THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

pubs.acs.org/JPCB

Interacting Lewis-X Carbohydrates in Condensed Phase: A First-Principles Molecular Dynamics Study

Rinaldo Zucca,[†] Mauro Boero,[‡,§,#] Carlo Massobrio,[‡] Carla Molteni,[¶] and Fabrizio Cleri*[,†]

[†]Institut d'Electronique, Microelectronique et Nanotechnologie (Cnrs UMR 8520), Université de Lille I, F-59652 Villeneuve d'Ascq, France

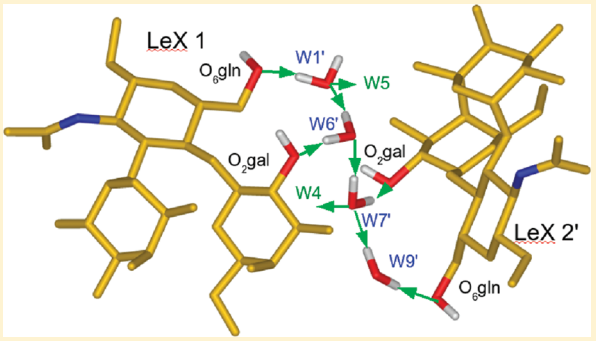
[‡]Institut de Physique et Chimie des Materiaux de Strasbourg (CNRS UMR 7504), F-64705 Strasbourg, France

[§]Research Center for Integrated Science, Japan Advanced Institute of Science and Technology (JAIST), 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan

[#]CREST, Japan Science and Technology Agency, Sanban-cho, Tokyo 102-0075, Japan

[¶]King's College London, Physics Department, Strand, London WC21R 2LS, United Kingdom

ABSTRACT: We performed first-principles molecular dynamics calculations at finite temperature, to study the interacting conformations of Lewis-X (LeX) trisaccharides in the crystalline phase. The calculated cell parameters and detailed atomic structure of the LeX molecule compare well to the experimental data obtained by X-ray diffraction. We identify and characterize the hydrogen-bond network, responsible for the mutual interaction of the LeX pairs, whereas we find the intramolecular conformation and stability to be mainly assured by dispersion forces. The relative contributions to the crystallization energy of the hydrogen bonds and of the dispersion forces are defined and quantified. From this study, candidate configurations for the fully hydrated, in vivo structures of homotypic LeX−LeX interactions at cell surfaces can be proposed. We discuss how these configurations could also be relevant for the adhesion and self-assembly of nanostructures.



INTRODUCTION

Carbohydrates are among the most important molecules for life. The surface of mammalian cells is covered by a dense layer of protein- and lipid-conjugated carbohydrates, the *glycocalix*,[1] known to be involved in the control of many normal and pathological processes.[2,3] Recently it has been proposed that some of these processes, in particular cell adhesion,[4] might involve a direct interaction between carbohydrates.[5,6] For the latter to be possible, it is required that such carbohydrate species display both adhesive and self-recognition capabilities, which, in turn, make them more generally attractive beyond the biological domain, as potential candidates to mediate surface-specific interactions. Such adhesion and recognition capabilities could, for example, be exploited to drive the selective self-assembly of nanoscale objects (nanoparticles, nanowires, etc.),[7] once coated with some carbohydrate-bearing molecular layer.

Notably, while carbohydrates are chemically much simpler than, say, proteins, their huge conformational variability makes even the finest details of the molecular structure relevant for recognition and adhesion. It was noted, for example, by Sharon and Lis that, while four nucleotides can make only 24 distinct tetranucleotides, four different monosaccharides can give rise to 35 560 unique tetrasaccharides.[8] As a prominent example of this specifity, the orientation of the hydroxyl groups in stereoisomers is generally acknowledged to be an essential feature of the diverse functional capabilities of sugars in the biological medium, leading to such far-reaching effects as the structuring of the hydrogen-bonding network around the solute molecule,[9] and their interaction with other biomolecules.[10,11] However, such subtle details often escape a direct experimental determination, and, on the other hand, they require highly accurate theoretical models to be fully captured.

One of the more extensively studied molecules in this respect is the Lewis-X (**LeX**), a trisaccharide found as the terminal moiety of several glycolipids and glycoproteins at the cell surface, and involved in cell adhesion and recognition processes.[3,4] The **LeX** molecule (Figure 1) is composed of an α-fucose (**Fucα**) and a β-galactose (**Galβ**) ring, linked to a β-glucose ring carrying a N-acetyl ($-$**NH**$-$**CO**$-$**CH$_3$**) group at ring position 2 (**GlcNAcβ**). In glycochemistry this sequence is indicated as **Galβ(1−4)-Fucα(1−3)GlcNAcβ**, meaning that **Fucα** and **Galβ** form glycosidic ($-$**C**$-$**O**$-$**C**$-$) bonds via their carbon C1 atom, respectively, and the C3 and C4 atoms of the glucose ring. (In the rest of the paper, the labels α and β, indicating isomeric forms, will be dropped since they will remain constant throughout.)

Received: June 14, 2011
Revised: August 28, 2011
Published: September 15, 2011

dx.doi.org/10.1021/jp2055816 | *J. Phys. Chem. B* 2011, 115, 12599–12606

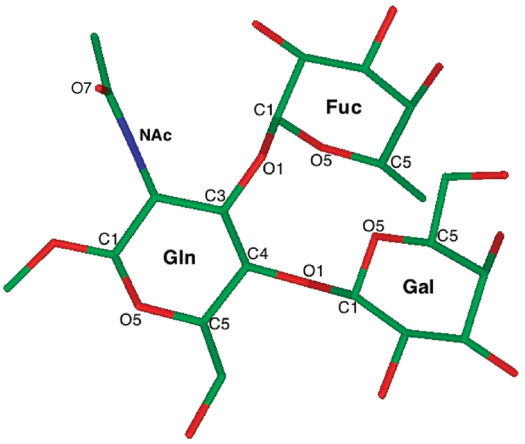


Figure 1. Stick model of the LeX trisaccharide, after relaxation by damped molecular dynamics with the CPMD code. Red = oxygen atoms, green = carbon, blue = nitrogen. Hydrogen atoms are not represented in the figure. Atom labels are given only for relevant atoms included in Table 1.

Due to its relevance in normal and pathologic cell adhesion, embryogenesis, cancer spreading[12,13] (for example, Ca^{2+}–mediated **LeX**–**LeX** homotypic interaction has been proposed as the main carrier of adhesion in the development of some forms of cancer, both by in vivo and in vitro experiments[14]), the **LeX**–**LeX** interaction was experimentally investigated by several techniques, among which nuclear magnetic resonance,[15,16] atomic force microscopy,[17] and surface plasmon resonance.[18] Such experiments provided relevant information about the type of **LeX**–**LeX** interaction (homotypic, and Ca^{2+}–mediated), and its strength, resulting in a quite weak dissociation constant $K_d = 5.4 \times 10^{-7}$ M, and an adhesion force of about 20 pN. However, the molecular configuration of the interacting pair could not be deduced from such techniques. A step forward in this direction was made possible by x-ray diffraction, a method requiring a long and delicate crystallization process to obtain suitable samples. Two independent x-ray studies, published almost simultaneously,[19,20] allowed to identify the **LeX** crystal structure in condensed phase.

The present study is aimed at complementing the experimental information by first-principles molecular dynamics (MD) simulations, providing a direct atomic-scale insight into the nature of the **LeX**–**LeX** interactions. A few previous studies of **LeX**–**LeX** interaction in water have been carried out by molecular mechanics with simplified, empirical force fields.[21a,b] However, thorough studies by first–principles MD of the crystalline structure have already proven very useful in assessing carbohydrate conformations,[21c] hence also in the present case the crystalline structure could be indeed suggestive of interacting **LeX**–**LeX** configurations. For the purpose of extracting from the crystallized structures molecular configurations relevant to both in vivo interactions and nanoparticle self–assembly, it should be first noted that all crystal structures include some water molecules, but no Ca^{2+} or other ions. The extent to which such structures compare to a fully hydrated, in vivo structure therefore requires further investigation. On the other hand, albeit the geometrical constraints imposed by the condensed phase could lead to differences in the **LeX** conformations, with respect to **LeX** in vivo (which are often associated in complex glycans with other molecules and proteins), we suggest that these same, highly constrained structures could be especially meaningful for the study of dense molecular coatings on nanoparticle surfaces.

COMPUTATIONAL DETAILS

From the crystallographic data of Perez et al.,[19] it is known that **LeX** crystallizes in the monoclinic space group $P2_1$ with unit cell parameters $a = 12.147(6)$ Å, $b = 27.552(9)$ Å, $c = 8.662(6)$ Å, and $\beta = 91.71°$. The asymmetric unit contains two independent **LeX** and 9 water molecules, and it is repeated twice in the irreducible unit cell, for a total of 350 atoms (4 N, 84 C, 60 O and 148 H, plus 18 O and 36 H of water). The following set of MD simulations makes use of one irreducible unit cell, with periodic boundary conditions, and without any constraints on atom dynamics and electronic degrees of freedom. Both the $T = 0$ K structural relaxation and subsequent MD simulations at finite temperature were carried out by means of the Car–Parrinello MD method, as implemented in the CPMD code.[22] All the simulations were done within the density functional theory (DFT) in the BLYP generalized–gradient approximation for the exchange–correlation functional.[23] The valence–core interaction was described by Troullier–Martins norm-conserving pseudopotentials.[24] The electronic wave functions were expanded in a plane wave basis set with a kinetic energy cutoff of 70 Ry. Brillouin-zone sampling was limited to the Γ-point. The use of a single k-point ($k = 0$) in the present case is acceptable, since the supercell is large enough to ensure an adequate sampling of the k-space, as extensively discussed in the literature.[25] Since DFT is known to fail in describing long–range dispersion forces, van der Waals (VdW) interactions were added by means of atom–centered spherical potentials, in the form parametrized by Grimme,[26] with an R^{-6} radial dependence. For the MD simulations, a time step of 0.096 fs, and a fictitious electronic mass of 0.165 amu for the numerical integration of the equations of motion ensured a good control of the conserved quantities. For structural relaxation, damped MD was continued until the maximum forces on each atom were in the range of 0.05 eV/Å. This threshold accounts for the extremely weak convergence of floppy modes in carbohydrates. For finite–temperature equilibration, we ran about 50 ps dynamics at $T = 300$ K, with the cell size and shape fixed at the $T = 0$ K minimum–energy configuration (see below), using the same MD parameters as above. Statistics were collected in the last 40 ps of each simulation.

Minimization of the cell parameters by means of damped MD at $T = 0$ K was carried out sequentially: we started with b and c fixed at the respective experimental values, and minimized the atomic and electronic structure as a function of the volume of the cell, by varying a; then, with a at its energy minimum, we minimized with respect to variation of b, and finally with respect to c. We obtained values of the cell parameters a, b, c, and β, nearly identical to the experimental ones, the difference not exceeding 0.1%. Notably, carrying out the minimization without the VdW empirical contribution, resulted in rather erratic values of the energy about an average energy–volume curve, and in final values of a, b, and c off by about 1 to 3% with respect to the experiment. The final conformation of the **LeX** molecule, extracted from the fully relaxed unit cell, is shown in Figure 1, with a definition of the nomenclature for bonds and angles.

By direct inspection of the molecular configurations, and by looking at the torsional angles Ψ and Φ, we saw that absence of the VdW term led to increased intramolecular fluctuations between the **Gal** and **Fuc** rings of each molecule. The much

Table 1. Comparison of Structural Quantities for the Two LeX Molecules in the $P2_1$ Unit Cell[a]

	CPMD T = 0 K		EXP		CPMD T = 300 K	
	LeX1	LeX2	LeX1	LeX2	LeX1	LeX2
			bond lengths			
C1 Gln−O5 Gln	1.484	1.462	1.433	1.419	1.489 ± 0.046	1.464 ± 0.029
C5 Gln−O5 Gln	1.476	1.479	1.418	1.434	1.483 ± 0.047	1.482 ± 0.025
C4 Gln−O1 Gal	1.474	1.475	1.419	1.440	1.483 ± 0.040	1.487 ± 0.030
O1 Gal−C1 Gal	1.432	1.425	1.421	1.418	1.442 ± 0.040	1.437 ± 0.039
C1 Gal−O5 Gal	1.464	1.463	1.437	1.417	1.476 ± 0.042	1.465 ± 0.040
C5 Gal−O5 Gal	1.490	1.490	1.432	1.480	1.494 ± 0.042	1.496 ± 0.050
C3 Gln−O1 Fuc	1.468	1.473	1.428	1.439	1.476 ± 0.034	1.474 ± 0.031
O1 Fuc−C1 Fuc	1.453	1.464	1.415	1.429	1.456 ± 0.036	1.471 ± 0.043
C1 Fuc−O5 Fuc	1.460	1.440	1.412	1.395	1.469 ± 0.043	1.443 ± 0.035
C5 Fuc−O5 Fuc	1.521	1.511	1.454	1.461	1.520 ± 0.036	1.516 ± 0.051
			bond angles			
C1−O5−C5 Gln	112.3	112.2	113.9	112.2	112.4 ± 3.5	112.8 ± 4.0
C4−O1−C1 Gal	115.8	115.6	117.0	117.7	115.4 ± 3.9	115.6 ± 3.1
C1−O5−C5 Gal	111.8	113.6	111.8	111.5	112.1 ± 3.9	113.9 ± 3.1
C3−O1−C1 Fuc	114.9	114.1	115.8	115.1	116.1 ± 4.1	114.5 ± 3.8
C1−O5−C5 Fuc	111.9	111.8	113.9	114.0	112.4 ± 3.4	111.9 ± 2.7
			torsion angles			
Ψ (Gal)	−77.1	−68.2	−80.0	−70.4	−78.1 ± 8.4	−69.9 ± 8.3
Φ (Gal)	−105.1	−105.6	−104.6	−107.8	−104.0 ± 6.5	−106.1 ± 6.4
Ψ (Fuc)	−72.4	−75.1	−72.3	−76.7	−72.3 ± 7.3	−75.1 ± 7.2
Φ (Gal)	140.1	138.1	138.7	139.0	138.9 ± 6.4	139.1 ± 6.3

[a] Experimental data (EXP, columns 4 and 5), and data calculated by Car−Parrinello molecular dynamics: after T = 0 K relaxation (CPMD T = 0 K, columns 2 and 3), and thermal averages at T = 300 K (columns 6 and 7). The experimental error in ref 20 is quoted at ±1%. In column 1, the definitions of the various bond lengths (Å), bond angles (deg), and torsional angles (deg) are given. For the corresponding nomenclature, see Figure 1. Torsional angles at the glycosidic bonds are defined as Φ(Gal) = O5−C1−O1−C4, Φ(Fuc) = O5−C1−O1−C3, Ψ(Gal) = C1−O1−C4−C5, Ψ(Fuc) = C1−O1−C3−C4 (all starting at Gal or Fuc, and ending at Gln).

smoother energy curve, as well as the very good agreement with the experimental volume, show that the VdW energy is correctly estimated with the Grimme parametrization, and that it plays a major role in the stabilization of the intramolecular degrees of freedom.

MOLECULAR DYNAMICS RESULTS

A detailed comparison of theoretical and experimental bond lengths, bond angles, and torsional angles is given in Table 1. It can be seen that the calculated bond lengths are systematically overestimated, by about 2−3%. Such error, likely originating from the known geometrical overcorrections of GGA,[27] is however not very significant, since it remains within the standard deviation of the bond−length values averaged in the MD simulations at the experimental temperature T = 300 K (see Table 1, last columns).

The bond angles centered at an oxygen atom, in each of the three sugar rings and in the two glycosidic bonds, are generally within ±2° and, in most cases, within ±1° from the experimental ones. The same holds for the torsional angles Ψ and Φ, while the finite−temperature MD averaged values agree with the experimental values at better than 1%. It should also be noted that the conformations of the two **LeX** molecules in the asymmetric unit are not identical: albeit small, the differences in the values of bond lengths, angles and torsions are beyond the statistical error, both in the experiment and in our simulation results (Table 1). Overall, the comparison with the experimental data is very satisfactory, proving the reliability of our approximations and computational methods.

Hydrogen Bonding Network. Having established the accuracy of the computational protocol, we could then address one first important question. Since hydrogen atoms are not visible in x-ray experiments, the position of H atoms and the actual hydrogen bonds network are deduced a posteriori, by an educated guess on the crystallographic structure (see ref 28 and references therein). Dynamical simulations, then, provide a reliable tool to assess and verify the stability of the hydrogen network, and allow for a direct inspection of the whole atomic positions. We adopted the definition for a hydrogen bond proposed by Mezei and Beveridge,[29] namely the distance between the donor (D) and acceptor (A) being less than 3.3 Å, the angles H−D−A and LP−A−D, involving the hydrogen atom (H), and the lone pair (LP) of the acceptor, being both less than 70.53°. For a better comparison with experiments, H−bonds were also evaluated with a more restrictive criterion, at angles smaller than 45°. Note that, for the doubly−bonded **C═O** oxygen in **Gln**, the angle involving the LP cannot be estimated.

With the 70.53° prescription, we could identify 40 H−bonds in each asymmetric unit (compared to 36 in the experimental work[19]), that is, 80 H−bonds in the unit cell, or 20 for each **LeX** molecule. In Tables 2 and 3 we list the structural features of each

dx.doi.org/10.1021/jp2055816 |J. Phys. Chem. B 2011, 115, 12599–12606

Table 2. Hydrogen Bonds Involving Water Molecules in the Asymmetric Unit[a]

r_{D-A}	θ(H)	θ(LP)	donor	acceptor
2.546	8.26	20.37	O2 Gal2′	W5
2.860	6.69	30.84	W5	W4
3.116	22.70	60.02	W3	W4 (§)
2.825	3.32	13.17	W4	O5 Fuc1
2.845	9.76	14.31	W4	W7′ ($)
2.894	11.72	20.36	W7′	O2 Gal2′
3.260	2.18	5.93	W5	W1′ (+)
2.906	12.14	14.43	W1′	O6 Gal1
2.941	13.66	46.44	O4 Fuc2′	W9′
2.880	5.60	15.22	W9′	O6 Gln2′
3.016	5.72	1.22	W9′	W7′
2.748	3.78	19.33	W7′	W6′ ($)
2.948	13.70	50.79	W8′	W6′ (#)
2.858	6.91	17.70	W6′	W1′ (+)
2.484	3.61	12.97	W1′	O6 Gln1
2.507	8.78	38.03	W6′	O2 Gal1
2.838	7.39	9.15	N2 Gln2	W2
2.961	3.63	39.60	W2	O2 Fuc1
2.936	5.79	19.90	W2	W8 (#)
2.820	5.02	7.19	W8	O3 Gal1′
2.986	7.28	4.89	N2 Gln1	W3 (§)
2.742	4.51	12.11	W3	O2 Fuc2

[a] Classified according to: donor−acceptor distance, r_{D-A} (column 1, Å); angle comprising the lone-pair, donor and acceptor oxygens (column 2, deg); angle comprising the hydrogen atom, donor, and acceptor oxygens (column 3, deg). In columns 4 and 5, the donor and acceptor atoms (O or N) in water and **LeX** molecules are listed. The index (1,2) refers to the two **LeX** in the asymmetric unit. The prime (′) symbol indicates a molecule lying in the repeated half of the unit cell. Each block of rows makes up a chain. Chains can branch, and reconnect at the symbols ($,+,§,#).

H−bond in one-half of the unit cell (asymmetric unit), the other half being equivalent.

Of these 40 bonds, 22 per asymmetric unit (see Tab. 2, 11 in each **LeX**) involve water molecules, and the remaining 18 (see Table 3, 9 in each **LeX**) directly connect **LeX** molecules, either across the facing pairs of **Fuc**−**Gal** rings (see also Figure 1, we call this 'head−to−head') or sideways. By restricting the angles to <45°, the number of H−bonds is reduced to 35 per unit, closer to the experimental work, however with a slightly different distribution, since in the X-ray structure 30 out of 36 H−bonds were supposed to imply at least one water molecule, while in our case the ratio is 20 out of 35. (Unfortunately, the criterion adopted to identify H-bonds in ref 19 was not specified.) Each of the two independent **LeX** in the asymmetric unit is involved in 24 or 25 H-bonds, respectively (reduced to 20 and 21 for angles <45°, compared to 21 in the experiment[19]).

H−bonding appears to give its most important contribution to the intermolecular energy since, with only one exception (the H−bond linking the **O6**−**Gal** and **O4**−**Fuc**, within each **LeX**), all such H−bonds connect different molecules. Although the bond−distance criterion of 3.3 Å could appear somewhat too broad, it is readily noted that the D−A average distances rarely exceed the value of 3.0 Å (three extreme cases, also corresponding to rather large values of the angles, are represented by the bonds in rows 12−14 of Table 3, having D−A distances of 3.22−3.26 Å).

Table 3. Hydrogen Bonds Directly Linking O Atoms in Adjacent LeX Molecules, in an Asymmetric Unit[a]

r_{D-A}	θ(H)	θ(LP)	donor	acceptor
2.345	15.53	23.90	O4 Fuc1	O3 Gal2′
2.724	16.45	44.38	O6 Gal1	O4 Fuc1
2.667	4.59	30.87	O6 Gal2	O4 Fuc2
3.121	12.96	21.71	O2 Gal1	O2 Gal2′
2.591	7.26	18.27	O4 Gal1	O6 Gal2′
3.200	12.87	42.55	O4 Gal2	O4 Fuc1
2.948	21.30	17.35	O2 Fuc2	O5 Gln1
3.005	12.46	10.67	O3 Gal2′	O6 Gln1
3.071	19.53	29.92	O3 Gal1	O6 Gln2′
2.815	12.82	15.12	O2 Fuc1	O1 Gln2
2.836	42.29	22.22	O2 Fuc1	O5 Gln2
3.226	55.64	47.57	O3 Gal1	O6 Gal2′
3.259	65.45	57.05	O4 Gal1	O4 Gal2′
3.238	52.68	17.72	O3 Gal2′	O3 Fuc1
2.590	3.39		O6 Gln1	O7 Gln2
2.632	7.51		O6 Gln2	O7 Gln1
2.543	13.43		O3 Fuc1	O7 Gln2
2.845	5.34		O3 Fuc2	O7 Gln1

[a] Columns and labels as in Table 2. The first block comprises head-to-head bonds, between facing molecules; the second block comprises sideways bonds, between parallel molecules; the third block comprises H-bonds involving the doubly bonded oxygen in **Gln** (on the **NAc** terminal).

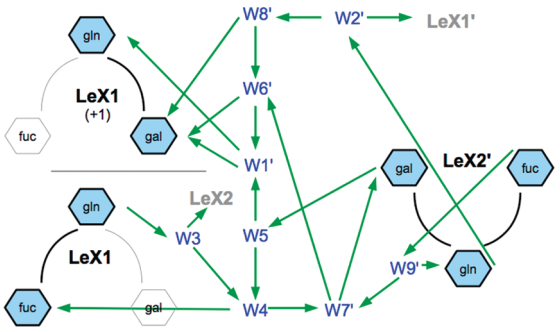


Figure 2. Schematic of the water-mediated hydrogen bond network. Green arrows point from donor to acceptor. The prime (′) indicates a molecule in the repeat half of the unit cell. The "+1" indicates a molecule in an adjacent unit cell.

Even beyond the numerical details, the data from Table 2 reveal that the H−bond network involving water molecules is complex and highly branched. A schematic representation of the water-H-bond network is shown in Figure 2. We find the same, unique example of a cooperative ring also identified in the experimental structure (**O2 Gal2′**−**W5**−**W4**−**W7′**−**O2 Gal2′**). Infinite chains are also found, i.e., connected series of H−bonds starting at one atom and reaching up to the periodically repeated image of the same atom in a neighboring unit cell (for example, the quite long **O3 Gal1**−**O6 Glu2′**−**W9′**−**W7′**−**W6′**−**W8′**−**O3 Gal1**, or the shorter **O2 Gal2′**−**W7′**−**W6′**−**O2 Gal1**−**O2 Gal2′**).

All such features are well-known in carbohydrate chemistry, under the name of *cooperative* H-bonds, and are supposed to affect molecular recognition in complex ways, for example providing the (human) taste of different sugars.[30,31]

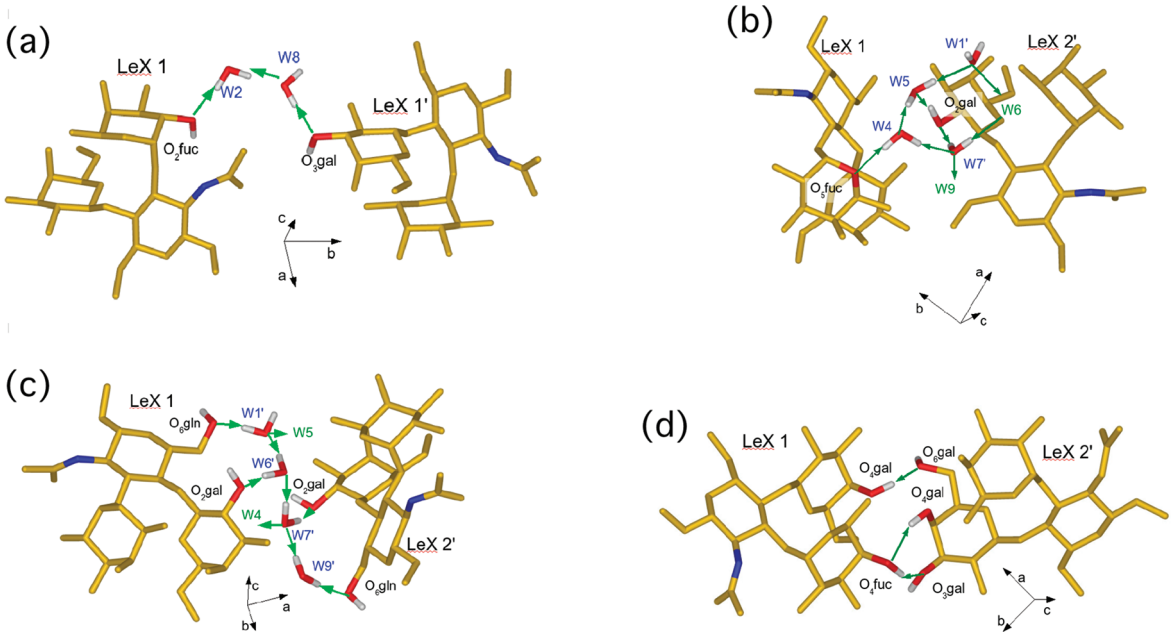


Figure 3. Examples of H-bonds in the **LeX** crystal unit cell. (a) Short water chain between the two **LeX1** in the repeated asymmetric units. (b) Branched water chain between the **LeX1** and the **LeX2** belonging to the two asymmetric units. (c) Long water chain between the same two **LeX** of (b). (d) Direct (nonwater-mediated) bonds between facing **LeX** in the two asymmetric units. The cell orientation is shown by the (*a*, *b*, *c*) reference frame.

As a general conclusion, we may observe that the intermolecular stability provided by H−bonds is based on a complex structuring of the water molecules surrounding the **LeX**, even in the relatively low hydration regime provided by the crystalline phase. The implications of such observation for the hydrated, in vivo configurations, as well as the role of the Ca^{2+} ions in water structuring and molecular adhesion, are a complex issue, open to further investigation.

Figure 3 shows four examples of H−bond networks, with a−c being examples of water-mediated chains, and (d) representing the three direct H−bonds between head-to-head **LeX** molecules (compare to the list given in Table 3, upper block). Water molecules in the crystal are arranged on the "hydrophilic" sides of the **LeX**, namely, in the regions lying mostly away from the **NAc** group.

Further **LeX**−**LeX** direct H−bonds (i.e., not involving water) exist also on the "hydrophobic" side nearing the acetyl $-\mathbf{COCH_3}$: these are the bonds listed in Tab. 3 (second and third block), also including the oxygen involved in the **C═O** double bond at the terminal of the **NAc** group (O7 in Figure 1). The overall picture is that each **LeX** is firmly embedded in a tight H−bond cage, comprising both water−mediated and direct bonds on the sides, plus a group of head−to−head direct bonds on the **Fuc+Gal** terminal. The latter are of special biological importance, since they are candidates to provide the main recognition and adhesion features. In fact, **LeX** determinants in glycans are always attached from the glucose terminal to the substrate, therefore the **Fuc+Gal** terminal head is the one mostly exposed to the biological interaction.

Upon MD simulation at finite temperature, it is seen that water molecules in the crystal display different behaviors, according to whether H−bonded to oxygen or nitrogen atoms. Only a few **W**−**LeX** bonds (2 out of 13 in Tab. 2) have a root mean squared fluctuation of <0.15 Å about the average at T = 300 K. Instead, most water molecules seem to straddle considerably about their

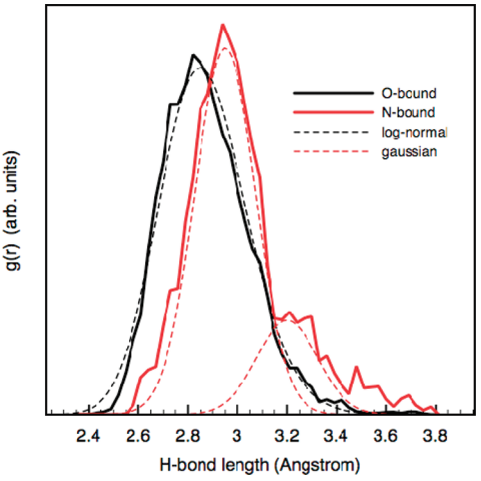


Figure 4. Pair-correlation functions for the H-bonds of O-bonded and N-bonded water molecules. Averages were taken over the last 40 ps of simulation, and over the four molecules in the irreducible unit cell. *r* is in Ångstrom, $g(r)$ in arbitrary units.

stable (T = 0 K) position. Figure 4 shows the partial pair−correlation functions, $g(r)$, of the O-bound and N-bound waters, averaged over 40 ps of simulation and over the four water molecules in the asymmetric unit, the distance r being measured relative to the closest donor or acceptor atoms in the **LeX**. The $g(r)$ of O-bound waters is nicely log-normal, with a fwhm spread of about 0.6 Å, indicative of a substantial mobility of the bound water about the equilibrium position. By looking at the individual $g(r)$ of the waters **W1**, **W2**, and **W3**, we notice that they are almost identical and simply shifted (by about 0.1−0.3 Å), meaning that the O atoms on each **LeX** are relatively immobile with respect to the motion of water molecules. For **W9**, in particular, the $g(r)$ corresponding to the H-bond with **O6**−**Gln2**

dx.doi.org/10.1021/jp2055816 |*J. Phys. Chem. B* 2011, 115, 12599–12606

and **O6−Gal2** are quite different from each other, pointing to a more complex relative dynamics of the two O atoms. On the other hand, the $g(r)$ of N- bound waters in Figure 4 is quite different, shifted to a larger average value of the bond length, and fitted by at least two Gaussian peaks. Both the log-normal and multi-Gaussian statistical distributions suggest that H-bond lengths (see Tables 2 and 3) indeed result from averages over more than one intermediate, metastable configuration. Notably, H-bonds are also metastable in time. Out of 40 H-bonds per asymmetric unit (22 involving water, 18 direct), we found at least 3 water−mediated and 5 direct bonds having a mean lifetime of less than 3/4 of the simulation time (40 ps), which may be considered a statistically significant measure of metastability.

Crystal Binding Energy: Hydrogen Bonding versus Dispersion Forces. The considerations above regarding the arrangement of the H−bond network in the crystalline structure of **LeX** led us to investigate a second relevant question: namely, the energetics of intermolecular vs intramolecular stabilization or, in other words, the competition between van der Waals and H−bond energy (although any such definition has some degree of ambiguity, since the H−bond energy always includes some VdW contribution). The VdW contribution to the stabilization energy was estimated as the difference between the total energies calculated with and without the VdW contribution in the crystal structure at the same minimum−energy configuration.

The binding energy in the unit cell was defined as $E_B = (E_{TOT} - 4E_{LeX} - 18E_{H_2O})/4 = -76.6$ kcal/mol, with E_{LeX} and E_{H_2O} the total energy, respectively, of one isolated **LeX** and **H_2O** molecule in vacuo, and E_{TOT} the total energy of the crystal, all computed at zero temperature and with the same cell size (presuming a negligible difference between the zero-point energy of free water with respect to crystallized water). By recalculating the same quantity, but without the VdW contribution, we obtained $E_{B^*} = -37.4$ kcal/mol. The VdW contribution amounts therefore to $E_{VdW} = -39.2$ kcal/mol, that is, slightly more than half of the total binding energy.

The remaining fraction of the energy, $E_{B^*} = E_B - E_{VdW}$, can be attributed to the sum of all the H-bonds plus the long-range electrostatic interactions, $E_{B^*} = E_H + E_{elec}$. The latter component of the energy was directly obtained from the DFT calculation, as the integral of the average electrostatic potential over the charge distribution (Poisson's equation), and in this case it results equal to $E_{elec} = 47.9$ kcal/mol (positive, i.e., building the crystal from the separate molecules *costs* electrostatic energy, to be balanced by H−bond and VdW energy). Therefore, the H-bond total energy is $E_H = -85.3$ kcal/mol, resulting in an average of about 4.3 kcal for each H-bond.

Since, as we already noted, all the H-bonds but one per **LeX** molecule (i.e., 76 out of 80 per unit cell) are between different molecules, it can be deduced that the *intermolecular* stabilization energy originates mostly from the very dense H-bond network. On the other hand, we already observed that both the unit cell size and shape, as well as the detailed structure of the individual **LeX** molecules, are quite off with respect to the experimental data, when minimized without the VdW interaction. Therefore, we can infer that the VdW energy, although only about half the H-bonding energy, mainly provides the *intramolecular* stabilization, with a secondary effect also on the intermolecular bonding. Clearly, such a classification of the various contributions to the energy must not be taken too strictly, since the H-bonds always contain some VdW energy,[32] and both are of fundamental electrostatic origin. The above estimates should serve just as a guidance in the forthcoming discussion of the adhesion energy.

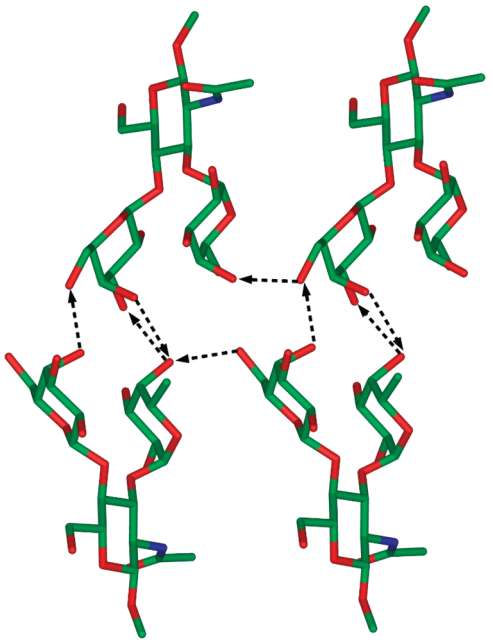


Figure 5. Representation of the head-to-head interacting **LeX** configuration, obtained by a particular cutting of the four **LeX** in the unit cell, from four neighboring cells. Hydrogen bonds are shown by dashed arrows (donor to acceptor). Color coding as in Figure 1.

Adhesion Energy. From the above considerations a working structure, to be used as the initial step for the study of the candidate in vivo configurations of the **LeX−LeX** interaction, can be extracted. Notably, such a structure could as well represent the basic building block of a nanoparticle assembly strategy, once the **LeX** are mounted on, for example, the terminal of an alkyl chain via the anomeric oxygen of the glucose ring, the alkyl chains forming in turn a dense monolayer on the nanoparticle surface. Therefore, this information is extremely relevant, but practically inaccessible to experimental analysis, at least at the level of detail which can be provided by first-principles simulations.

Figure 5, extracted by cutting four facing **LeX** molecules from a multiple repeat of neighboring unit cells, shows a possible head-to-head interacting configuration. Water molecules are removed for clarity. Each pair of facing **LeX** in this projection is held together by three direct H-bonds (arrows in the Figure), plus one H-bond linking to the side neighbor. Note that each molecule is also linked to its other side neighbors by twelve other H-bonds (not shown, they connect **LeX** in periodic repeats of the unit cell). For the purpose of nanoparticle assembly, however, only the three direct, head-to-head bonds per molecule give an effective contribution to adhesion (of course, with an additional contribution from background VdW forces, not quantified in this estimate), whereas the side H-bonds contribute to intermolecular rigidity and, likely, ensure the best accessibility of molecular surfaces.

On the basis of the above estimates, one could expect a value of adhesion energy of the order of about 13 kcal/mol pair. This value is based simply on the contribution of the three direct H-bonds, disregarding other contributions, such as VdW, and electrostatic. (Moreover, in the free-standing configurations, other forces could play a role, such as desolvation of the lipophilic

side of **LeX**.) Anyway, for a guess surface density of the alkyl chains, carrying one **LeX** each, of ~0.01 $Å^{-2}$,[33] the above estimated value turns into 5.6×10^{-3} eV/$Å^2$, or about 0.09 J/m^2.

This estimate could be compared to recent values of surface adhesion energies of interacting **LeX**−**LeX** pairs, measured for **LeX** mounted on different long-chain ceramide molecules, and coated onto large spherical vesicles of several micrometers diameter, to mimic adjoining cell membranes.[13] The quite peculiar arrangement of that experiment notwithstanding, adhesion energies of 47.5 μJ/m^2, or about 3×10^{-6} eV/$Å^2$, were deduced for a surface density of 10^{-4} $Å^{-2}$. This is about a factor of 20 smaller than our estimate above, when transformed into equivalent surface density. Notably, however, a more than 5-fold variability of the experimental results was observed, as a function of the type of supporting molecules, showing an important effect of the steric accessibility of the **LeX** during the interaction. Moreover, it should also be noted that, for the very low values of the surface density in the experiment, the **LeX** are not enough tightly packed to be stabilized by lateral H−bonds. The intrinsic dynamics of the **LeX**−**LeX** interaction should therefore become relevant as well, leading to a reduced interaction probability and incomplete adhesion, compared to the ideal upper limit of our theoretical analysis. Finally, it should be recalled that our estimate of the H−bond energy includes a majority of cooperative bonds, meaning that the energy of individual bonds could be somewhat smaller. Therefore, although affected by several differences and uncertainties, our (over)estimate is broadly consistent with the experimental results.

DISCUSSION AND CONCLUSIONS

In summary, we carried out a first-principles computational study of the **LeX** trisaccharide in condensed phase, by means of Car−Parrinello molecular dynamics. The resulting equilibrium structure of the **LeX** crystal at T = 300 K compare well with the experimental X-ray data. We discussed the detailed structure of the H-bonding network, as well as H-bonding and van der Waals energies, and identified candidate structures of interacting **LeX**−**LeX** pairs, possibly relevant for cell adhesion and nanoparticle self−assembly. It was observed that the main contribution of the VdW energy is in the stabilization of the individual molecules in the unit cell, while the relative positioning and orientation of the molecules with respect to one another is mostly assured by the tight H-bond network, with a high number (per molecule) of water-mediated and direct H-bonds. However, as already mentioned, such a distinction is just a guideline for the interpretation of the results, since also H-bonds have a VdW component,[32] and both interactions are of electrostatic origin.

Such condensed phase structures correspond to conditions of low hydration and absence of ions, such as Ca^{2+}, known to play a relevant role in the self-assembly and recognition processes. They can be considered as a reasonable starting point to study the dynamics of self-assembly, likely more interesting for the surface coating of nanoparticles, in which case the surface density should be high enough for the tight H-bond network to maintain the right molecular orientation and therefore ensure the most relevant interaction. We observed that the head-to-head, direct (i.e., non water-mediated) H-bonds provide in this case the largest fraction of the adhesion force, with a minor contribution arising from VdW forces, while the side H-bonds, both direct and water mediated, provide directional stability and accessibility of the more hydrophilic **LeX** surface. On the other hand, the same configurations are probably less relevant when compared to in vivo conformations, in which the driving force for the formation of the **LeX**−**LeX** complex is probably the desolvation of the complementary lipophilic surfaces of the **LeX** in water, the stabilization being likely provided by a complex interrelation of Ca^{2+} ions and H-bond network.

The dynamics of free **LeX** pairs in water is now the next important step to be assessed to fully characterize carbohydrate interactions, starting from the results of the present study. Moreover, at this stage the role of the Ca^{2+} ions in the adhesion switching process remains to be clarified. Clearly, the ion dynamics in such a complex system will require a separate and more detailed study.

AUTHOR INFORMATION

Corresponding Author
*E-mail: fabrizio.cleri@univ-lille1.fr.

ACKNOWLEDGMENT

R.Z. thanks a postdoc grant from IEMN Cnrs, Axe-1 "Nanostructures", and a supporting grant from the CPER Region Nord-Pas de Calais. Computer resources were provided by IDRIS Orsay and CINES Montpellier Supercomputing centers, under Grant 2010095007 to F.C. The ENEA Centro Ricerche Frascati (Italy) is also acknowledged for a generous computer time allocation. Y. Guerardel (Glycobiology Unit, University of Lille I) is thanked for attracting our interest to this problem and for many inspiring discussions. Discussions with R. Blossey (Interdisciplinary Research Institute, Lille) are also acknowledged.

REFERENCES

(1) Gabius, H.; Gabius, S. *Glycoscience*; Chapman and Hall: Weinheim, Germany, 1997.
(2) Varki, A. *Glycobiology* **1993**, *3*, 97–130.
(3) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683–720.
(4) Hakomori, S.-I. *Cancer Res.* **1996**, *56*, 5309–5314.
(5) Bourne, Y.; van Tilbeurgh, H.; Cambillau, C. *Curr. Opin. Struct. Biol.* **1992**, *3*, 681–6.
(6) Hakomori, S.-I. *Pure Appl. Chem.* **1991**, *63*, 473–82.
(7) de la Fuente, J. M.; Penades, S. *Glycoconj. J.* **2004**, *21*, 149–163.
(8) Sharon, N.; Lis, H. *Science* **1989**, *246*, 227–234.
(9) Frank, D. S.; Yen, W. Y. *Discuss. Faraday Soc.* **1957**, *24*, 133.
(10) Lopez de la Paz, M.; Ellis, G.; Perez, M.; Perkins, J.; Jimenez-Barbero, J.; Vicent, C. *Eur. J. Org. Chem.* **2002**, *21*, 840.
(11) Crowe, J. H.; Crowe, L. M.; Carpenter, J. M.; Aurell Wistrom, C. *Biochem. J.* **1987**, *242*, 1.
(12) Edwards, N. J.; Monteiro, M.; Faller, G.; Walsh, E. J.; Moran, A. P.; Roberts, I. S.; High, N. J. *Mol. Microbiol.* **2000**, *35*, 1530–1539.
(13) Gourier, C.; Pincet, F.; Perez, E.; Zhang, Y.; Zhu, Z.; Mallet, J.-M.; Sinay, P. *Angew. Chem., Int. Ed.* **2005**, *44*, 1683–1687.
(14) (a) Eggens, J.; Fenderson, B. A.; Toyokuni, T.; Dean, B.; Stroud, M. R.; Hakomori, S.-I. *J. Biol. Chem.* **1989**, *264*, 9476–9484. (b) Kojima, N.; Fenderson, B. A.; Stroud, M. R.; Goldberg, R. I.; Habermans, R.; Toyokuni, T.; Hakomori, S.-I. *Glycoconj. J.* **1994**, *11*, 238–248.
(15) Henry, B.; Desvaux, H.; Pristchepa, M.; Berthault, P.; Zhang, Y.-M.; Mallet, J.-M.; Esnault, J.; Sinay, P. *Carbohydr. Res.* **1999**, *315*, 48–62.
(16) Geyer, A.; Gege, C.; Schmidt, R. R. *Angew. Chem., Int. Ed.* **2000**, *39*, 3245–3249.

(17) Tromas, C.; Rojo, J.; de la Fuente, J. M.; Barrientos, A. G.; García, R.; Penadés, S. *Angew. Chem., Int. Ed.* **2001**, *40*, 1554–1557.
(18) Hernaiz, M. J.; de la Fuente, J. M.; Barrientos, A. G.; Penadés, S. *Angew. Chem., Int. Ed.* **2002**, *41*, 1554–1557.
(19) Pérez, S.; Mouhous-Riou, N.; Nifant'ev, N. Z.; Tsvekov, Y. E.; Bachet, B.; Imberty, A. *Glycobiology* **1996**, *6*, 537–542.
(20) Yvelin, F.; Zhang, Y.—M.; Mallet, J. M.; Robert, F.; Jeannin, Y.; Sinay, P. *Carbohydr. Lett.* **1996**, *1*, 475–480.
(21) (a) Gourmala, C.; Luo, Y.; Barbault, F.; Zhang, Y.; Ghalem, S.; Maurel, F.; Tao Fan, B. *THEOCHEM* **2007**, *821*, 22–39. (b) Luo, Y.; Barbault, F.; Gourmala, C.; Zhang, Y.; Maurel, F.; Hu, Y.; Tao Fan, B. *J. Mol. Model.* **2008**, *14*, 901–910. (c) Molteni, C.; Parrinello, M. *Chem. Phys. Lett.* **1997**, *275*, 409–413.
(22) CPMD, Copyright IBM Corp. (1990—2009) and MPI für Festkörperforschung Stuttgart (1997—2001). www.cpmd.org.
(23) (a) Becke, A. D. *Phys. Rev. A* **1998**, *38*, 3098. (b) Lee, C; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(24) Troullier, N.; Martins, J.-L. *Phys. Rev. B* **1991**, *43*, 1993.
(25) Prendergast, D.; Grossman, J. C.; Galli, G. J. *Chem. Phys.* **2005**, *123*, 014501.
(26) Grimme, S. J. *Comput. Chem.* **2006**, *27*, 1787.
(27) Johnson, B. G.; Gill, P. M. W.; Pople, J. A. *J. Chem. Phys.* **1993**, *98*, 5612–5626.
(28) Boero, M.; Ikeda, T.; Ito, E.; Terakura, K. *J. Am. Chem. Soc.* **2006**, *128*, 16798–16807.
(29) Mezei, S.; Beveridge, D. L. *J. Chem. Phys.* **1981**, *74*, 622–628.
(30) Jeffrey, G. A. *Food Chem.* **1996**, *56*, 241–246.
(31) Dashnau, J. L.; Sharp, K. A.; Vanderkooi, J. M. *J. Phys. Chem.* **2005**, *109*, 24152–24159.
(32) Lin, I. C.; Seitonen, A. P.; Coutinho-Neto, M. D.; Röthlisberger, U. *J. Phys. Chem. B* **2009**, *113*, 1127–1131.
(33) Cleri, F.; Letardi, S. *Appl. Phys. A: Mater. Sci. Process.* **2007**, *86*, 293–300.