H- and bridged bonds analysis for mucin runs

January 4, 2015

Contents [hide]

GalNAc H-Bonding Analysis

To investigate the structural stabilization of the n-acetyl-galactosylated moieties, we considered the (previously suggested?) possibility of hydrogen bonding between the N-acetyl H in the GalNAc and the O in the nearby backbone (Figure YY). Although there was little evidence for this sort of hydrogen bond (see region labeled 1 and enclosed by black lines), there was a substantial population of conformations nearby (region labeled 2).

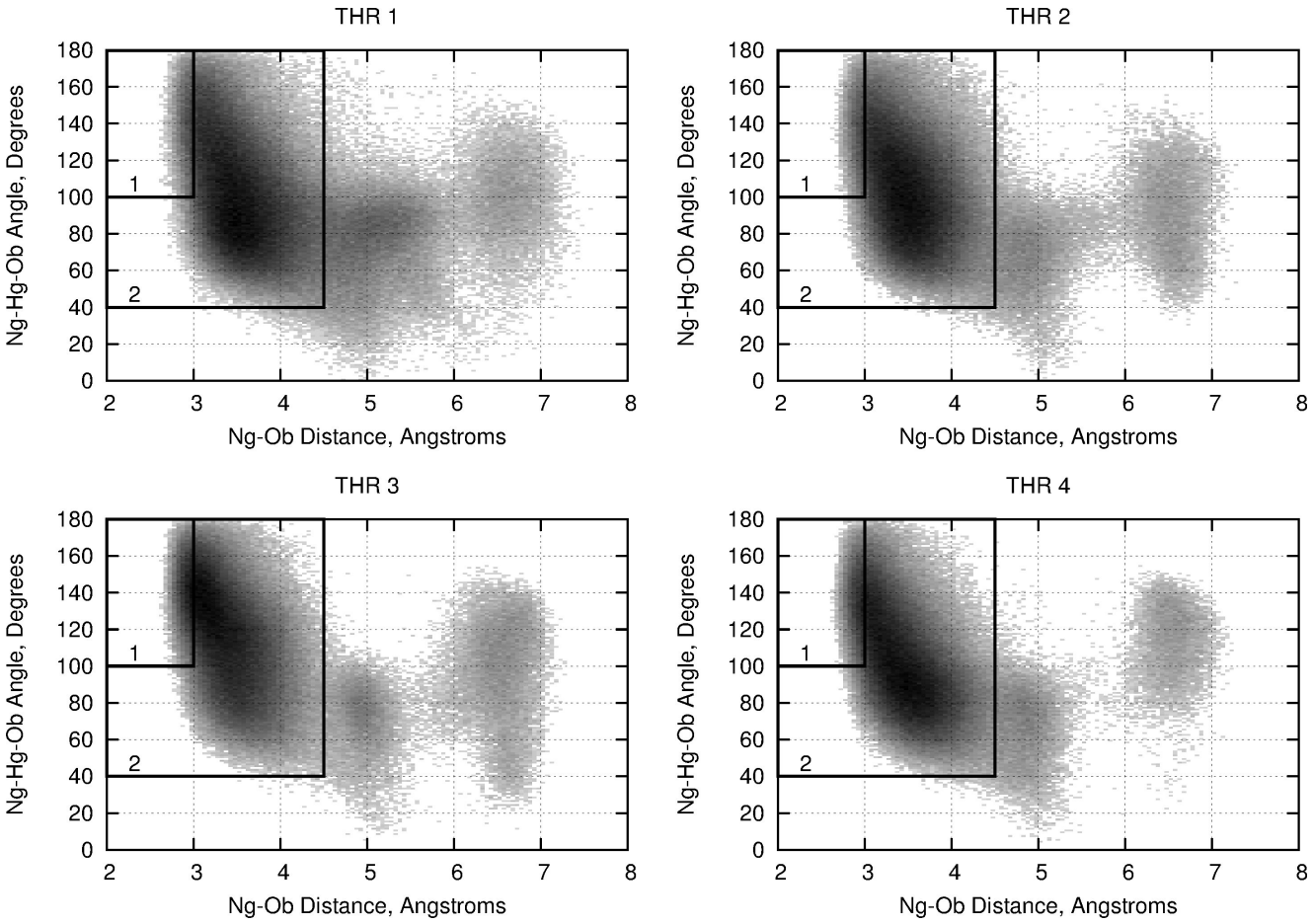


Figure YY: Hydrogen bonding analysis for interactions between the HN ("Hg") of GalNAc's NAc with the O in the local backbone ("Ob"). The NAc nitrogen is represented as "Ng". Black lines enclose region 1, where hydrogen bonding is defined as occurring, and region 2, where the bulk of the population exists.

Analysis of Region 2

To investigate the structural features in the region, coordinates corresponding to region 2 were saved into separate trajectories, one for each glycosylation site. While visualizing the atoms in proximity to Hg and Ob (see Figure YY caption), a possible reason for the structural stability was noted. In region 2, representing about 90% of the simulation frames, the GalNAc and its associated threonine are oriented in a manner similar to that shown in Figure X. That is, the Hg and Ob form a pseudo-tetrahedral arrangement with the backbone H ("Hb") and the glycosidic O ("Og"). Presumably, the two oxygens are stabilized by the presence of the two hydrogens and vice-versa. The apparent bias in the force field toward less-negative values of Psi (see Figure Something) does not invalidate these structures. Indeed, if anything, the force field would then be minimizing the importance of these structures because the less negative values of Phi move the Hb away from the interaction.

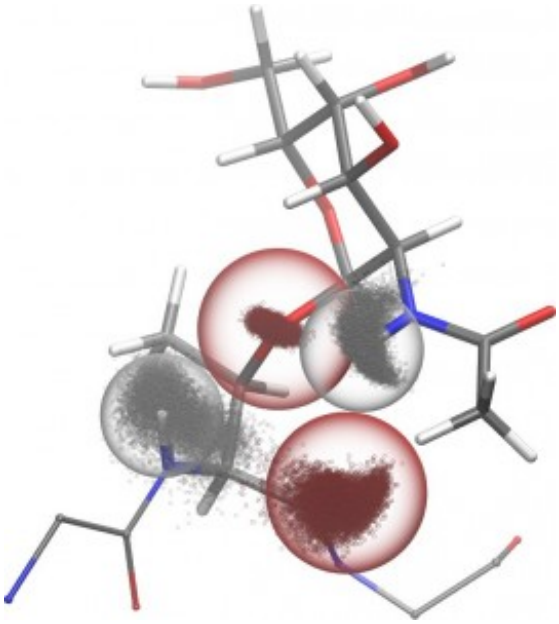


Figure X: Illustration of a possible stabilizing structure for the N-acetyl galactosylated moiety. For this image, the heavy atoms for the GalNAc and THR residues at site 3, from only those frames from region 2 (see Figure YY), were aligned using VMD. The RMSDs for those atoms were then taken with respect to the average aligned structure. From the RMSD results, the frame closest to the average (with the lowest RMSD, frame index 15536) was selected and is illustrated with thick sticks. The backbone atoms in the flanking residues at that frame are shown as thin sticks and small spheres. The large transparent spheres represent approximate van der Waals radii of Hg, Hb, Og and Ob at the reference frame. The small, transparent spheres represent the range of positions of the Hg, Hb, Og and Ob in region 2. Specifically, they represent every 20th frame out of the total of 278213 frames, after alignment as described. Since they are drawn as transparent spheres, regions of darker color represent regions that are populated more frequently.

Possible Bridging Water

Previously, the stabilization of GalNAc to threonine has been studied in a similar manner, but only for a single GalNAc attached to a THR. The results of those studies suggested that the relative orientations of the GalNAc and THR were stabilized by a water molecule whose oxygen formed a bridging hydrogen bond between the equivalent Hg and Hb. To conserve resources, the coordinates of waters were not saved in our main simulations. However, water coordinates were saved during the equilibration phase. The systems were amply equilibrated, so we searched for evidence of a bridging water in the well-equilbrated portions of those trajectories, a coordinate subset of 28,832 frames (compare to the 355,642 frames in the main simulations).

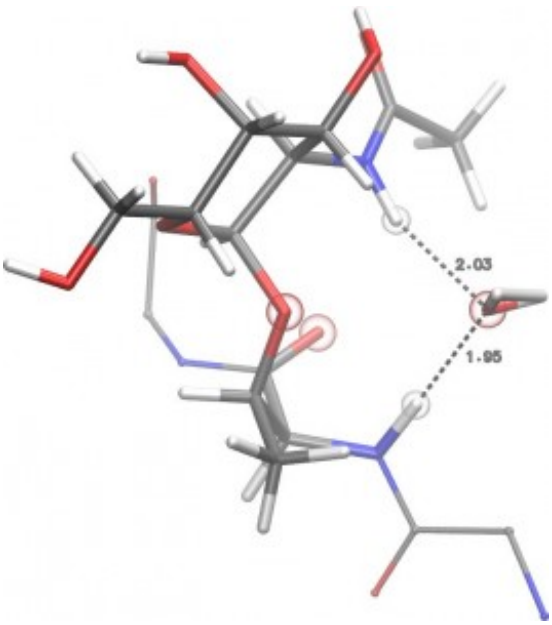
Because of the relatively small number of simulation frames, the frames were almost entirely found in region 2. We constructed a program to scan these frames for water molecules forming a bridge with Hb and Hg. We did find evidence of them (see Table ZZ and Figure WW). However, with the 4-way interaction being present essentially 100% of the time in these simulations, it is difficult to say whether the bridging waters act to stabilize the interaction or if they are merely form bridges occasionally due to the proximity of two appropriate hydrogens.

Table ZZ: The percentage of frames from the equilbrated portions of the equilibration phase in which bridging waters were found, noted per glycosylation site.

Site #	1	2	3	4
% Bridging	13	19	21	30

Edit

Figure WW: Example frame showing a bridging water. The dashed lines connect the oxygen of the bridging water with the two hydrogens involved in the bridged bonds. The transparent spheres highlight the atoms involved in the proposed 4-way interaction plus the oxygen of the bridging water.



Examination of structural differences between frames containing a bridging bond and those that do not are also inconclusive. Figure VV shows

heat maps of the Hb-Hg distances versus Og-Ob at site 3 for frames with and without a bridging water. The bridges seem only to form when the two hydrogens are close to each other, but the other interactions seem stable over a wide range of distances.

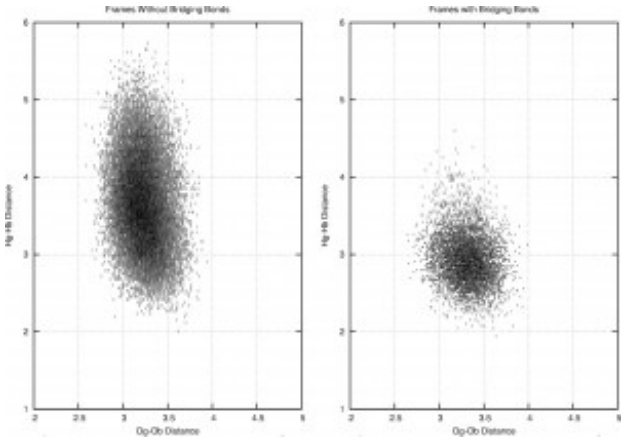


Figure VV: (need to fix fonts) Plots of the Hb-Hg distance (y axes) versus Ob-Og (x axes). Distances are in Angstroms. Distributions without (left) and with (right) a bridging water.

We could also consider the fraction of frames that any individual water molecule was involved in a bridging bond (Figure UU). The distribution is weighted heavily toward a single, or just a few, frames for any individual water molecule. This implies that even if the bridging bonds are important, that any individual water molecule is unlikely to occupy the bridging position for a long time or repeatedly.

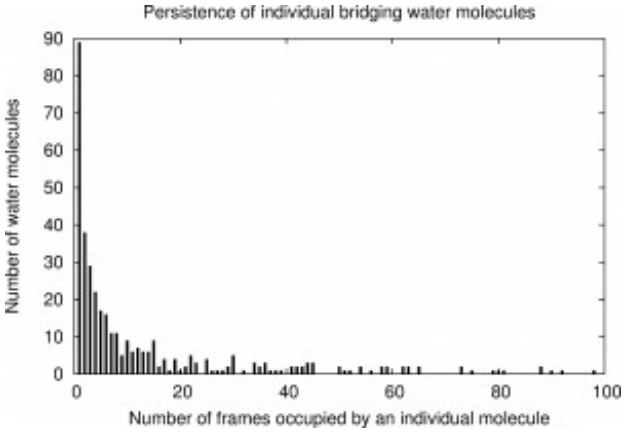


Figure UU: An indication of the persistence of any individual water molecule in a bridged bond. The x-axis represents the total observed number of frames, consecutive in the simulation or not, in which a particular water molecule occupied a bridging position between Hb and Hg. The y-axis is the number of molecules that occupied the position for that number of frames. Not shown are six points above 100 frames, ranging from 109 to 312, that corresponded to only one molecule.

Implicit Solvent Simulations

To further investigate the role of individual water molecules in the stabilization of the N-acetyl-galactosylated peptide lengths, we ran simulations of the molecules in implicit solvent (IS). To gauge the relative significance of the dielectric shielding alone, we also ran simulations in a water-like dielectric, but with no other solvent approximation. Since none of the currently available IS methods has been optimized for use with carbohydrates, we used all approximations to the Poisson-Boltzmann equation that are available in AMBER 14. To choose which methods to analyze, we measured the distances between the backbone nitrogens in Thr-3 and Lys-7. Those methods, *ALPB2* and *ALPB7*, whose distance distributions best approximated those from simulations with explicit waters were selected

for further analysis (see Figure ZZZ, below). The ALPB2 method makes a good approximation to the shape of the explicit simulation's distribution, and the ALPB7 method has a distribution average that is closest to that in the explicit simulation. The IGB2 and IGB7 methods also performed well. Comparison images for all simulations are available in the SI.

The ALPB2 and ALPB7 methods produced heat-maps that strongly resemble the heat-map for the explicit-water simulations. The dielectric-only heat-maps, however, are quite different. Although the IS methods are not tuned for use on carbohydrates, these results allow the tentative conclusion that the bulk properties of water play a significant role in structural stabilization of the N-acetyl-galactosylated moieties, but the actions of individual water molecules play a less important role, at least for this particular system. The results argue that the direct interactions of the GalNAc with the backbone, in the environment of aqueous solution, are the largest contributing factors to the structural stabilization.

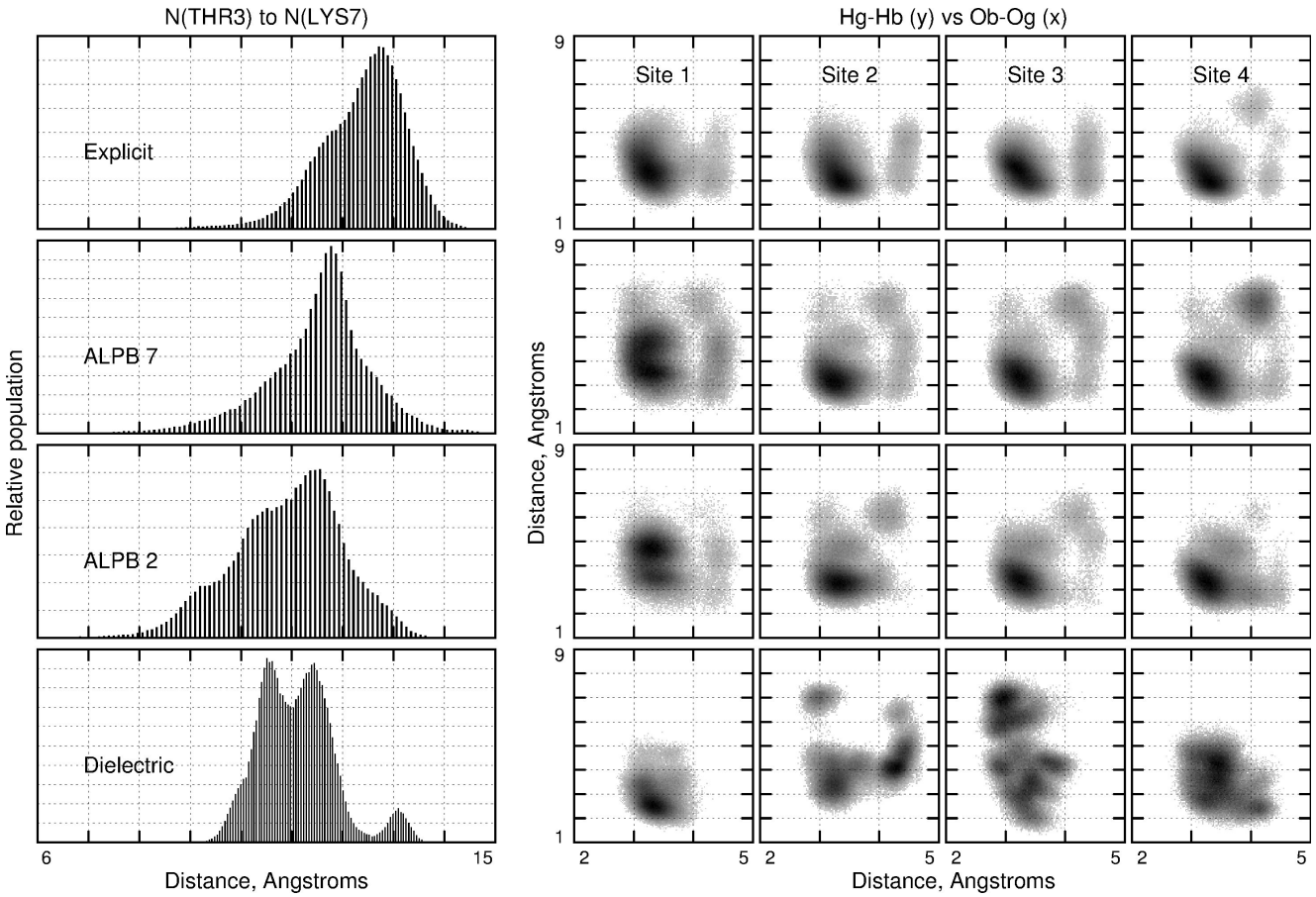


Figure ZZZ: Comparison of results from the explicit, implicit and dielectric-only runs. The top row of graphs all correspond to the explicit water simulations. The second and third rows correspond to simulations in implicit solvent using the Analytical Linearized Poisson Boltzmann (ALPB) methods labeled 2 and 7 in the AMBER manual. The bottom row of graphs correspond to a simulation in which there was no attempt to model solvent other than the imposition of a water-like dielectric. The left column of graphs shows the

distribution of THR-flanking, N-to-N lengths for each type of simulation. The right-most columns contain heat-maps of the distances (y axes) between the two hydrogens in the observed 4-way interaction versus the distances (x axes) between the two oxygens. The distributions are broken down per glycosylation site as noted.

(earlier stuff below here)

The basic argument goes like this:

There was thought that maybe the GalNAc's NAc H was making a H-bond with the O in the backbone. In the simulations, that doesn't happen very much. However, *something* is going on because there is a huge portion of time where there is almost an H-bond there. Another paper had suggested that there was an important bridging water between the NAc H and the backbone H. So, we looked for that in the sims where waters were saved. Turns out, not so much. However, there appears to be a cool 4-way thingy happening between the NAc H, the glycosidic O, the backbone O and the backbone H. So, to see if a bridging water might be implicated, we ran some implicit solvent simulations. These are all kinda bad because there are no good IS variants for carbs. However, we thought we'd try anyway because there can't be a bridging water in those runs. Turns out that a few varieties of IS do a not-terrible job of capturing the N-to-N length in the peptide backbone. So, we checked the 4-way presence in those. Turns out that they adopt that conformation a lot, too, at least in the sims where the lengths are similar (I think, and maybe others… still working on plots). So, we still can't be certain, but it looks as if that 4-way thing is the most important part of why the GalNAcs don't move around a lot.

Figures:

- H-bond heat maps for the 4 positions (already made, but add information regarding limits for H-bonding and the huge section of almost H-bond). This is for Ng-Hg-Ob angle versus Ng-Ob distance.
- Table showing fraction of frames with a bridged H-bond at each site.
- Nice image of the 4-way interaction.
- One with a bridging water.
- Heat maps of distances and/or "H-bond" maps like above with and without the bridging water
- NN length vs HH-OO heat map per each site and all that by simulation method (Exp, alpb7 and alpb2 are the ones for the main text).