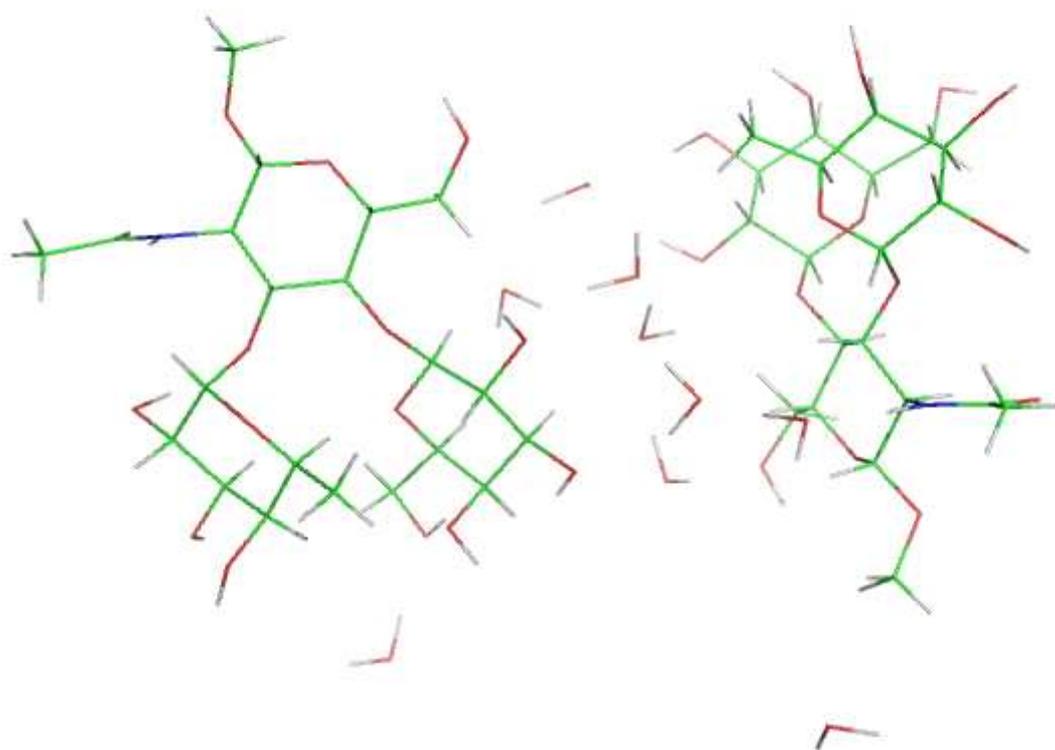


b

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

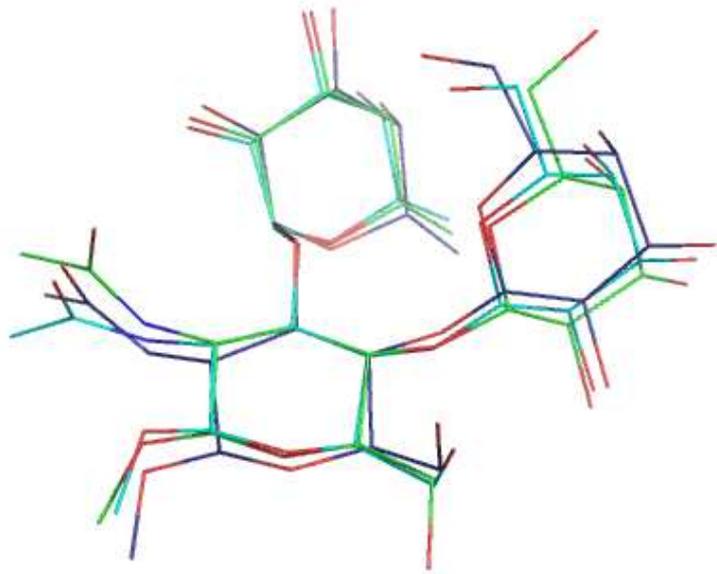


Peer Review Only

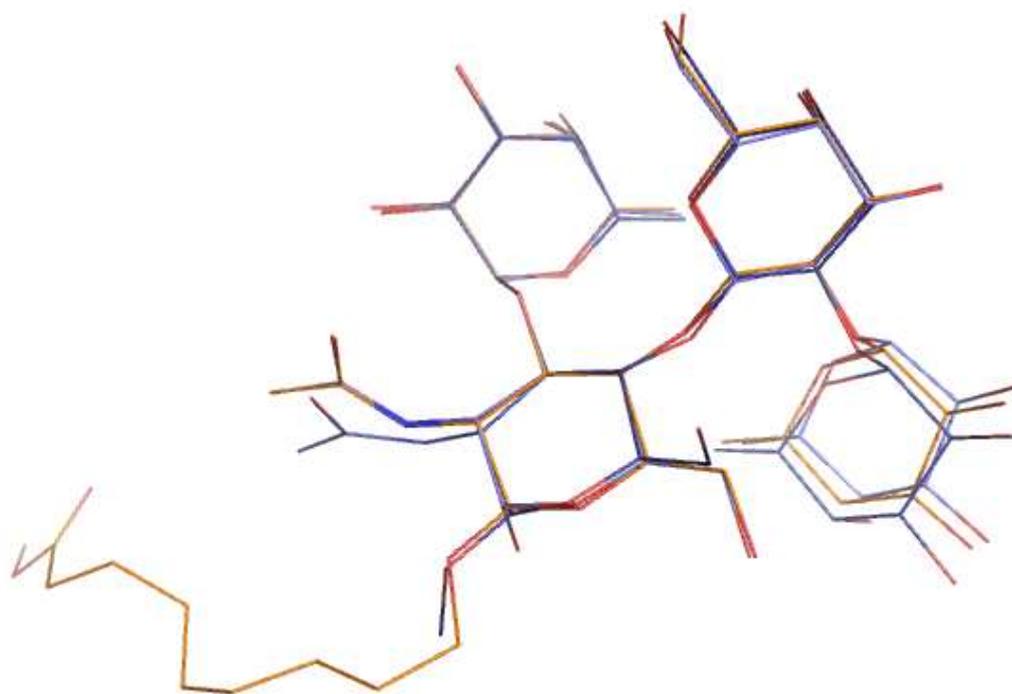


View Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

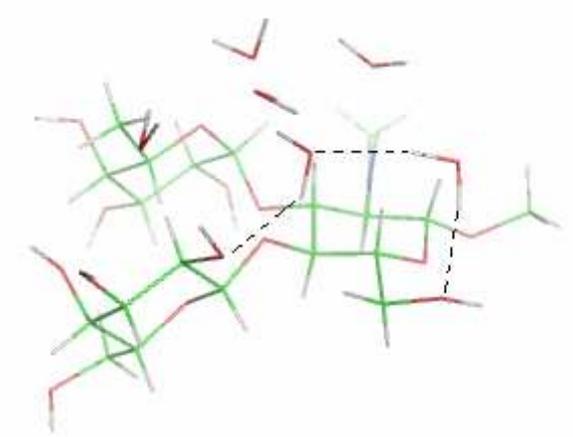


Review Only



Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



or Peer Review Only

atom	Le ^x			Le ^y		
	T _{max} / ps	T _{av} / ps	significance	T _{max} / ps	t _{av} / ps	Significance
Fucp(1)-O2	62.4	2.74	99	42.4	2.65	99
O3	27.7	2.23	99	33.6	2.26	99
O4	45.0	3.12	99	38.8	3.13	99
O5	6.1	0.42	98	6.2	0.41	98
GlcNAcp(2)-	16.3	0.86	95	13.7	0.86	95
O1						
O3	18.4	0.87	74	15.1	0.85	74
O4	7.6	0.51	61	3.1	0.27	1
O5	12.1	0.64	89	11.5	0.65	89
O6	25.9	2.28	99	35.8	2.67	99
CO	22.2	2.15	99	23.4	2.09	99
N	17.4	1.20	92	26.8	1.21	92
Galp(3)-O2	61.0	3.19	99	7.7	0.44	56
O3	30.9	2.40	99	28.1	2.72	99
O4	40.9	2.31	98	45.0	2.35	98
O5	2.9	0.23	35	3.90	0.24	36
O6	31.8	2.45	99	34.7	2.40	99
Fucp(4)-O2				45.8	2.66	99
O3				21.9	2.05	99
O4				46.4	2.82	99
O5				23.0	1.18	88