

Ring Conformational Energetics for Rhamnogalacturonan-II Monosaccharides: Implications for Structural Dynamics of Complex Carbohydrates

Vivek Bharadwaj

Vivek.Bharadwaj@nlr.gov

National Laboratory of the Rockies

Sean Brooks

National Laboratory of the Rockies

Lintao Bu

National Laboratory of the Rockies

Michael Crowley

National Laboratory of the Rockies

Deepak Sharma

University of Georgia

Breeanna Urbanowicz

University of Georgia <https://orcid.org/0000-0001-5247-4513>

Article

Keywords:

Posted Date: August 5th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10398105/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Ring Conformational Energetics for Rhamnogalacturonan-II Monosaccharides: Implications for Structural Dynamics of Complex Carbohydrates

Sean E. Brooks¹, Lintao Bu¹, Deepak Sharma², Michael F. Crowley¹, Breeanna Urbanowicz^{2,3}, and Vivek S. Bharadwaj^{1*}

¹Renewable Resources and Enabling Sciences Center, National Laboratory of the Rockies, Golden, CO 80401, USA

²Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602, USA

³Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA

*Correspondence to: vivek.bharadwaj@nlr.gov

Abstract

Ring puckering governs the three-dimensional structure of glycans, and conformational changes in a single monosaccharide can reshape the architecture and function of an entire polysaccharide. Structural characterization in glycobiology is challenged by constituent sugar diversity, their varied linkages, and the conformational flexibility of sugar rings. Rhamnogalacturonan-II (RG-II), found in the cell walls of all vascular plants, is the most structurally complex polysaccharide known in nature, comprising 14 sugars linked through 21 glycosidic bonds. Here we resolve the ring conformational energetics of every sugar and uniquely bonded variant in RG-II using quantum mechanical calculations at the MP2/cc-pVDZ level. The resulting free energy atlases reveal shared conformational traits among pyranoses and show that exocyclic substituents substantially impact ring energetics, an important consideration for modeling glycosidic linkages in complex carbohydrates. Integrating these calculations with molecular dynamics simulations and nuclear magnetic resonance measurements, we identify ring pucker interconversion of arabinopyranose as the molecular origin of distinct spectroscopic signatures across RG-II glycoforms. We further show that pucker transformations at the monosaccharide level propagate to govern side chain structure in the assembled polysaccharide. These findings establish a generalizable computational framework for linking monosaccharide energetics to polysaccharide structure, with broader implications for glycan-based biomaterials and therapeutics.

1. INTRODUCTION

Carbohydrates are a diverse class of biological molecules that are ubiquitous in all kingdoms of life. In contrast to other biopolymers such as DNA and proteins, carbohydrates exhibit more complex challenges to structural characterization. While DNA contains four base pairs and proteins are composed of up to 20 common amino acids¹, carbohydrates can be constructed from over 100 different monosaccharides². Beyond monosaccharide diversity, carbohydrates also exhibit a much larger variety of internal linkages and conformations and may be attached through glycosidic linkages on up to five or six carbons in a furanose or pyranose sugar, respectively¹. In contrast, a DNA nucleotide or peptide can form only two respective covalent bonds¹. The structural complexity of carbohydrates is further complicated by the fact that the cyclic nature of each monosaccharide enables it to adopt one of numerous possible ring conformations or pucker states (20 for furanoses and 38 for pyranoses)³.

This complexity is perhaps best exemplified by the pectic polysaccharide rhamnogalacturonan-II (RG-II), which is often considered the most complex carbohydrate in nature. RG-II is present in all vascular plants and exists as a dimer in the cell wall, comprising up to 4% of total cell wall mass⁴. The ability of RG-II to dimerize via boric acid-mediated crosslinking is critical for plant growth and development^{5–8}. However, determining the structure of carbohydrates such as RG-II poses experimental challenges because traditional analytical methods are often insufficient. X-ray crystallography, for instance, requires rigid crystals that highly flexible carbohydrates usually cannot form⁹. Nuclear magnetic resonance (NMR) faces similar challenges resulting from the complex stereochemistry and heterogeneous inter-residue linkages found in carbohydrates¹⁰. Computational approaches, involving quantum mechanics (QM) and molecular dynamics (MD), enable sampling across the full conformational space, which can help identify low energy conformers that may be inaccessible to experimental methods¹¹.

Ring conformations can be described by Cremer-Pople (CP) puckering parameters, which define a series of pucker states based on the relative locations of ring atoms above and below the central ring plane in spherical coordinates (**Figure 1.a,b**)³. Five-membered rings are described by 20 canonical pucker states with two degrees of freedom: pseudorotation angle (Φ) and puckering amplitude (Q)³. Canonical pucker states for five-membered rings are categorized as envelope (E) or twist (T)³. Furanoses often exhibit highly flexible conformational dynamics, with thermodynamically favorable pucker states depending on the location and type of exocyclic groups^{13–15}. Six-membered rings are described by 38 canonical pucker states with three degrees of freedom: pseudorotation angle (Φ), phase angle (θ), and total puckering amplitude (Q)³. Canonical pucker states for six-membered rings are categorized as chair (C), boat (B), skew (S), envelope (E), or half-chair (H)³. Theoretical^{16,17} and experimental¹⁸ literature demonstrates that the two chair pucker states (${}^1C_4$ and ${}^4C_1$) are nearly universally preferred in six-membered heterocyclic rings. Similar to furanoses, preference for the ${}^1C_4$ chair versus the ${}^4C_1$ chair is largely determined by the location, orientation, and type of exocyclic groups¹⁷.

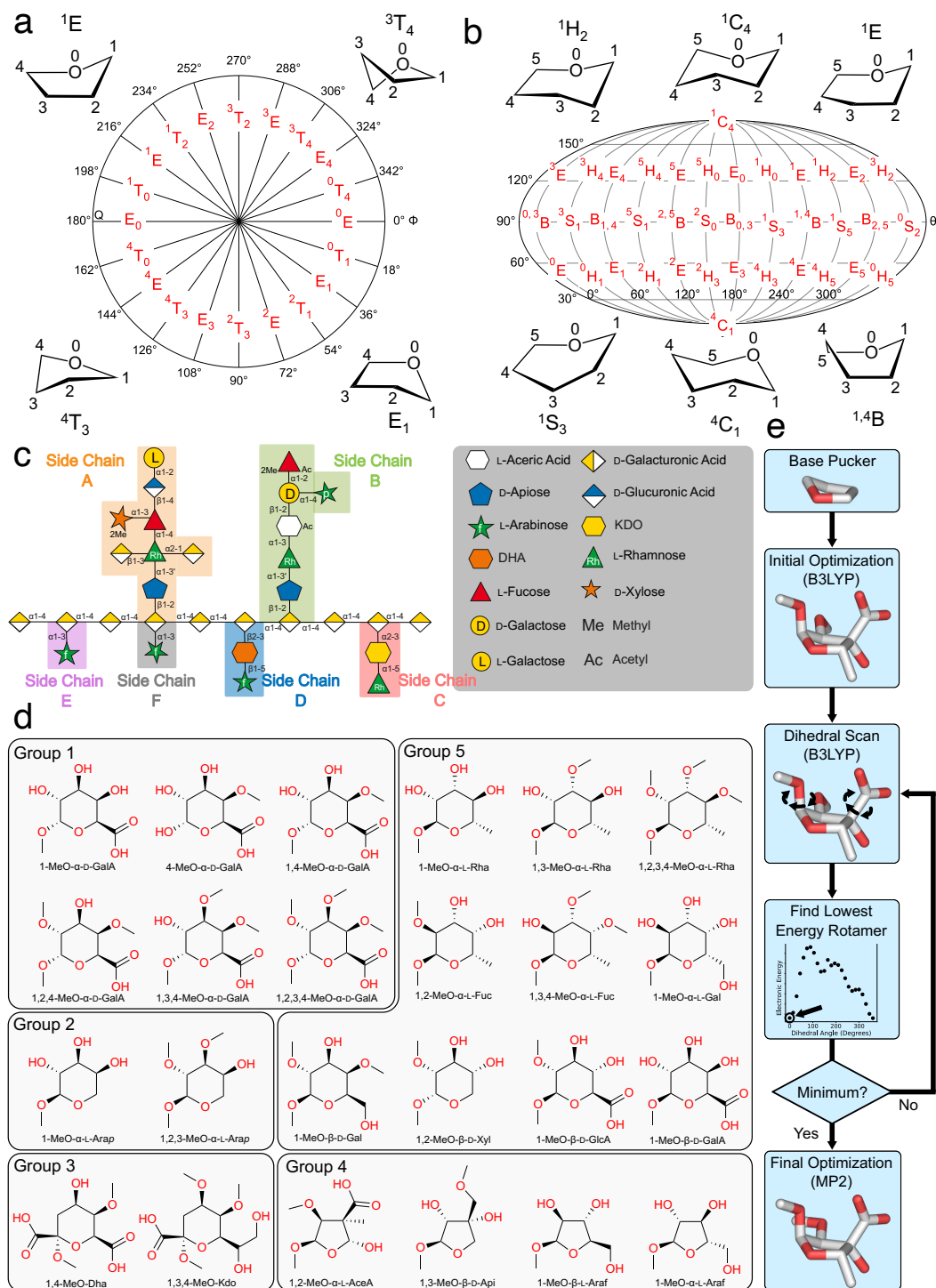


Figure 1. Structural and Computational Framework for RG-II Conformational Analysis (a) Polar plot representing canonically defined pucker states for furanoses. (b) Molleweide projection plot representing canonically defined pucker states for pyranoses. (c) Schematic of the primary structure of celery RG-II represented according to the Symbol Nomenclature for Glycans (SNFG) convention¹². (d) Structures and abbreviated nomenclature for all sugars in RG-II. Monosaccharides were grouped into five structural classes: galacturonic acids (Group 1), arabinopyranoses (Group 2), Kdo- and Dha-containing pyranoses (Group 3), furanoses (Group 4), and all remaining monosaccharides (Group 5). (e) Calculation procedure flowchart.

The monosacchide diversity in RG-II makes it an ideal system to examine the influence of ring puckering on the overall structural dynamics of a complex carbohydrate. RG-II contains 4 furanose sugars and 11 pyranose sugars, including some uncommon sugars such as apiose (3-C-(hydroxymethyl)- β -D-glycero-tetrofuranose)¹⁹, Dha (3-deoxy-D-lyxo-hept-2-ulopyranosaric acid)²⁰, and Kdo (3-deoxy-D-manno-oct-2-ulopyranosonic acid)²¹. Furthermore, the same sugar may be present in different forms within RG-II based on the unique set of glycosidic linkages connecting to neighboring sugars. For instance, α -D-GalA occurs as substituted variants with one, two, three, or four glycosidic linkages depending on location in the backbone and side chains (**Figure 1.c**)²². RG-II variants, such as the substituted arabinopyranose (α -L-Arap) found in side chain B of papaya RG-II, provide further case studies for relating computationally derived structural dynamics to experimental observations.

A quantum mechanics (QM) study has not yet been conducted to determine free energy landscapes for all RG-II monosaccharides. Mazeau and Pérez investigated the conformational energy landscapes of aceric acid (3-C-carboxy-5-deoxy-L-xylofuranose) and apiose with the MM3(92) force field²³. However, they omitted analysis of the two arabinofuranose (Araf) sugars and all pyranose sugars that constitute the majority of RG-II's monomeric units, instead opting to treat the pyranose sugars together in a fragment-based simulation approach²³. Our group's prior QM work on RG-II characterized the reaction mechanism underpinning RG-II dimerization²⁴, as well as predicted peak shifts in ¹³C and ¹H NMR spectra^{24,25}. Additionally, we showed that MD simulations combined with experimental NMR measurements can be used to characterize the structure and monosaccharide sequence of RG-II²².

While QM has not been applied to study the sugars in RG-II, this computational technique has proven useful for probing free energy landscapes in other glycans. Mayes and coworkers used density functional theory (DFT) to explore the conformational free energy landscape of pucker states in five common pyranose sugars: xylose, α -glucose, β -glucose, β -mannose, and β -N-acetylglucosamine¹⁶. Numerous studies have also been conducted to scan free energy surfaces in a variety of monosaccharides using classical MD^{11,26–29} and enhanced sampling techniques^{17,30}. While classical physics-based methods are useful for scanning a vast conformer space, QM simulations typically provide higher accuracy *ab initio* energy predictions, with the added consideration of electronic effects, and are therefore of significant interest as a complimentary computational approach³¹.

In this study, we undertake a comprehensive evaluation of ring conformational energetics and present ring pucker free energy atlases for all RG-II monosaccharides using QM calculations. To ensure their relevance for understanding overall RG-II structure, we account for the influence of neighbors for each monosaccharide by capping every glycosidic linkage with a methoxy group.

We have developed an open source code-base (SUCER-qm: <https://github.com/vbharadwcomosci/sucerQM>) that allows users to easily generate structural coordinates for all pucker states of a ring molecule using just a SMILES string, as well as generate direct inputs for QM calculations. This procedure creates all pucker states in any substituted five- or six-membered ring without having to add exocyclic groups manually, as was previously standard^{32,33}.

We find that, in general, furanose energy surfaces allow for greater flexibility of the ring compared to pyranoses. Comprehensive evaluation of 20 pyranoses reveals that the chair conformations ¹C₄ and ⁴C₁ are the most favorable for all of them. We also find that the presence of exocyclic

groups on the same sugar can alter the energy atlas and impact plausible ring-pucker interconversion pathways. QM calculations combined with microsecond-timescale MD and experimental NMR were then used to explain the phenomenon of ring-pucker interconversion observed for substituted α -L-Arap in papaya RG-II. Finally, we demonstrate the impact that pucker state transitions can have on the overall structure of RG-II, which may influence interactions with other cell wall biopolymers, physiological solvents, and enzymes.

2. RESULTS AND DISCUSSION

When decomposed into individual sugars, RG-II yields 24 unique, methoxy-capped cyclic molecules, of which 4 are furanoses and 20 are pyranoses. To simplify our analysis, the sugars are classified into five groups: Group 1, galacturonic acids; Group 2, arabinopyranoses; Group 3, the unusual pyranoses Kdo and Dha; Group 4, furanoses; and Group 5, the remaining sugars, including rhamnose, galactose, fucose, xylose, and glucuronic acid (**Figure 1.d**). The zero-point energies, enthalpies, free energies, electronic energies, and fully unconstrained optimizations from our calculations are reported in **SI Sections S1, S2, S3, S4, and S5**, respectively, for all sugars ordered group-wise. Here, we discuss notable observations from our analysis on sugars that are integral to the structure and function of RG-II.

2.1 Chair conformations are the most stable for pyranose sugars, and a variety of metastable pucker states occur in the CP-equatorial region

The distribution of high- and low-energy pucker states across all pyranose sugars demonstrates a strong preference for chair (${}^1C_4$ and ${}^4C_1$) (**Figure 2.a**). Every pyranose sugar preferred either the ${}^1C_4$ or ${}^4C_1$ chair as its lowest energy pucker state, and only two sugars, Kdo and Dha, did not include both chairs as their two lowest energy pucker states. More sugars preferred the ${}^4C_1$ chair than the ${}^1C_4$ chair, which can be explained by the large number of uronic acids that are included in this study. For the uronic acids, the ${}^4C_1$ chair positions the bulky carboxylic group in the more stable equatorial orientation, reducing steric clashes.

Of the ten non-uronic acid constituents (aldoses) of RG-II, seven prefer the ${}^1C_4$ chair, whereas a galactose (1-MeO- β -D-Gal), xylose (1-MeO-Xyl), and 1,2,3-MeO-Arap prefer the ${}^4C_1$ chair as the lowest energy pucker state. The mixed preference for ${}^1C_4$ and ${}^4C_1$ chairs exhibited by the aldoses is expected considering that they lack the bulky carboxylic group that requires the ${}^4C_1$ chair to be energetically favorable in uronic acids. Among groups containing the same sugar with different substitution patterns, every group showed a preference for the same chair across all differently substituted variants except for the Arap sugars, which exhibit a lowest energy pucker state of ${}^1C_4$ in 1-MeO-Arap compared to ${}^4C_1$ in 1,2,3-MeO-Arap.

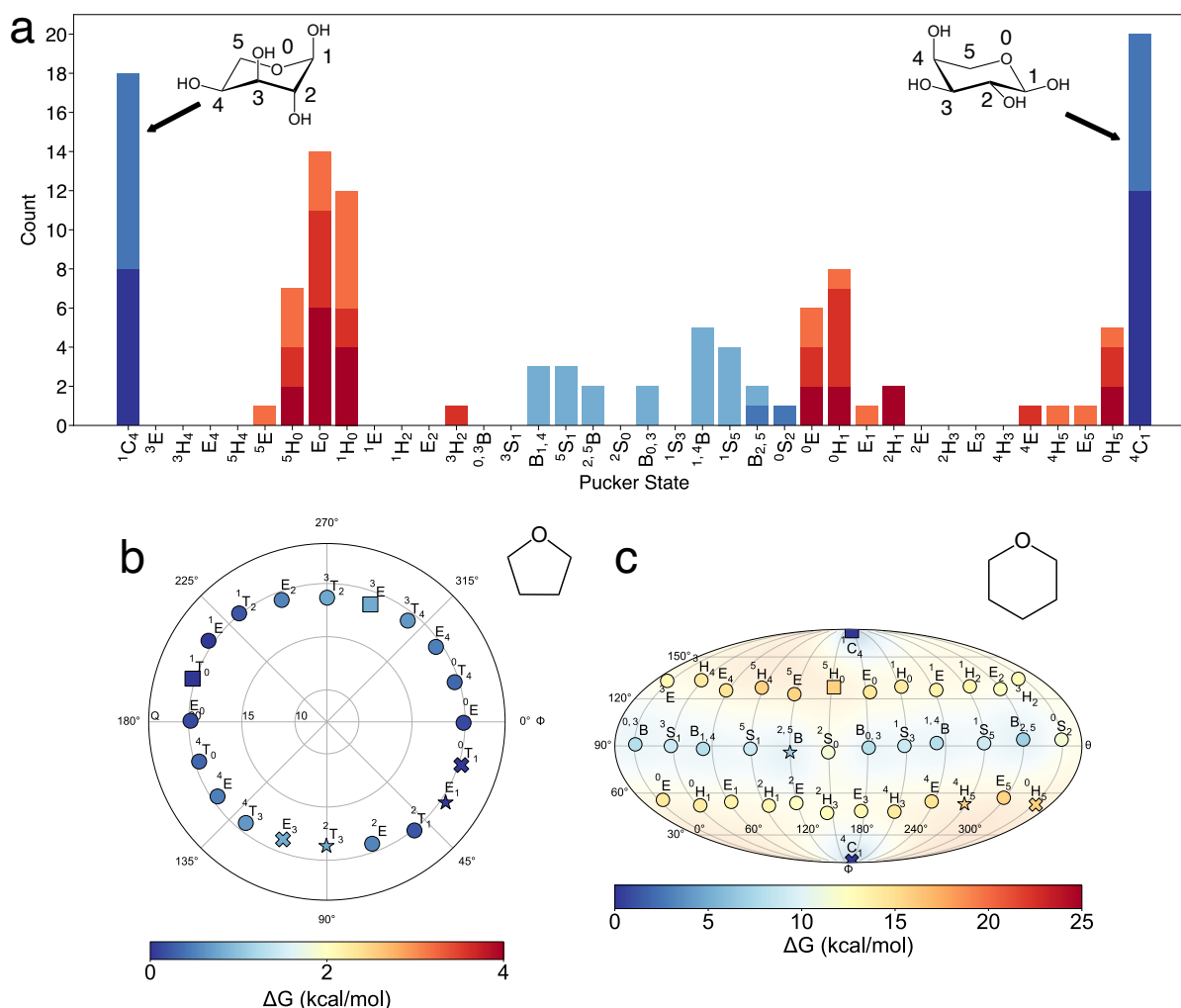


Figure 2. Conformational Energy Landscapes and Pucker-State Preferences of Carbohydrate Rings
 (a) Distribution of lowest and highest energy pucker states across all pyranose sugars. Blue represents the low energy pucker states, with the darkest blue representing the lowest energy pucker state, the second darkest blue being second lowest energy pucker state, and lightest blue being the third lowest energy pucker state. Red represents the high energy pucker states, with the darkest red representing the highest energy pucker state, the second darkest red being second highest energy pucker state, and lightest red being the third highest energy pucker state. (b) Free energy landscape for the unsubstituted furanose ring. (c) Free energy landscape for the unsubstituted pyranose ring.

A variety of metastable pucker states (local energy minima) appear in the CP-equatorial region, most prominently ${}^1,4B$ and ${}^1S_5$. Equatorial pucker states are vital for understanding conformational dynamics since they influence the favorability and nature of ring-pucker interconversion pathways^{34,35}. This is particularly relevant when considering glycan-enzyme interactions. Glycans bind with proteins to undergo reactive chemical transformations and for molecular recognition, which entail significant modifications to the glycan conformation³⁶. In glycosidase-catalyzed hydrolysis, for instance, active site binding is known to result in the glycan substrate adopting an unorthodox pucker state that activates the glycosidic bond for cleavage³⁶. Thus, metastable equatorial pucker states enable energetically feasible interconversion pathways that provide the catalytic pathways of glycan- and carbohydrate-active enzyme functions.

The highest energy pucker states (**Figure 2.a**) are dominated by conformations in which the ring oxygen (atom number 0) is out of plane. Specifically, envelope (^0E and E_0) and half chair ($^5H_0$, $^0H_5$, $^0H_1$, and $^1H_0$) pucker states exhibit the greatest incidence of high energy conformations. The energetics of a given pucker state are determined by a combination of factors, including ring strain, steric effects from exocyclic groups, and presence of hydrogen bonding. We calculated the pucker energetics for an unsubstituted furanose (**Figure 2.b**) and pyranose ring (**Figure 2.c**) as baselines to isolate the energetic contributions from ring strain. This enables us to deconvolute energetic contributions from other factors, including namely steric effects and hydrogen bonding.

Interestingly, the E_0 and $^1H_0$ pucker states, which are the two most frequently occurring high energy pucker states among substituted pyranoses, do not appear among the highest energy pucker states for the unsubstituted pyranose ring. This indicates that steric effects from exocyclic group orientation, rather than ring strain, significantly contribute to the overall energetics of certain pucker states. For furanoses, our analysis of pucker energetics for the unsubstituted ring reveals relatively small (~ 1 kcal/mol) energy differences between states (**Figure 2.b**).

2.2 Conformational energetics are influenced by exocyclic groups

RG-II⁴ and RG-I³⁷ are branched polysaccharides consisting of sugars connected to each other with multiple glycosidic linkages. One such sugar is galacturonic acid (GalA), which forms the homogalacturonan backbone of RG-II, and is also found as a terminal sugar in side chain A (Group 1, **Figure 3.a**). Previous studies investigating sugar pucker energetics typically omit glycosidic linkages, instead opting to model these bonds as simple hydroxyl groups^{16,28}. Other QM studies that analyze the energetic effect of glycosidic linkages do so with simplified furanose-³⁸ and pyranose-like^{39–41} model compounds that are not sugars, typically methoxy-substituted tetrahydrofuran and tetrahydropyran variants. The presence of α -D-GalA variants that include one, two, three, or four glycosidic linkages (represented by exocyclic methoxy groups) on various ring carbons enables us to examine the energetic and structural effects of ring substitution for real sugars in RG-II (**Figure 3.d-h**). We also model a fully unsubstituted α -D-GalA variant (**Figure 3.c**) for comparison, even though it is not part of the RG-II structure.

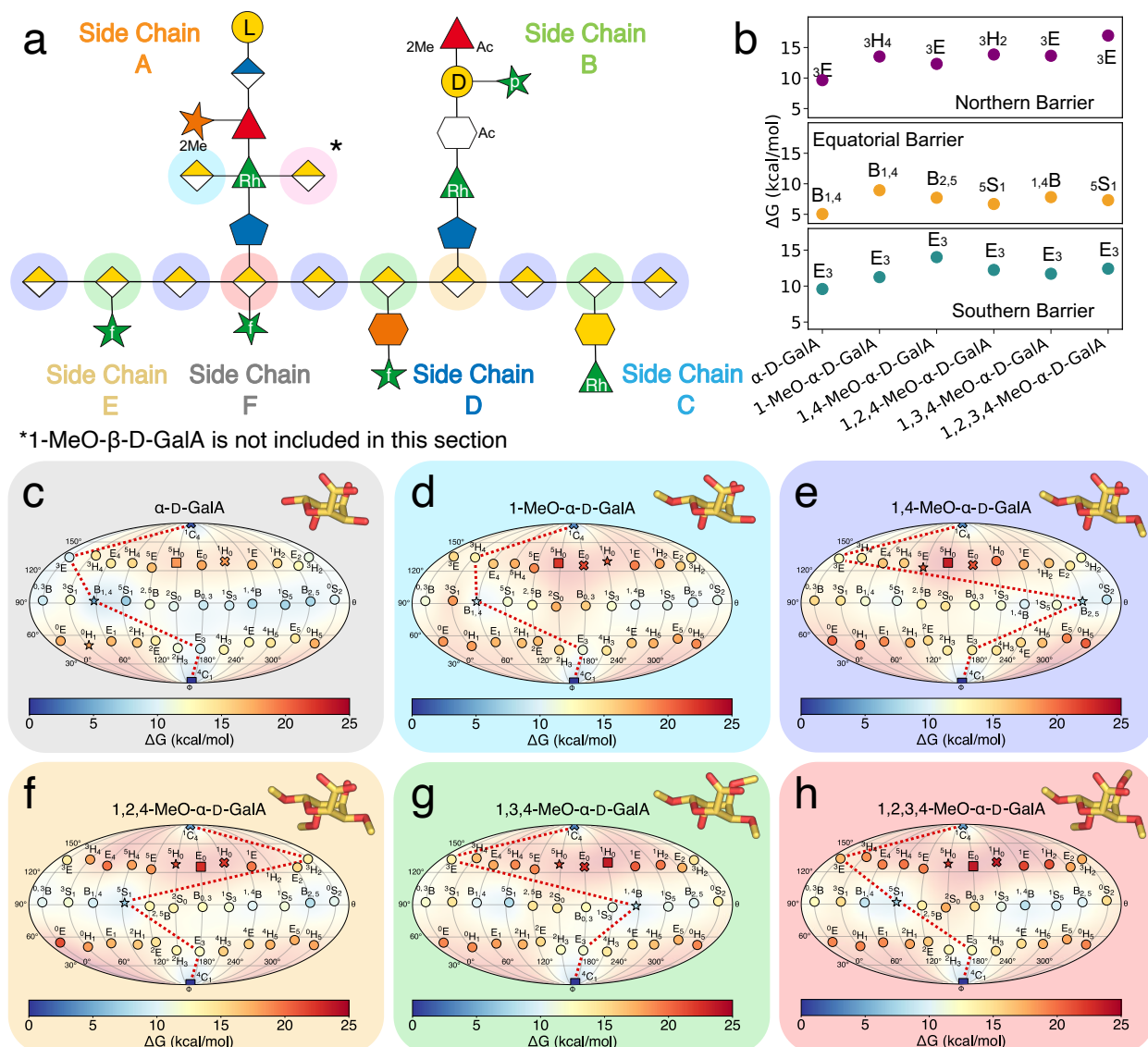


Figure 3. Ring-Pucker Energetics of Backbone and Side-Chain α -D-GalA Residues in RG-II (a) Color-coded locations of α -D-GalA variants within the RG-II backbone and side chains. The backbone contains 1-MeO- α -D-GalA, 1,4-MeO- α -D-GalA, 1,2,4-MeO- α -D-GalA, 1,3,4-MeO- α -D-GalA, and 1,2,3,4-MeO- α -D-GalA, and side chain A also contains 1-MeO- α -D-GalA. (b) Minimum energy barriers in the CP-northern, CP-equatorial, and CP-southern regions for each α -D-GalA variant. (c-h) Free energy landscapes for all α -D-GalA variants in RG-II (*1-MeO- β -D-GalA is not included). The free energy of each pucker state is reported relative to the lowest energy pucker state for that sugar (${}^4C_1$ is the lowest energy pucker state for every α -D-GalA variant). Values of ϕ and θ for each marker are plotted based on their final values from the constrained optimization. The red dotted line connects the lowest energy pucker state in each region of the plot to show minimum ring-pucker interconversion barriers in the CP-northern, CP-equatorial, and CP-southern regions. Interpolation between the energy of each discrete calculation is displayed in the background as color shading to provide an overview of the free energy surface. The three special markers are as follows for pucker state free energies: (1) square, lowest or highest; (2) cross, second lowest or highest; and (3) star, third lowest or highest.

A comparison of pucker free energy atlases for α -D-GalA variants in RG-II is shown in **Figure 3.c-h**. One of the most distinctive features of the conformational landscapes is the marked variation

in the minimum CP-northern interconversion barrier among α -D-GalA variants. (**Figure 3.b**), which, in general, increases with the number of glycosidic linkages. Minimum energy barriers in each region of the plot are important for describing kinetically relevant ring-pucker interconversion pathways (depicted with red dashes on the CP plot in **Figure 3.c-h**), where a ring progresses from one pucker state to another, such as during enzyme catalysis¹⁶.

The largest sequential increase in the minimum energy barrier is observed when moving from unsubstituted α -D-GalA to 1-MeO- α -D-GalA. For α -D-GalA (**Figure 3.c**), the minimum energy barrier in the CP-northern region is ³E (9.65 kcal/mol), whereas it is ³H₄ for 1-MeO- α -D-GalA (13.53 kcal/mol). This barrier difference of nearly 4 kcal/mol demonstrates the significant energetic contribution of the anomeric methoxy group compared to the unsubstituted ring. While previous theoretical literature has established the energetic contribution of anomeric glycosidic linkages in pyranose-like model compounds⁴², we present a side-by-side comparison of pucker energetics for a real sugar both with and without an anomeric glycosidic linkage.

Even more drastically, the most substituted variant, 1,2,3,4-MeO- α -D-GalA, has a CP-northern minimum energy barrier at ³E (16.96 kcal/mol), which represents an increase of 3.4 kcal/mol compared to the substituted 1-MeO- α -D-GalA, and 7.3 kcal/mol compared to the unsubstituted α -D-GalA. A larger kinetic barrier to ring-pucker interconversion suggests that highly substituted α -D-GalA variants connecting the side chains to the backbone are expected to demonstrate less ring flexibility than terminal variants in side chains. These findings challenge QM sugar methodology that focuses on modeling only the unsubstituted, isolated variant of a sugar¹⁶ when considering it in the context of a larger polysaccharide construct. Here, we observe that ϕ values for the CP-northern minimum energy barrier are narrowly confined to a region of 70° between 330° and 40°. Thus, while the barrier energy may change significantly between variants, the pucker state at which the barrier occurs shows only slight deviation.

We also find that metastable pucker states in the CP-equatorial region further characterize the thermodynamics of ring-pucker interconversion. In general, as the number of methoxy groups increases, the relative free energy of metastable states in the CP-equatorial region decreases (**Figure 3.b**). The exception to this observation occurs when comparing B_{1,4} (5.05 kcal/mol) in unsubstituted α -D-GalA to B_{1,4} (8.94 kcal/mol) in substituted 1-MeO- α -D-GalA, which represents an increase of 3.89 kcal/mol. In this case, addition of the anomeric methoxy group significantly increases the relative free energy of the metastable CP-equatorial barrier, leading to reduced thermodynamic stability. However, comparison across substituted variants indicates that the presence of additional methoxy groups decreases the relative free energy of these metastable states. For instance, 1,2,3,4-MeO- α -D-GalA exhibits a CP-equatorial barrier of 7.32 kcal/mol at ⁵S₁, compared to 8.94 kcal/mol at B_{1,4} for the less substituted 1-MeO- α -D-GalA, a difference of 1.62 kcal/mol. More highly substituted sugars appear to face a greater thermodynamic barrier to interconversion since the presence of lower energy CP-equatorial pucker states increases the probability that the sugar becomes trapped in a metastable conformation and cannot progress to one of the two chairs.

Free energy barriers across α -D-GalA variants demonstrate that, in general, as the number of glycosidic linkages increases, the maximum free energy barrier along the ring interconversion pathway increases. Simultaneously, the relative free energy of metastable pucker states in the CP-equatorial region decreases as the number of glycosidic linkages increases, which indicates a thermodynamically stabilizing effect for more highly substituted sugar variants. Taken together,

these findings indicate that highly substituted α -D-GalA variants connecting the backbone to side chains are expected to exhibit reduced conformational flexibility.

Large energetic differences are also observed when comparing the unsubstituted α -D-GalA variant with the anomERICALLY substituted 1-MeO- α -D-GalA variant. While less drastic, significant energy differences also appear among substituted α -D-GalA variants in all regions of the CP plot. We note that the energetic trends observed in this section are specific to α -D-GalA and may not be directly generalizable to other sugars. Nevertheless, we demonstrate that including specific glycosidic linkage patterns may be important for determining accurate pucker state energetics.

2.3 Implications of substitution in *Arap* across RG-II variants

The structure of RG-II is generally conserved among all species of plants, but some variations have been observed in nature⁴³. For instance, RG-II side chain B may be substituted with rhamnoses in papaya or unsubstituted in other plants such as celery⁴³. This structural variation provides further opportunity to examine the effect of substitution on ring-pucker interconversions. While we have already analyzed substitution for the α -D-GalA variants in RG-II, we hypothesize that substitution patterns may impact the free energy landscapes of other sugars differently. Furthermore, understanding the conformational accessibility of pucker states in various sugars is also relevant when interpreting NMR spectra. Here, we combine QM simulations with microsecond MD and experimental NMR measurements to investigate how substitution impacts ring-pucker interconversions in *Arap*.

Comparing free energy landscapes for unsubstituted (celery) and substituted (papaya) *Arap* shows that the substituted *Arap* variant is significantly stabilized in the CP-equatorial region (**Figure 4.a,b**). For instance, B_{0,3} has an energy of 3.39 kcal/mol in substituted *Arap*, compared to 8.24 kcal/mol in unsubstituted *Arap*. Likewise, the minimum energy barrier in the CP-northern region is lower in substituted *Arap* (10.27 kcal/mol) than unsubstituted *Arap* (11.75 kcal/mol).

Based on the QM atlases, the substituted *Arap* variant is expected to demonstrate increased conformational flexibility due to its energetically accessible equatorial region and comparatively low maximum energy barrier. This hypothesis was confirmed with both microsecond-scale MD simulations and experimental NMR measurements.

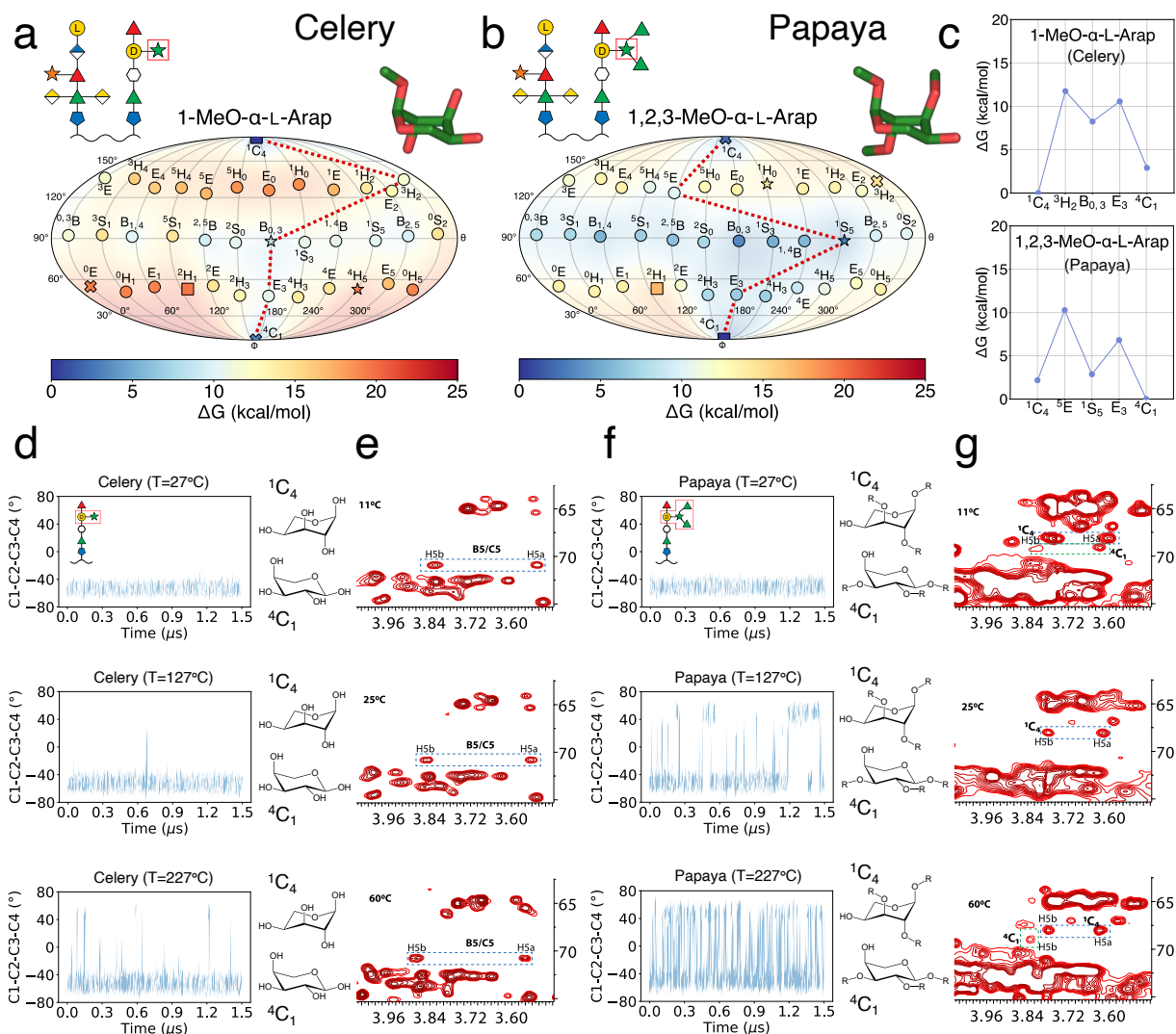


Figure 4. Substitution-Dependent Ring Flipping of Arabinopyranose Residues in RG-II Side Chain
B (a-b) Free energy landscapes for Arap variants in RG-II from different plant species: celery (unsubstituted) and papaya (substituted), highlights in a red box. Plot construction and marker scheme are the same as described in Figure 3. (c) Lowest-energy pucker interconversion pathways for unsubstituted and substituted Arap variants (d) MD simulations showing lack of ring flip for the unsubstituted Arap in celery. The red box shows sugars included in the simulation. (e) Experimental NMR shifts for unsubstituted monomeric RG-II in celery. (f) MD simulations showing ring flip for the substituted Arap in papaya. The red box shows sugars included in the simulation. (g) Experimentally determined NMR shifts for substituted monomeric RG-II in papaya.

To observe ring interconversion, MD simulations were conducted on side chain B residues containing D-Gal and Arap residues, as well as the two connected rhamnose (Rha) residues for the substituted variant (**Figure 4.d,f**). At a temperature of 27°C, chair conformations are observed for both the unsubstituted and substituted side chain B variants, with no ring interconversion occurring throughout the 1.5 microsecond timeframe. As the temperature is increased to 127°C and 227°C, significant and frequent ring interconversion is observed for the substituted side chain variant, whereas interconversion is not observed for the unsubstituted side chain variant. These

results provide further evidence that *Arap* substitution results in greater conformational flexibility, which enables ring interconversion.

To experimentally validate the conformational behavior of *Arap*, HSQC spectra of monomeric RG-II (mRG-II) isolated from either celery or papaya were recorded at three temperatures (11°C, 25°C, and 60°C). The analysis focused on the C5–H5 correlations of the *Arap* residue, as C5 is a methylene carbon bearing two diastereotopic protons (H5a and H5b) that are highly sensitive to changes in ring conformation. In pyranose rings, the equilibrium between the ${}^1C_4$ and ${}^4C_1$ chair conformations can influence the chemical environment of these protons, and under conditions where conformational exchange slows on the NMR timescale, signals corresponding to minor conformers may become observable.

Distinct temperature-dependent conformational behavior is observed, which appears to correlate with differences in substitution patterns of the *Arap* residue. In celery-derived mRG-II, where the *Arap* residue is substituted only at the anomeric position, the HSQC spectra across all three temperatures show only a single set of C5–H5 correlations (**Figure 4.c**). This indicates that a single conformer predominates under the experimental conditions, most consistent with the ${}^4C_1$ chair conformation. The absence of additional peaks suggests that the unsubstituted ring experiences minimal steric or electronic constraints in the ${}^4C_1$ chair, allowing this conformation to dominate, and preventing the detection of alternative conformers within the accessible NMR timescale. Experimentally observed NMR peak shifts corresponding to the ${}^1C_4$ and ${}^4C_1$ chair conformations were verified using Gauge-Including Atomic Orbitals (GIAO) calculations²⁵ with implicit D₂O solvation to match experimental conditions, which are presented in **SI Section 7**.

The papaya-derived mRG-II displayed a notably different conformational preference. At 11°C, inspection of the HSQC spectrum indicates that the ${}^1C_4$ conformation is the predominant form of the *Arap* residue (**Figure 4.f**). A clear additional peak corresponding to one of the C5 protons of a minor conformer is observed, while the second C5 proton signal is weaker but still detectable. At 25°C, only signals corresponding to the ${}^1C_4$ conformer remained clearly visible, suggesting that those from the minor conformer are broadened or overlapped as conformational exchange becomes faster. Interestingly, at 60°C, signals attributable to the ${}^4C_1$ conformer become detectable again in addition to the dominant ${}^1C_4$ peaks. In this case, a distinct additional peak corresponding to one C5 proton is clearly resolved, whereas the second proton signal remains weak or overlapped. This behavior may result from subtle temperature-induced chemical shift changes that alter peak overlap conditions, allowing previously obscured signals of the minor conformer to become visible.

These experimental observations suggest that the conformational equilibrium of the *Arap* residue in RG-II is strongly influenced by its substitution pattern and local structural environment within the polysaccharide⁴⁴. In the unsubstituted *Arap* residue of celery mRG-II, the ring adopts a stable ${}^4C_1$ conformation across the examined temperature range; however, papaya mRG-II appears to favor the ${}^1C_4$ conformation, indicating that its specific substitution pattern and glycosidic linkages may stabilize this chair geometry.

We hypothesize substituents on the pyranose ring significantly influence ring puckering by altering the axial or equatorial orientation of glycosidic bonds and modifying steric interactions within the RG-II side chain. Indeed, our experimental results agree with the QM and MD simulations, which both suggest that the substituted *Arap* variant enables greater conformational flexibility than the unsubstituted variant. These results for *Arap* are in contrast with findings for the α -D-GalA variants

presented in Section 2.2, in which increasing substitution was found to decrease conformational flexibility. A critical difference is the bulky carboxylic group on C5 of α -D-GalA that causes significant steric effects that are not present in Arap. This further emphasizes the importance of modeling unique substitution patterns for each sugar individually because energy trends can vary widely and unpredictably by sugar due to both complex steric and electronic effects.

2.4 Hydrogen bonding produces extensive, low-energy pucker conformations for Dha and Kdo despite bulky exocyclic groups

Dha and Kdo are uncommon, highly substituted sugars found in RG-II (Group 3) that exhibit unique hydrogen bonding patterns. They are also the only sugars in RG-II that do not include both the ${}^1C_4$ and ${}^4C_1$ chairs as their two lowest energy pucker states. Both exhibit ${}^4C_1$ as their lowest energy pucker state, but the second lowest energy is $B_{2,5}$ for Dha and ${}^0S_2$ for Kdo. These CP-equatorial pucker states are energetically favorable to the ${}^1C_4$ chair because both Dha and Kdo possess bulky substituents on their C1 and C5. Dha is decorated with a carboxylic acid group on each of C1 and C5, respectively, whereas Kdo has a carboxylic acid group on C1 and a (1,2)-hydroxyethyl group on C5. The ${}^4C_1$ chair is energetically favorable because it is the only pucker state that allows the bulky exocyclic groups on C1 and C5 to both occupy the equatorial orientation (Figure 5.c,d).

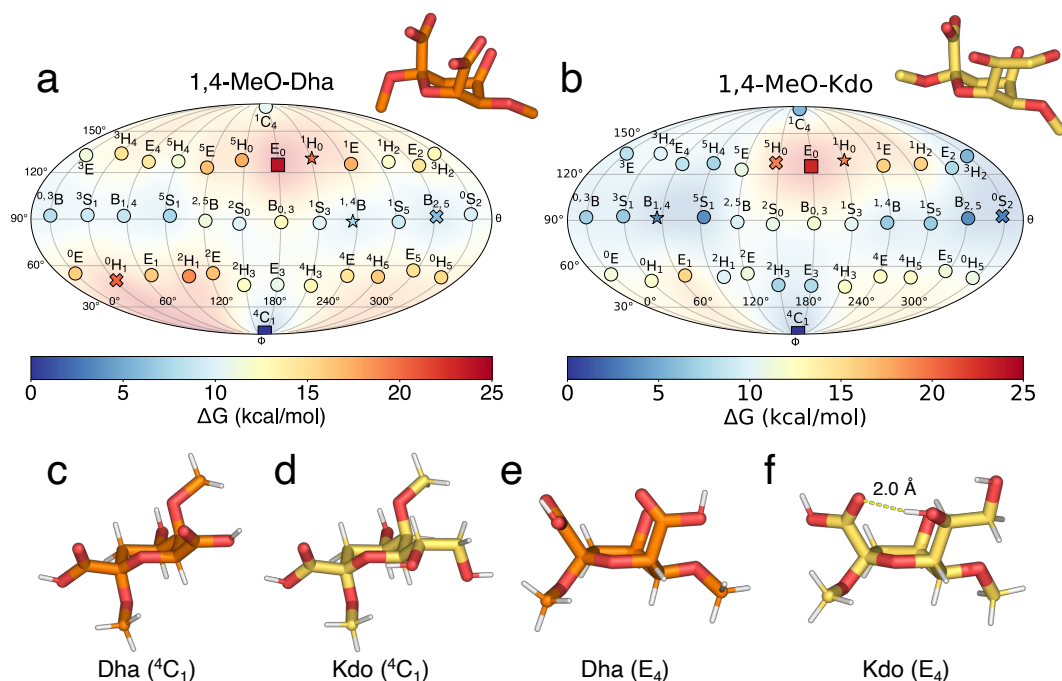


Figure 5. Distinct Pucker Energetics for Dha and Kdo. (a-b) Free energy atlases for Dha and Kdo. Note the single high energy pucker region in the CP-northern region centered around E_0 . Kdo demonstrates a relatively higher number of low energy states across the CP-northern and CP-southern regions compared to Dha. (c-d) ${}^4C_1$ chair pucker state for Dha and Kdo, respectively, emphasizing the axially oriented C1 and C5 exocyclic groups. (e-f) Representations of the E_4 pucker state for Dha and Kdo, respectively, to show differences in hydrogen bonding patterns.

Interestingly, the relatively large 10.47 kcal/mol free energy difference between the ${}^1C_4$ and ${}^4C_1$ pucker states in Dha suggests that chair-to-chair pucker-state interconversion is not thermodynamically favorable. Instead, interconversion between the ${}^4C_1$ chair and metastable states in the CP-equatorial region is expected. This is corroborated by unconstrained optimizations that are presented in **SI Section 4**, which show that more Dha pucker states in the CP-northern region collapse to metastable pucker states in the CP-equatorial region than the ${}^1C_4$ chair. For most other sugars, including the α -D-GalA variants, nearly all high energy CP-northern pucker states collapse to the ${}^1C_4$ chair, rather than the CP-equatorial region.

In contrast, Kdo exhibits relatively lower free energy pucker states across the CP-northern region for ϕ values between 300° and 100° , as well as lower energy metastable pucker states in the CP-equatorial region. This suggests that chair-to-chair pucker-state interconversion is likely more probable for Kdo than Dha. Indeed, during unconstrained optimizations, more CP-northern pucker states collapse to the ${}^1C_4$ chair for Kdo (**SI Figure 5.3b**) than in Dha (**SI Figure 5.3a**), which indicates that pucker-state interconversion may be energetically favorable for certain transition pathways.

Kdo exhibits significant thermodynamic stability across the CP-equatorial and CP-northern regions due to extensive hydrogen bonding networks. This is enabled by the (1,2)-hydroxyethyl group on C5 that can form up to two hydrogen bonds with exocyclic groups on adjacent ring carbons. For instance, a comparison of E_4 pucker state geometries reveals a hydrogen bond that is present in Kdo (**Figure 5.f**) and absent in Dha (**Figure 5.e**). Consequently, the relative free energy of E_4 is 8.47 kcal/mol for Kdo, compared to 14.20 kcal/mol for Dha. Thus, the range of thermodynamically accessible pucker states in Kdo may have substantial implications for side chain C flexibility, possibly influencing interactions with other carbohydrates via the terminal Rha residue attached to Kdo.

2.5 Apiose pucker energetics likely facilitate RG-II dimerization via conformational flexibility

Among the furanoses, apiose (1-MeO- β -D-Api) is particularly significant in the context of RG-II's structure as the site of the borate crosslink on side chain A to form the RG-II dimer in planta⁷. From a pucker energetics perspective, apiose presents the most contiguous low energy pucker states of all four furanose sugars in RG-II. Pucker state ${}^1E$ represents the global energy minimum with three contiguous pucker states, including ${}^1T_2$ (0.02 kcal/mol), E_2 (0.52 kcal/mol) and ${}^3T_2$ (0.64 kcal/mol), occurring within 1 kcal/mol of ${}^1E$. The relatively bulky ethoxy group on C3 appears to influence apiose conformational energetics. Low energy pucker states in quadrant III predominantly prefer axial orientation of the C1 methoxy group and quasi-axial orientation of the C3 ethoxy group.

Previous QM studies of boric acid-mediated cross-linking of apiose revealed the sequence of ring-pucker states traversed during the reaction mechanism, highlighting the dynamic conformational changes associated with borate ester formation²⁴. Apiose was found to initially adopt the ${}^4T_3$ pucker state when modeled in proximity to boric acid. During various stages of the esterification reaction, the pucker state varied between ${}^4T_3$, E_3 , and E_0 before adopting an E_0 pucker state in the final diborate complex²⁴. As shown in **Figure 6.a**, each transition pucker state along the reaction coordinate exhibits a <3 kcal/mol energy difference from the global minimum at ${}^1E$. Thus, even though the previously observed pucker states are different than the most stable pucker states identified in this study, the difference may be attributed to the interactions of apiose with

boric acid. We hypothesize that the underlying energetics facilitating multiple accessible pucker states may be key to the conformational flexibility required for facilitating RG-II dimerization at the apiose residue.

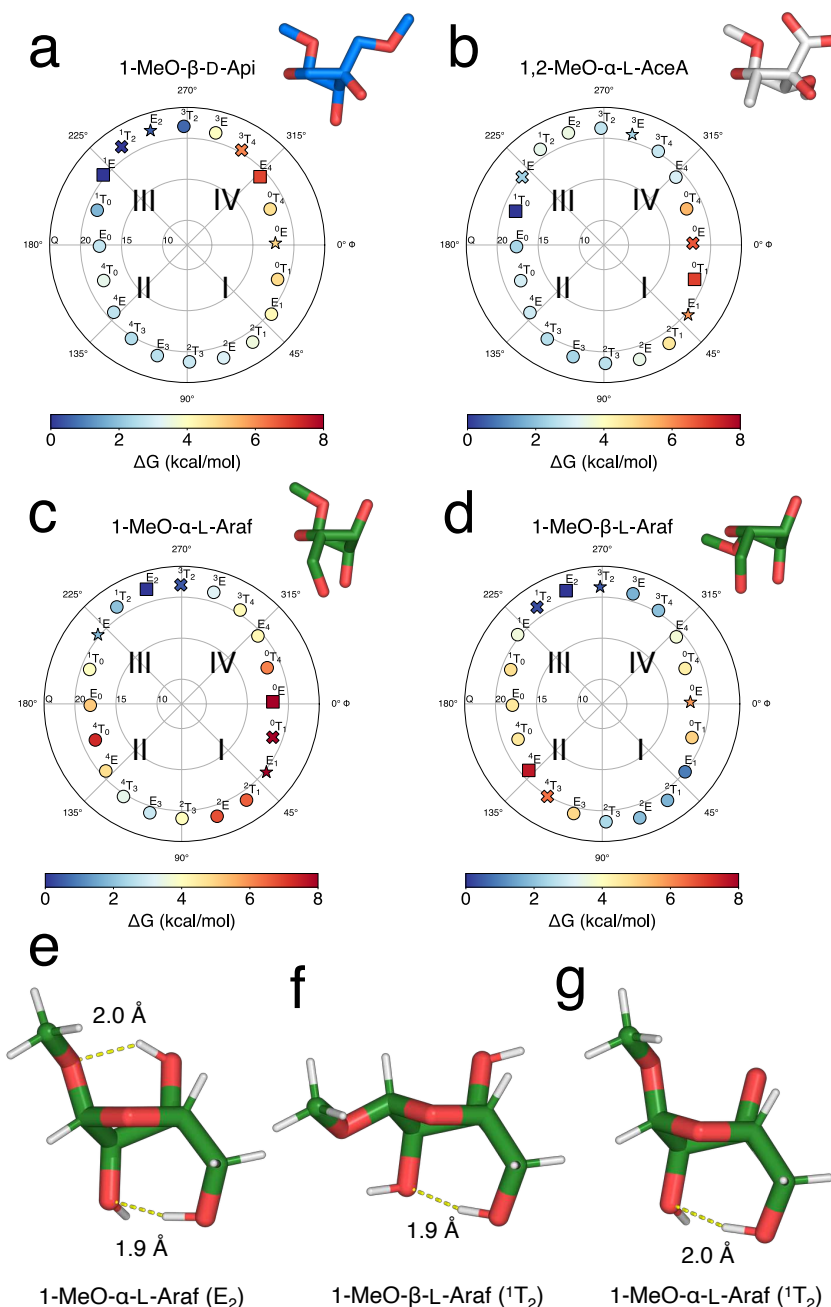


Figure 6. Hydrogen Bonding Networks Shape the Conformational Preferences of Furanoses in RG-II (a-d) Free energy landscapes for all furanoses in RG-II. The marker scheme is the same as described in Figure 3. (e) Hydrogen bonding networks for the lowest energy pucker state of α-L-Araf. (f) Hydrogen bonding for the lowest energy pucker state of β-L-Araf. (g) Hydrogen bonding networks in a low-energy pucker state of α-L-Araf that exhibits only one hydrogen bond.

2.6 Hydrogen bonds contribute to two distinct regions of thermodynamic stability in β -L-Araf and α -L-Araf

The highest energy barriers for any furanose sugar are in the range of 8 – 10 kcal/mol. Ring interconversion pathways therefore face a relatively small maximum thermodynamic barrier when compared to pyranoses. Hence, it is likely that a wider range of possible pucker states may be experimentally observed for the furanose sugars. Indeed, the energy landscapes of apiose (**Figure 6.a**) and aceric acid (**Figure 6.b**) are characterized by a single, extensive region of thermodynamically stable pucker states. In contrast, conformational energy landscapes for Araf each exhibit two distinct regions of thermodynamic stability (**Figure 6.c,d**), characterized by hydrogen bonding, with the flexible hydroxyethyl group on C4 acting as a hydrogen bond donor and the hydroxy group on C2 acting as a hydrogen bond acceptor (**Figure 6.f,g**). The most notable difference between the two free energy surfaces occurs for the ^0E pucker state, which has a higher relative energy in α -L-Araf (9.28 kcal/mol) than β -L-Araf (5.87 kcal/mol). This occurs because α -L-Araf can form a maximum of two hydrogen bonds due to its geometry (**Figure 6.e**), while β -L-Araf can form only one hydrogen bond. Furthermore, if a hydrogen bond is not possible for a given pucker state in β -L-Araf, it will exhibit a lower relative energy because the global energy minimum pucker state, E_2 , has only one hydrogen bond.

In contrast, the global energy minimum pucker state for α -L-Araf, E_2 , is stabilized by two hydrogen bonds, so a pucker state that does not have any hydrogen bonds will appear at a higher relative energy value since the minimum is more stabilized than in β -L-Araf. Thus, the high energy region encompassing ^0E , $^0T_1$, and E_1 is observed for α -L-Araf but not for β -L-Araf. These observations have important implications because α -L-Araf is a terminal residue on side chains E and F, and β -L-Araf is a terminal residue on side chain D. Understanding the energetics of the Araf species is important when considering that they are highly exposed to interactions with other biomolecules.

2.7 Pucker state changes affect overall RG-II structure

The influence of ring puckering extends beyond a single monosaccharide to affect the conformation of the polysaccharide. We examine the influence of pucker state geometries on RG-II's overall conformation by modifying the pucker states of two sugars in side chain A – apiose and a highly substituted rhamnose (1,2,3,4-MeO-Rha). Apiose was selected due to its importance for forming the RG-II dimer *in planta*, and 1,2,3,4-MeO-Rha was selected because it allows us to examine conformational changes resulting from a highly substituted sugar. The complete RG-II structure was reconstructed in CHARMM⁴⁵ using the QM-determined lowest energy pucker state for each sugar (**Figure 7.a,d**). A second low-energy pucker state was selected for apiose and 1,2,3,4-MeO-Rha, respectively, and side chain A was rebuilt using each of these new pucker states without any other structural modifications. Geometries for the rebuilt side chains are shown in **Figure 7.b,e**. The structural difference due to pucker modification of a single sugar is quantified by measuring the root-mean square deviation (RMSD) of side chain A.

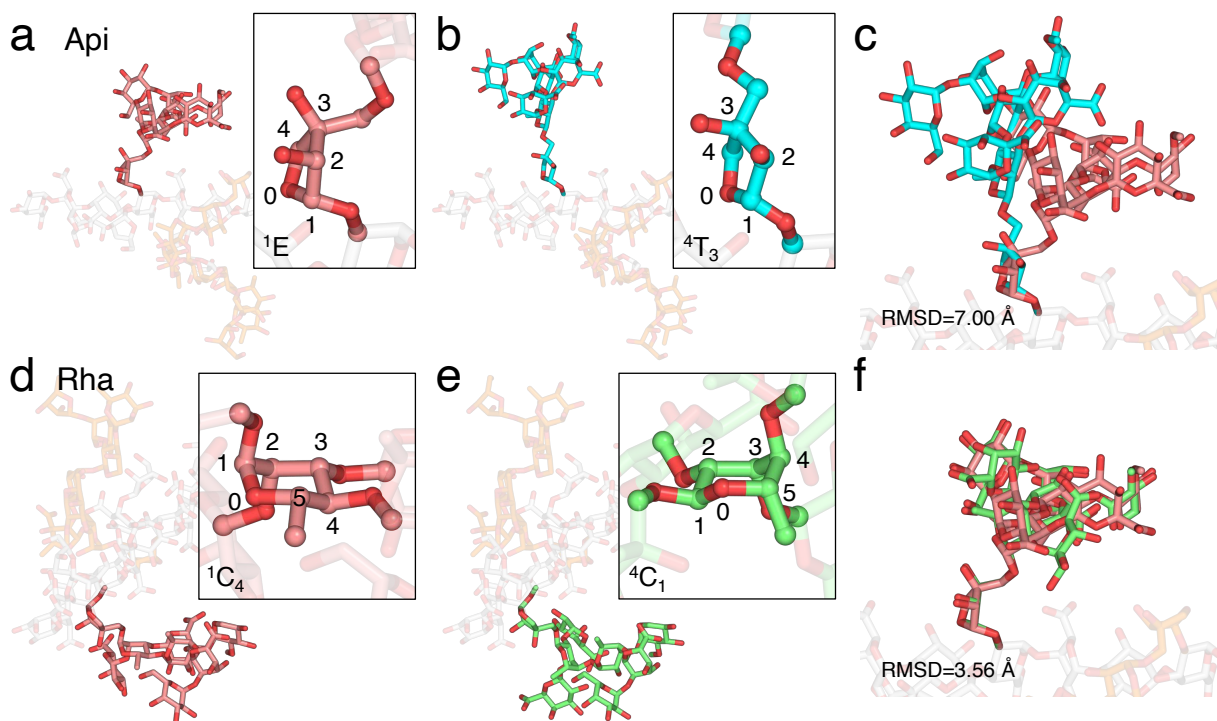


Figure 7. Pucker State Changes in Api and 1,2,3,4-MeO-Rha Influence RG-II Structure (a) Full RG-II structure with apiose in the ${}^1E$ pucker state, which was identified as the lowest energy conformation through QM simulations. (b) Apiose in the ${}^4T_3$ pucker state, which was identified in a previous study as the conformation required to undergo esterification to form the RG-II dimer²⁴. (c) Overlay of side chain A geometries shown in a,b. The RMSD value is calculated between these two variations of side chain A and not the entire RG-II molecule. (d) Full RG-II structure with 1,2,3,4-MeO-Rha in the ${}^1C_4$ pucker state identified as the lowest energy conformation through QM simulations. (e) 1,2,3,4-MeO-Rha in the ${}^4C_1$ pucker state. (f) Overlay of side chain A geometries shown in d,e.

The change in apiose pucker state significantly influences the conformation of side chain A. This is primarily caused by the orientation of the C3 glycosidic linkage, which changes from axial in the ${}^1E$ pucker state to equatorial in the ${}^4T_3$ pucker state. During this pucker change, the directions of the C1 and C3 linkages each go from approximately 10° to 120° . Consequently, side chain A is forced into a more parallel position relative to the ring in apiose, which results in the 7.00 Å RMSD (**Figure 7.c**).

Surprisingly, the change in 1,2,3,4-MeO-Rha pucker state influences the conformation of side chain A to a lesser extent. The RMSD between side chain A conformations is 3.56 Å, which is nearly half that observed for the apiose pucker state change (**Figure 7.f**). Although the C1 glycosidic linkage changes orientation from axial to equatorial, the rest of the side chain changes conformation minimally because the C4 glycosidic linkage simultaneously reorients from equatorial to axial. The approximately 90° angle between the C1 and C4 linkages is preserved when the pucker change occurs, which minimizes structural changes in the side chain.

These results indicate that pucker state changes for even a single sugar can have significant impacts on side chain structure. While these effects appear localized in our static reconstruction, a full MD study is necessary to determine how the overall ensemble of RG-II structures is impacted by changes in side chain geometry throughout a dynamic simulation. We find that in some cases, such as the apiose pucker change, large motion of other parts of the molecule is

required, which may kinetically slow down the transition. In the case of 1,2,3,4-MeO-Rha, a pucker change can be accommodated with very little modification to the structure of the rest of the molecule and likely occurs more frequently.

For sugars with multiple energetically favorable pucker states, including apiose and 1,2,3,4-MeO-Rha, considering multiple energetically favorable pucker states appear to more completely model carbohydrate structure, unlike traditional static reconstruction that relies on a single, low energy pucker state. In the context of RG-II, conformational changes are a key point of consideration when understanding its function in the plant cell wall. More generally, structural changes are also important when considering interactions between carbohydrates and proteins, such as during enzymatic cleavage of glycosidic linkages and immunological recognition.

3. Conclusions

Carbohydrates represent one of the most structurally diverse class of biomolecules. In glycobiology, the structural architecture of carbohydrates is fundamental to their physiological function(s). Even subtle variations in monosaccharide composition, type of glycosidic linkage, anomeric configuration, degree of polymerization, and ring pucker can dramatically alter a carbohydrate's physical properties, molecular recognition capabilities, and biological role. Given this tight coupling between structure and function, the comprehensive characterization of every structural facet—including monomer sequence, linkage position, and monomeric ring conformation—is essential.

In this study, we employ a systematic approach to elucidate pucker energetics for all monosaccharides in RG-II using high-level QM calculations. We then evaluate the impact of energetics on the overall structure and dynamics of RG-II. By probing the canonical pucker states for every uniquely substituted sugar within RG-II, we provide complete free energy atlases for twenty-one sugar structures. We find that using methoxy groups to approximate glycosidic linkages can significantly impact a sugar's free energy landscape, making it an important consideration for accurately modeling sugar structures in complex carbohydrates.

Furthermore, a multi-disciplinary approach was used to demonstrate the impact of ring substitutions on conformational dynamics for MeO-Arap. Our MD simulations bridge the QM energetic predictions with NMR experimental observations, revealing that ring interconversions of MeO-Arap on side chain B are responsible for the distinct NMR signatures of RG-II glycoforms.

Finally, we demonstrate that the impact of pucker energetics of monosaccharides extends beyond their own conformation towards governing the structure of larger polysaccharide constructs. Indeed, we found that reconstruction of the RG-II polysaccharide from the lowest energy sugar geometries demonstrated that changes in the pucker state of a single sugar can significantly impact side chain structure. More broadly, our findings provide a computational approach that is important to consider when studying structure-function relationships in other glycans and complex carbohydrates, with wide-ranging applications from development of bio-inspired polymeric materials to anticancer therapeutics.

4. Methods

Sugar Conformer Generation

All sugars were constructed using the RDKit package in Python⁴⁶. Canonical CP pucker states were generated using the Ring Analysis/Ring Reconstruction packages³³. A flowchart showing the sugar construction and subsequent optimization process is depicted in **Figure 1.e**. First, a base furanose ring was generated from a SMILES string and passed to the Ring Reconstruction package, which returned Cartesian ring atom coordinates for the desired pucker state. The sugar of interest was then produced from a second SMILES string. The SMILES string was constructed such that each ring carbon participating in a glycosidic bond was functionalized with a methoxy group as opposed to the hydroxyl group that would be attached to the carbon in a canonical model of the sugar. For sugars containing at least one carboxylic acid group, each carboxylic acid group was modeled in the protonated state.

Methoxy groups were included to account for the steric effects from neighboring sugars that are present in RG-II, which would be ignored using the canonical sugar structure²⁴. Fifty conformers of the sugar were generated using RDKit's conformer generation toolkit, and the conformer with the lowest root mean square deviation (RMSD) of its ring atom coordinates from the base ring was selected. This step was conducted to preserve the stereochemistry of each sugar during the following Cartesian transformations. Coordinates for each ring atom in the chosen sugar conformer were set equal to coordinates of the corresponding ring atoms in the base molecule to ensure that the desired sugar aligned directly with the pucker coordinates generated by the Ring Reconstruction package. A force field (MMFF94) was applied to the sugar while maintaining fixed positions of its ring atoms to preserve the initial pucker state⁴⁷. The Q value for each pucker state was fixed at 0.36 radians throughout all constrained optimizations for furanoses to agree closely with optimized values observed in previous literature^{23,33}.

Geometry Optimizations and Frequency Calculations

Gaussian 16 was used for all QM geometry optimizations and frequency calculations. Water solvation effects were included in all calculations *via* the polarizable continuum model with the integral equation formalism (IEFPCM). Puckering parameters were measured pre- and post-optimization using the pucker method in Amber CPPTRAJ⁴⁸. The initial geometry of each pucker state was optimized using DFT at the B3LYP level of theory with a 6-31+G(d,p) basis set to provide a low-energy reference point for further optimizations^{49–52}. Two dihedral angles in the ring were fixed to maintain the pucker state during optimization for furanose sugars and three were fixed for pyranose sugars.

Dihedral scans were performed on all major rotatable exocyclic bonds at the B3LYP level of theory with a 6-31+G(d,p) basis set to determine the global energy minimum as recommended by previous DFT benchmarking studies⁵³. Phenolic hydroxyl groups were exempt from manual rotation due to their relatively simple geometry and to reduce computational cost. Dihedral scans were started at the initial optimized geometry, and geometry optimizations were separately conducted for each exocyclic group in 15° increments^{24,38}. The geometry representing the lowest single-point electronic energy across all exocyclic groups was selected as a starting point for the next dihedral scan. Dihedral scans starting at the lowest energy geometry identified during the prior scan were repeated until no lower electronic energy states were identified. Ring atom

dihedral angles were fixed (as described for the initial optimization) during all scans to maintain the initial pucker state.

A final geometry optimization and frequency analysis was conducted using second-order Møller-Plesset perturbation theory (MP2) and cc-PVDZ basis set with tight convergence parameter^{54–61}. The final optimization was said to represent the global energy minimum geometry for each pucker state.

A second set of fully unconstrained optimizations was conducted using the same calculation methods as the final optimization but with all ring dihedral constraints lifted. These calculations were performed to assess the accuracy of the constrained optimizations and are presented in the Supplemental Information (SI) Section 4. We observe that the unconstrained optimizations validate our extrapolated free energy surfaces for each sugar since all high-energy pucker states optimize to accessible local energy minima in the CP-equatorial or CP-polar regions.

We also generated minimum energy transition pathway profiles for pucker interconversions for all sugars by identifying the lowest energy pucker states in the CP-northern, CP-equatorial, and CP-southern regions. These results are presented in **SI Section 6**.

MD Simulations

Two L-Arap models were generated to represent the fragmental structure of side chain B of native RG-II from celery or papaya (Fig. 4C, 4E): (i) a single-substitution model containing D-Gal and L-Arap, and (ii) a triple-substitution model containing D-Gal, L-Arap, and two Rha residues linked sequentially. Sugar residues were constructed sequentially by covalent attachment to L-Arap, rotation about the glycosidic linkage to achieve sterically and energetically favorable conformations, and energy minimization to yield a relaxed starting geometry.

Systems were solvated in cubic water boxes (side length 39.4 Å) containing ~6,100 atoms. Initial equilibration was performed for 200 ps at 300 K in the NPT ensemble, followed by production simulations of 1.5 μs at 300 K, 400 K, and 500 K in the NVT ensemble. All simulations were conducted using CHARMM program^{62,63} with the all36_carb carbohydrate force field, employing DOMDEC method⁶⁴ for parallel dynamics. A 2 fs integration timestep was used with SHAKE⁶⁵ to constrain covalent bonds involving hydrogen atoms. Periodic boundary conditions were applied, with a 12 Å cutoff for nonbonded interactions and the particle mesh Ewald (PME) method⁶⁶ for electrostatics.

NMR Data Collection

HSQC NMR spectra were acquired on a Bruker Neo spectrometer operating at a frequency of 900 MHz equipped with a 5 mm cryogenically cooled probe. Prior to analysis, monomeric RG-II samples (9–10 mg) were subjected to a single D₂O exchange by dissolution in 1 mL of 99.9% D₂O (Cambridge Isotope Laboratories, Andover, MA, USA), followed by lyophilization. The freeze-dried samples were subsequently re-dissolved in 0.6 mL of D₂O and transferred to 5 mm NMR tubes for spectral acquisition. HSQC spectra were recorded at three different temperatures (11 °C, 25 °C, and 60 °C) to evaluate temperature-dependent conformational behavior of the Arap residues. All NMR data were processed and analyzed using MestReNova software (Mestrelab Research S.L., Santiago de Compostela, Spain).

We also conducted QM calculations to corroborate the experimentally observed chemical shifts directions that signify pucker interconversions between the chair conformations. See **SI Section 7** for QM GIAO methodology details.

RG-II 3-D Structure Generation in CHARMM from Optimized Sugars

The three-dimensional structure of the complete RG-II molecule can be constructed from the sugars in any pucker state. The structure generated from the lowest-energy pucker states may represent the most likely structure of RG-II, which is the most probable among an ensemble of structures found at biologically relevant temperatures. We built up the complete RG-II structure by adding one sugar at a time starting with the GalA backbone 10-mer. Each sugar was bonded to the existing structure, and, while fixing the pucker state of the newly added sugar ring and the rest of the existing molecule, the geometry of the new sugar was adjusted to have optimal glycosidic linkage bond lengths and angles. The newly added sugar was then rotated around the glycosidic bond attaching it to its neighboring sugar to reduce van der Waals and electrostatic conflicts using the CHARMM carbohydrate force field^{38–40,67}.

Once the bonds, angles, and obvious conflicts were resolved, the molecular structure with the newly added sugar was energy-minimized, keeping the pucker states fixed. This process was followed by adding one sugar at a time to generate the backbone and then each of the side chains in the following order: E, F, A, B, C, D. The resultant structure is a valid representation of the possible structure of RG-II, but only one member of a large ensemble of possible low-energy conformations. A deeper analysis of the structure is beyond the scope of this study but will be forthcoming.

To compare the effects of pucker changes in the apiose and rhamnose of side chain A, the first structure constructed was of sugars all in their QM-lowest-energy pucker states. A structure for the change in apiose pucker was generated by deleting the apiose in side chain A and then adding the apiose back in by the method above using the alternate pucker state. This process involved two steps. First, the apiose was added to the backbone by the method above and rotated to position it as close to the previous orientation as possible. Second, the rest of the side chain A was bonded to the apiose and re-oriented to position the bond lengths and angles correctly, after which all bonds in side chain A were rotated to orient as close to the previous structure as possible. The closest structure with the new pucker state for apiose had a side chain A RMSD of about 7 Angstroms. The same process was followed for the alternate pucker state of rhamnose, except that there were two GalAs also bonded to the rhamnose that needed to be reoriented after pucker state change.

Acknowledgement of Funding

This work was authored in part by Alliance for Sustainable Energy, LLC, the manager and operator of the National Laboratory of the Rockies for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. The authors acknowledge the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Sciences of the United States Department of Energy through Grant DESC0008472 for funding structural studies of RG-II. The authors are also grateful for the access and use of NLR's Computational Sciences' resources (Kestrel) supported by the DOE Office of EERE under Contract Number DE-AC36-08GO28308. Sean Brooks acknowledges funding support from the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internship (SULI) program.

Data Availability

The code to generate the sugar pucker conformations from SMILES strings is available at:

<https://github.com/vbharadwcomosci/sucerQM>

All optimized structures generated in this study are also available in the final_optimizations folder of the Github repository:

https://github.com/vbharadwcomosci/sucerQM/tree/main/final_optimizations

Author Contributions

V.S.B. conceptualized and designed the study, S.B. performed QM Calculations, L.B. performed MD simulations, D.K. performed NMR experiments, and M.F.C. built the RG-II structure using CHARMM. S.B. and V.S.B. performed the data analysis and interpretation, wrote, and edited the manuscript with assistance from all authors. B.U. assisted with writing and editing. Visualizations were generated by S.B.

Author Information

Corresponding Author
vivek.bharadwaj@nlr.gov

Ethics approval and consent to participate: Not applicable

Competing Interest: The authors declare no competing interests

References

- ¹ Michael M. Cox and David L. Nelson, *Lehninger Principles of Biochemistry*, 6th ed. (Macmillan Learning, 2012).
- ² R.D. Cummings, “A periodic table of monosaccharides,” *Glycobiology* **34**(1), cwad088 (2024).
- ³ D. Cremer, and J.A. Pople, “A General Definition of Ring Puckering Coordinates,” *Journal of the American Chemical Society*, (1975).
- ⁴ T. Matsunaga, T. Ishii, S. Matsumoto, M. Higuchi, A. Darvill, P. Albersheim, and M.A. O’Neill, “Occurrence of the Primary Cell Wall Polysaccharide Rhamnogalacturonan II in Pteridophytes, Lycophytes, and Bryophytes. Implications for the Evolution of Vascular Plants,” *Plant Physiology* **134**(1), 339–351 (2004).
- ⁵ Y. Oiwa, K. Kitayama, M. Kobayashi, and T. Matoh, “Boron deprivation immediately causes cell death in growing roots of *Arabidopsis thaliana* (L.) Heynh.,” *Soil Science and Plant Nutrition* **59**(4), 621–627 (2013).
- ⁶ M. Kobayashi, T. Matoh, and Ji. Azuma, “Two Chains of Rhamnogalacturonan II Are Cross-Linked by Borate-Diol Ester Bonds in Higher Plant Cell Walls,” *Plant Physiol.* **110**(3), 1017–1020 (1996).
- ⁷ T. Koshiba, M. Kobayashi, and T. Matoh, “Boron Nutrition of Tobacco BY-2 Cells. V. Oxidative Damage is the Major Cause of Cell Death Induced by Boron Deprivation,” *Plant and Cell Physiology* **50**(1), 26–36 (2009).
- ⁸ T. Koshiba, M. Kobayashi, K. Matsuoka, T. Fujiwara, and T. Matoh, “Boron nutrition of cultured tobacco BY-2 cells. VII. Rapid induction of metabolic dysfunction in cells deprived of boron as revealed by microarray analysis,” *Soil Science and Plant Nutrition* **59**(2), 189–194 (2013).
- ⁹ J.L. De Meirelles, F.C. Nepomuceno, J. Peña-García, R.R. Schmidt, H. Pérez-Sánchez, and H. Verli, “Current Status of Carbohydrates Information in the Protein Data Bank,” *J. Chem. Inf. Model.* **60**(2), 684–699 (2020).
- ¹⁰ J.Ø. Duus, C.H. Gotfredsen, and K. Bock, “Carbohydrate Structural Determination by NMR Spectroscopy: Modern Methods and Limitations,” *Chem. Rev.* **100**(12), 4589–4614 (2000).
- ¹¹ C. J N, O. Guvench, A.D. MacKerell Jr, T. Yamaguchi, and S.S. Mallajosyula, “Refinement of the Drude Polarizable Force Field for Hexose Monosaccharides: Capturing Ring Conformational Dynamics with Enhanced Accuracy,” *J. Chem. Theory Comput.* **20**(20), 9161–9177 (2024).
- ¹² S. Neelamegham, K. Aoki-Kinoshita, E. Bolton, M. Frank, F. Lisacek, T. Lütkeke, N. O’Boyle, N.H. Packer, P. Stanley, P. Toukach, A. Varki, R.J. Woods, The SNFG Discussion Group, A. Darvill, A. Dell, B. Henrissat, C. Bertozzi, G. Hart, H. Narimatsu, H. Freeze, I. Yamada, J. Paulson, J. Prestegard, J. Marth, J.F.G. Vliegthart, M. Etzler, M. Aebi, M. Kanehisa, N. Taniguchi, N. Edwards, P. Rudd, P. Seeberger, R. Mazumder, R. Ranzinger, R. Cummings, R. Schnaar, S. Perez, S. Kornfeld, T. Kinoshita, W. York, and Y. Knirel, “Updates to the Symbol Nomenclature for Glycans guidelines,” *Glycobiology* **29**(9), 620–624 (2019).
- ¹³ K. Gaweda, A. Plazinska, and W. Plazinski, “Conformations of saturated five-membered heterocycles evaluated by MP2 calculations,” *Chem Heterocycl Comp* **56**(12), 1599–1604 (2020).
- ¹⁴ K. Nester, and W. Plazinski, “Deciphering the conformational preferences of furanosides. A molecular dynamics study,” *Journal of Biomolecular Structure and Dynamics* **38**(11), 3359–3370 (2020).
- ¹⁵ D. Walczak, A. Sikorski, D. Grzywacz, A. Nowacki, and B. Liberek, “Identification of the furanose ring conformations and the factors driving their adoption,” *Carbohydrate Research* **526**, 108780 (2023).
- ¹⁶ H.B. Mayes, L.J. Broadbelt, and G.T. Beckham, “How Sugars Pucker: Electronic Structure Calculations Map the Kinetic Landscape of Five Biologically Paramount Monosaccharides and Their Implications for Enzymatic Catalysis,” *J. Am. Chem. Soc.* **136**(3), 1008–1022 (2014).

- ¹⁷ I. Alibay, and R.A. Bryce, “Ring Puckering Landscapes of Glycosaminoglycan-Related Monosaccharides from Molecular Dynamics Simulations,” *J. Chem. Inf. Model.* **59**(11), 4729–4741 (2019).
- ¹⁸ J. Cybulska, A. Brzyska, A. Zdunek, and K. Woliński, “Simulation of Force Spectroscopy Experiments on Galacturonic Acid Oligomers,” *PLoS ONE* **9**(9), e107896 (2014).
- ¹⁹ M.A. O’Neill, D. Warrenfeltz, K. Kates, P. Pellerin, T. Doco, A.G. Darvill, and P. Albersheim, “Rhamnogalacturonan-II, a Pectic Polysaccharide in the Walls of Growing Plant Cell, Forms a Dimer That Is Covalently Cross-linked by a Borate Ester,” *Journal of Biological Chemistry* **271**(37), 22923–22930 (1996).
- ²⁰ T.T. Stevenson, A.G. Darvill, and P. Albersheim, “3-deoxy-d-lyxo-2-heptulosaric acid, a component of the plant cell-wall polysaccharide rhamnogalacturonan-II,” *Carbohydrate Research* **179**, 269–288 (1988).
- ²¹ W. York, and A.G. Darvill, “3-deoxy-d-manno-2-octulosonic acid (KDO) is a component of rhamnogalacturonan II, a pectic polysaccharide in the primary cell walls of plants,” (n.d.).
- ²² V. Bharadwaj, S. Koj, L. Bu, M. Crowley, D. Sharma, B. Urbanowicz, M. O’Neill, and M. Pena, “Probing Sequence-Structure Paradigms in Complex Carbohydrates – A Case Study on Rhamnogalacturonan-II,” (2024).
- ²³ K. Mazeau, and S. Pérez, “The preferred conformations of the four oligomeric fragments of Rhamnogalacturonan II,” *Carbohydrate Research* **311**(4), 203–217 (1998).
- ²⁴ V.S. Bharadwaj, M.F. Crowley, M.J. Peña, B. Urbanowicz, and M. O’Neill, “Mechanism and Reaction Energy Landscape for Apiose Cross-Linking by Boric Acid in Rhamnogalacturonan II,” *J. Phys. Chem. B* **124**(45), 10117–10125 (2020).
- ²⁵ V.S. Bharadwaj, L.P. Westawker, and M.F. Crowley, “Towards Elucidating Structure–Spectra Relationships in Rhamnogalacturonan II: Computational Protocols for Accurate ¹³C and ¹H Shifts for Apiose and Its Borate Esters,” *Front. Mol. Biosci.* **8**, 756219 (2022).
- ²⁶ C. J. N., and S.S. Mallajosyula, “Impact of Polarization on the Ring Puckering Dynamics of Hexose Monosaccharides,” *J. Chem. Inf. Model.* **63**(1), 208–223 (2023).
- ²⁷ G.J. Davies, A. Planas, and C. Rovira, “Conformational Analyses of the Reaction Coordinate of Glycosidases,” *Acc. Chem. Res.* **45**(2), 308–316 (2012).
- ²⁸ K. Gaweda, and W. Plazinski, “Tautomeric and epimeric equilibria of aldo- and ketohexoses studied by the MD simulations and QM calculations,” *Carbohydrate Research* **474**, 8–15 (2019).
- ²⁹ O. Guvench, D. Martin, and M. Greene, “Pyranose Ring Puckering Thermodynamics for Glycan Monosaccharides Associated with Vertebrate Proteins,” *International Journal of Molecular Sciences* **23**(1), 473 (2022).
- ³⁰ E. Autieri, M. Sega, F. Pederiva, and G. Guella, “Puckering free energy of pyranoses: A NMR and metadynamics-umbrella sampling investigation,” *The Journal of Chemical Physics* **133**(9), 095104 (2010).
- ³¹ P. Zhang, L. Shen, and W. Yang, “Solvation Free Energy Calculations with Quantum Mechanics/Molecular Mechanics and Machine Learning Models,” *J. Phys. Chem. B* **123**(4), 901–908 (2019).
- ³² J. Rihon, S. Reynders, V. Bernardes Pinheiro, and E. Lescrinier, “The pucker.rs toolkit to facilitate sampling the conformational space of biomolecular monomers,” *Journal of Cheminformatics* **17**(1), 53 (2025).
- ³³ L. Chan, G.R. Hutchison, and G.M. Morris, “Understanding Ring Puckering in Small Molecules and Cyclic Peptides,” *J. Chem. Inf. Model.* **61**(2), 743–755 (2021).
- ³⁴ A. Tankrathok, J. Iglesias-Fernández, R.J. Williams, S. Pengthaisong, S. Baiya, Z. Hakki, R.C. Robinson, M. Hrmova, C. Rovira, S.J. Williams, and J.R. Ketudat Cairns, “A Single Glycosidase

Harnesses Different Pyranoside Ring Transition State Conformations for Hydrolysis of Mannosides and Glucosides,” *ACS Catal.* **5**(10), 6041–6051 (2015).

³⁵ G. Speciale, A.J. Thompson, G.J. Davies, and S.J. Williams, “Dissecting conformational contributions to glycosidase catalysis and inhibition,” *Current Opinion in Structural Biology* **28**, 1–13 (2014).

³⁶ V.S. Bharadwaj, B.C. Knott, J. Ståhlberg, G.T. Beckham, and M.F. Crowley, “The hydrolysis mechanism of a GH45 cellulase and its potential relation to lytic transglycosylase and expansin function,” *Journal of Biological Chemistry* **295**(14), 4477–4487 (2020).

³⁷ A. Kaczmarek, P.M. Pieczywek, J. Cybulska, and A. Zdunek, “Structure and functionality of Rhamnogalacturonan I in the cell wall and in solution: A review,” *Carbohydrate Polymers* **278**, 118909 (2022).

³⁸ E.P. Raman, O. Guvench, and A.D. MacKerell, “CHARMM Additive All-Atom Force Field for Glycosidic Linkages in Carbohydrates Involving Furanoses,” *J. Phys. Chem. B* **114**(40), 12981–12994 (2010).

³⁹ O. Guvench, E. Hatcher, R.M. Venable, R.W. Pastor, and A.D. MacKerell, “CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses,” *J. Chem. Theory Comput.* **5**(9), 2353–2370 (2009).

⁴⁰ O. Guvench, S.S. Mallajosyula, E.P. Raman, E. Hatcher, K. Vanommeslaeghe, T.J. Foster, F.W. Jamison, and A.D. MacKerell, “CHARMM Additive All-Atom Force Field for Carbohydrate Derivatives and Its Utility in Polysaccharide and Carbohydrate–Protein Modeling,” *J. Chem. Theory Comput.* **7**(10), 3162–3180 (2011).

⁴¹ H.L. Woodcock, D. Moran, R.W. Pastor, A.D. MacKerell, and B.R. Brooks, “Ab Initio Modeling of Glycosyl Torsions and Anomeric Effects in a Model Carbohydrate: 2-Ethoxy Tetrahydropyran,” *Biophysical Journal* **93**(1), 1–10 (2007).

⁴² R.S. Bitzer, A.G.H. Barbosa, C.O. Da Silva, and M.A.C. Nascimento, “On the generalized valence bond description of the anomeric and exo-anomeric effects: an ab initio conformational study of 2-methoxytetrahydropyran,” *Carbohydrate Research* **340**(13), 2171–2184 (2005).

⁴³ M.A. O’Neill, I. Black, B. Urbanowicz, V. Bharadwaj, M. Crowley, S. Koj, and M.J. Peña, “Locating Methyl-Etherified and Methyl-Esterified Uronic Acids in the Plant Cell Wall Pectic Polysaccharide Rhamnogalacturonan II,” *SLAS Technology* **25**(4), 329–344 (2020).

⁴⁴ J.N. Glushka, M. Terrell, W.S. York, M.A. O’Neill, A. Gucwa, A.G. Darvill, P. Albersheim, and J.H. Prestegard, “Primary structure of the 2-O-methyl- -L-fucose-containing side chain of the pectic polysaccharide, rhamnogalacturonan II,” (n.d.).

⁴⁵ B.R. Brooks, C.L.B. Iii, A.D.M. Jr, L. Nilsson, R.J. Petrella, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, and M. Karplus, “CHARMM: The Biomolecular Simulation Program,” (2010).

⁴⁶ Greg Landrum, “An overview of the RDKit — The RDKit 2025.03.6 documentation,” (2025).

⁴⁷ T.A. Halgren, “Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94,” *J. Comput. Chem.* **17**(5–6), 490–519 (1996).

⁴⁸ D.R. Roe, and T.E. Cheatham, “PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data,” *J. Chem. Theory Comput.* **9**(7), 3084–3095 (2013).

⁴⁹ C. Lee, W. Yang, and R.G. Parr, “Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density,” *Phys. Rev. B* **37**(2), 785–789 (1988).

⁵⁰ A.D. Becke, “Density-functional thermochemistry. III. The role of exact exchange,” *The Journal of Chemical Physics* **98**(7), 5648–5652 (1993).

- ⁵¹ P.J. Stephens, F.J. Devlin, C.F. Chabalowski, and M.J. Frisch, "Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields," *J. Phys. Chem.* **98**(45), 11623–11627 (1994).
- ⁵² W.J. Hehre, R. Ditchfield, and J.A. Pople, "Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules," *The Journal of Chemical Physics* **56**(5), 2257–2261 (1972).
- ⁵³ G.I. Csonka, "Proper basis set for quantum mechanical studies of potential energy surfaces of carbohydrates," *Journal of Molecular Structure: THEOCHEM* **584**(1–3), 1–4 (2002).
- ⁵⁴ Chr. Møller, and M.S. Plesset, "Note on an Approximation Treatment for Many-Electron Systems," *Phys. Rev.* **46**(7), 618–622 (1934).
- ⁵⁵ M.J. Frisch, M. Head-Gordon, and J.A. Pople, "A direct MP2 gradient method," *Chemical Physics Letters* **166**(3), 275–280 (1990).
- ⁵⁶ M.J. Frisch, M. Head-Gordon, and J.A. Pople, "Semi-direct algorithms for the MP2 energy and gradient," *Chemical Physics Letters* **166**(3), 281–289 (1990).
- ⁵⁷ M. Head-Gordon, J.A. Pople, and M.J. Frisch, "MP2 energy evaluation by direct methods," *Chemical Physics Letters* **153**(6), 503–506 (1988).
- ⁵⁸ S. Sæbø, and J. Almlöf, "Avoiding the integral storage bottleneck in LCAO calculations of electron correlation," *Chemical Physics Letters* **154**(1), 83–89 (1989).
- ⁵⁹ M. Head-Gordon, and T. Head-Gordon, "Analytic MP2 frequencies without fifth-order storage. Theory and application to bifurcated hydrogen bonds in the water hexamer," *Chemical Physics Letters* **220**(1–2), 122–128 (1994).
- ⁶⁰ A.D. McLean, and G.S. Chandler, "Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, $Z=11-18$," *The Journal of Chemical Physics* **72**(10), 5639–5648 (1980).
- ⁶¹ R. Krishnan, J.S. Binkley, R. Seeger, and J.A. Pople, "Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions," *The Journal of Chemical Physics* **72**(1), 650–654 (1980).
- ⁶² Alexander D. MacKerell, Bernard R. Brooks, Charles Brooks, Lennart Nilsson, Benoit Roux, Youngdo Won, and Martin Karplus, "CHARMM: The Energy Function and Its Parameterization," *Encyclopedia of Computational Chemistry*, (1998).
- ⁶³ W. Hwang, S.L. Austin, A. Blondel, E.D. Boittier, S. Boresch, M. Buck, J. Buckner, A. Caflisch, H.-T. Chang, X. Cheng, Y.K. Choi, J.-W. Chu, M.F. Crowley, Q. Cui, A. Damjanovic, Y. Deng, M. Devereux, X. Ding, M.F. Feig, J. Gao, D.R. Glowacki, J.E. Gonzales, M.B. Hamaneh, E.D. Harder, R.L. Hayes, J. Huang, Y. Huang, P.S. Hudson, W. Im, S.M. Islam, W. Jiang, M.R. Jones, S. Käser, F.L. Kearns, N.R. Kern, J.B. Klauda, T. Lazaridis, J. Lee, J.A. Lemkul, X. Liu, Y. Luo, A.D. MacKerell, D.T. Major, M. Meuwly, K. Nam, L. Nilsson, V. Ovchinnikov, E. Paci, S. Park, R.W. Pastor, A.R. Pittman, C.B. Post, S. Prasad, J. Pu, Y. Qi, T. Rathinavelan, D.R. Roe, B. Roux, C.N. Rowley, J. Shen, A.C. Simmonett, A.J. Sodt, K. Töpfer, M. Upadhyay, A. Van Der Vaart, L.I. Vazquez-Salazar, R.M. Venable, L.C. Warrensford, H.L. Woodcock, Y. Wu, C.L. Brooks, B.R. Brooks, and M. Karplus, "CHARMM at 45: Enhancements in Accessibility, Functionality, and Speed," *J. Phys. Chem. B* **128**(41), 9976–10042 (2024).
- ⁶⁴ A. Hynninen, and M.F. Crowley, "New faster CHARMM molecular dynamics engine," *J. Comput. Chem.* **35**(5), 406–413 (2014).
- ⁶⁵ V. Kräutler, W.F. Van Gunsteren, and P.H. Hünenberger, "A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations," *J. Comput. Chem.* **22**(5), 501–508 (2001).
- ⁶⁶ T. Darden, D. York, and L. Pedersen, "Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems," *The Journal of Chemical Physics* **98**(12), 10089–10092 (1993).

⁶⁷ K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A.D. Mackerell, "CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields," *J Comput Chem* **31**(4), 671–690 (2010).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Sl.pdf](#)