

Integrated Experimental and Theoretical Studies of Saccharide Structure

Thomas Ethan Klepach

Publication Date

22-07-2008

License

This work is made available under a All Rights Reserved license and should only be used in accordance with that license.

Citation for this work (American Psychological Association 7th edition)

Klepach, T. E. (2008). *Integrated Experimental and Theoretical Studies of Saccharide Structure* (Version 1). University of Notre Dame. <https://doi.org/10.7274/v405s754q1d>

This work was downloaded from CurateND, the University of Notre Dame's institutional repository.

For more information about this work, to report or an issue, or to preserve and share your original work, please contact the CurateND team for assistance at curate@nd.edu.

INTEGRATED EXPERIMENTAL AND THEORETICAL STUDIES OF
SACCHARIDE STRUCTURE

VOLUME II

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by

Thomas E. Klepach

Anthony S. Serianni, Director

Graduate Program in Chemistry and Biochemistry

Notre Dame, Indiana

July 2008

CONTENTS

Volume II

Figures	ix
Tables	xxiv
Schemes	xxviii
Chapter 7: Hydroxyl Group Configuration / Conformation and $^2J_{\text{CCC}}$ Spin-Couplings in Saccharides: DFT, NJC and NSA Analyses of Experimental Data	
7.1. Abstract	272
7.1.1. Introduction.....	273
7.2. Experimental.....	277
7.2.1. Experimental ^{13}C - ^{13}C Spin-Coupling Measurements in DMSO	277
7.3. NMR titration of experimental $^2J_{\text{C1,C3}}$ by ionization of the O2 hydroxyl	278
7.4. Computational	278
7.4.1. Selection and Geometric Optimization of Model Compounds.....	278
7.4.2. Theoretical Calculations of ^{13}C - ^{13}C Spin-Coupling Constants	282
7.4.3. Natural Bonding Orbital Calculations	282
7.4.4. Theoretical SSCC Parameterization and Analysis of Experimental Couplings.....	282
7.5. Results and Discussion.....	283
7.5.1. The Dependence of $^2J_{\text{C1,C3}}$ on Configuration.....	283
7.5.2. The Dependence of $^2J_{\text{C1,C3}}$ on Conformation.....	284
7.5.3. The Effect of Deoxygenation on $^2J_{\text{CC}}$	291
7.5.4. The Effect of <i>N</i> -acetylation on $^2J_{\text{CC}}$	296
7.5.5. C-C Bond Angle and Bond Length Considerations Regarding $^2J_{\text{CC}}$	297
7.5.6. NBO Analysis: Lone Pair Donation into C-C σ^* Orbitals	301

7.5.7. The Influence of C-O rotation on $^2J_{C2,C4}$ in β -Glc.....	310
7.5.8. Natural J -coupling Analysis I: Configurational Dependence of $^2J_{C1,C3}$	314
7.5.9. Natural J -coupling Analysis II: Conformational Dependence of $^2J_{C1,C3}$	323
7.5.10. Coupling Trend Validation in Conformationally Determined Lacotside System	337
7.5.11. O2 Ionization and $^2J_{C1,C3}$	338
7.5.12. The Conformational Analysis of ϕ , ξ and ζ in Methyl- β -D- glucopyranose	339
7.6. Conclusions	399
7.7. References	406
Chapter 8: Crystal Structure Methyl β -allolactoside (Methyl 6-O-beta-D-[1- ^{13}C]- galactopyranosyl-beta-D-glucopyranoside) Monohydrate.....	415
8.1. Introduction	415
8.2. Comment	416
8.3. Experimental.....	426
8.3.1. Reagents.....	426
8.3.2. Synthesis of o-Nitrophenyl β -D-[1- ^{13}C]-galactopyranoside (2).....	427
8.3.3. Synthesis of Disaccharide I_E via Enzyme-catalyzed Transglycosylation	427
8.3.4. Crystal growth of I_E	428
8.3.5. Data Collection.....	429
8.3.6. Instrumentation	430
8.3.7. Refinement.....	430
8.3.8. Crystal Data.....	430
8.4. References	431
Chapter 9: Synthesis.....	434
9.1. Definition of Terms.....	434
9.2. Primary Structure	436
9.3. Secondary Structure	437
9.4. Higher Order Structure.....	438
9.5. Modular Conformational Analysis.....	442
9.6. Biological Applications.....	444
9.7. References	446

FIGURES

- Figure 7.1. Conformational Dependence of $^2J_{C1,C3}$ on ξ . Definitions of ϕ , ξ and ζ are found in the text. Specific values for ϕ and ζ are represented by the coloration indicated in the legend.286
- Figure 7.2. Comparison of Spin-Spin Couplings Calculated *In Vacuo* and *In Solvo* as a Function of ξ in 2^C . Data depicted with circles and squares are structures in which C2-C3-O3-H (ζ) = 60° and 180° respectively. Solid and outlined shapes were data calculated *In Vacuo* and *In Solvo* respectively. Definition of ξ in the text.....290
- Figure 7.3. Projection Resultant for $^2J_{C1,C3}$ as a Function of C1, C2 and C3 Hydroxyl Substitution. The projection resultant is a unitless measure. C1 through C3 hydroxyl substitutions patterns are indicated by color where ‘Exptl.’ denotes an experimentally measured coupling and ‘DFT’ an averaged ab initio value. C1 through C3 hydroxyls not mentioned are absent (i.e. O1, O3 indicates a 2-deoxy structure). Data is not sorted by configuration.295
- Figure 7.4. Comparison of Experimental and DFT $^2J_{C1,C3}$ Values in **14** - **17**. Diamonds represent *ab initio* coupling values from the Series 6 calculations, whereas the horizontal lines are experimental couplings. Color representations are blue: b-GlcNAc (**15**), green: a-GlcNAc (**14**), orange: b-AlloNAc (**17**), and red: a-AlloNAc (**16**).....297
- Figure 7.5. The Subtended Bond Angle and Length Dependence of $^2J_{C1,C3}$, $^1J_{C1,C2}$ and $^1J_{C2,C3}$ in 1^C - 8^C . Geometric parameters and spin-spin coupling constants derived from Series 1 calculations. Colors indicate C1 through C3 hydroxyl configuration as per legend.298
- Figure 7.6. Dependence of $^2J_{C1,C3}$ on Additive Subtended Bond Lengths for 1^C , 2^C and 5^C . Color scheme for data is identical to Figure 7.1.301
- Figure 7.7. Electron Occupancy in C-O σ^* NBO’s in 1^C and 2^C as a Function of C2-C1-O1-H Dihedral. Panels A and C depict data for 1^C and panels C and D contain

data for 2^C . The labeling color scheme is the same as Figure 7.1. The clustering in panels A and B are indicative of the exoanomeric effect, whereas the relative ranges in the electron occupancy in panels C and D is indicative of the presence (panel C) or absence (panel D) of the endoanomeric effect. 303

Figure 7.8. Principal Component Analysis (PCA) of NBO Occupancies in 2^C 305

Figure 7.9. Combined NBO Electron Occupancy for 1^C , 2^C and 5^C as a Function of ξ . Combined electron occupancy is the sum of the individual occupancies for the $\sigma_{C1,C2}$, $\sigma_{C2,C3}$, $\sigma_{C1,C2}^*$ and $\sigma_{C2,C3}^*$ NBOs. Coloring scheme is identical to that in Figure 7.1. Dihedral ξ defined in text. 307

Figure 7.10. Normalized Plots of $2J_{C1,C3}$ vs. Sums of Selected NBO Occupancies in 2^C . The conformational color scheme is identical to Figure 1. (A) The $s_{C1,C2}$, $s_{C2,C3}$, $s_{C1,C2}^*$ and $s_{C2,C3}^*$ NBO occupancies were summed for a given conformer. The grey dotted lines indicate geometrically related conformers. (B) The s and s^* NBO occupancies for the C1-O5, C1-C2, C2-C3, C3-C4, C1-H1, C3-H3 bonds were summed along with $s_{C2,H2}$ for the 27 conformers of 2^C 309

Figure 7.11. Calculated $2J_{C2,C4}$ in 9^C as a function of the H3-C3-O3-H dihedral. Color scheme is analogous to that found in Fig. 7.1, except shifted with respect to the coupling path. For example, the C1-O1 torsion in Fig. 7.1 is the C2-O2 torsion here, etc. 312

Figure 7.12 Configuration Dependent Variation in the Vicinal $J(\sigma_{C,C})$. The differences in the pairwise steric exchange energy between $\sigma_{C2,C3}$ and $\sigma_{C1,O1}$ (A) and $\sigma_{C1,C2}$ and $\sigma_{C3,O3}$ (B) are significant contributors to the configuration dependent changes in the magnitude of $2J_{C1,C3}$ between the epimeric pairs 2^C and 1^C , and 1^C and 5^C respectively. The orbitals depicted are pre-orthogonal NBOs. The bonding (σ) orbitals have three lobes with a node at each of the bonded carbon or oxygen nuclei. The coloration denotes orbital phasing and is consistent only within an individual PNBO. The schematic diagrams show the relative hydroxyl orientations obscured by the presence of orbitals. 318

Figure 7.13. Configuration Dependence of $2J_{C1,C3}$ as Mediated by Vicinal $J(\text{deloc})$. The differences in orbital interactions $\sigma_{C2,C3} \rightarrow \sigma_{C1,O1}^*$ (A) and $\sigma_{C1,C2} \rightarrow \sigma_{C3,O3}^*$ (B) are the largest contributors to the configuration dependent changes in the magnitude of $2J_{C1,C3}$ between the epimeric pairs 2^C and 1^C , and 1^C and 5^C respectively. The bonding (σ) orbitals have three lobes with a node at each of the bonded carbon nuclei, and the antibonding (σ^*) orbitals have four lobes with a

node each at the bonded carbon and oxygen nuclei, as well as a third node between them. The coloration denotes orbital phasing and is consistent only within an individual PNBO. The schematic diagrams show the relative hydroxyl orientations obscured by the presence of orbitals.....321

Figure 7.14. O2 Lone Pair Orbitals from NBO Analysis of **2C**. Pseudo π shaped orbital (A) and pseudo σ shaped orbital (B) on O2 are the largest contributors to the ξ conformation dependent changes in the magnitude of $^2J_{C1,C3}$ between the two gauche and the trans conformers in **2C**. The dihedrals ϕ , ξ and $\zeta = 180^\circ$ for both structures depicted. The coloration denotes orbital phasing.325

Figure 7.15. The ξ Dependent Variation in the Pairwise Steric Exchange Energy Between O2lp π and $\sigma_{C1,C2}$ or $\sigma_{C2,C3}$. The overlap in the **2C** PNBOs for O2lp π and $\sigma_{C1,C2}$ (A) or O2lp π and $\sigma_{C2,C3}$ (B) are a reflection of the angle between the bilateral plane of symmetry for O2lp π (dotted line in central scheme) and either the C1-C2 or C2-C3 bond vectors respectively. This in turn is directly related to the pairwise steric exchange energies ($E_{i,j}$) which are listed in kcal/mol. An entry of *bt* indicates that the $E_{i,j}$ was below the 0.2 kcal/mol threshold. The corresponding $J^{(L)}$ contributions are listed in Hz for comparison. The bar above the *E* and the *J* indicates that the reported value is averaged over the 9 conformers of C1-O1 and C3-O3. The number in the parenthesis is the standard deviation in the last reported digit. The orbital coloration denotes phasing and is consistent only within an individual PNBO. The ϕ and ζ conformations depicted (definitions in text) all = 180° , whereas ξ conformation is consistent within a column.328

Figure 7.16. The ξ Dependent Variation in the Second Order Perturbation Energy Between O2lp π and $\sigma_{C1,C2}^*$ or $\sigma_{C2,C3}^*$ in **2C**. The overlap in the **2C** PNBOs for O2lp π and $\sigma_{C1,C2}^*$ (A) or O2lp π and $\sigma_{C2,C3}^*$ (B) vary as a function of ξ . This in turn is directly related to the 2nd-order perturbation energy ($E_{i,j}$) which are listed in kcal/mol. The corresponding $J^{(deloc)}$ is listed in Hz for comparison. An entry of *bt* indicates that the $E_{i,j}$ or the $J^{(deloc)}$ were below the respective 0.2 kcal/mol or 0.1 Hz thresholds. The bar above the *E* and the *J* indicates that the reported value is averaged over the 9 conformers of C1-O1 and C3-O3. The number in the parenthesis is the standard deviation in the last digit. The antibonding ($\sigma_{C,C}^*$) orbitals have four lobes with a node at each of the bonded carbon nuclei, as well as a third node between them. The ϕ and ζ conformations depicted (definitions in text) are all = 180° , whereas ξ conformation is consistent within a column.333

Figure 7.17. The Normalized Relationship Between ${}^2J_{C1,C3}$ and the Sum of Select ${}^2J_{C1,C3}^{(L)}$ and ${}^2J_{C1,C3}^{(deloc)}$. The select ${}^2J_{C1,C3}^{(L)}$ and ${}^2J_{C1,C3}^{(deloc)}$ terms summed and normalized for the 27 conformations of **2C** (A) are $\sigma_{C1,C2}$, $\sigma_{C2,C3}$ and $\sigma_{C1,C2}^*$, $\sigma_{C2,C3}^*$, $\sigma_{C1,O1}^*$, $\sigma_{C3,O3}^*$ respectively with the removal of $\sigma_{C1,O1}^*$ for **1C** (B) and the removal of both $\sigma_{C,O}^*$ terms for **5C** (C).....336

Figure 7.18. Calculated ${}^2J_{C1,C3}$ as a Function of ϕ , ξ and ζ with Parameterization Overlays. The ξ (H2-C2-O2-H) and ϕ (C2-C1-O1-H) dimensions are explicitly indicated while the ζ (C2-C3-O3-H) dimension is implicitly indicated by the vertical stacking of the hypersurfaces. Magenta and teal spheres indicate ${}^2J_{C1,C3}$ calculated by DFT for $\zeta = \pm 180^\circ$ and 0° respectively. Spheres have a 0.5 Hz diameter. DFT data for intermediate ζ values omitted for clarity. Solid surfaces represent best fit to the DFT data for $\zeta = \pm 180^\circ$ (indigo), $\pm 150^\circ$ (blue), $\pm 120^\circ$, (green), $\pm 90^\circ$ (yellow), $\pm 60^\circ$ (orange), $\pm 30^\circ$ (fuchsia) and 0° (red).341

Figure 7.19. Isosurface of ${}^2J_{C1,C3}$ in the ϕ , ξ and ζ conformational space. Grey spheres are combinations of ϕ , ξ and ζ (definitions in text) that lead to a ${}^2J_{C1,C3}$ of 4.6 Hz as determined by DFT. The spheres have a 15° radius and are digitized every 5° in the ϕ and ξ dimensions and every 30° in the ζ dimension.....348

Figure 7.20. The Effect of μ , σ and w on Population Distribution. The parameter μ controls the location of the peak maximum across the rotational itinerary whereas σ dictates the peak width (A). The parameter w dictates the fractional weighting of a peak with respect to others about the same rotational itinerary (B). The two curves in (B) bound areas of different size whereas the two curves in (A) have the same area.351

Figure 7.21. Population Histograms for Conformational Models about the Dihedrals ϕ , ξ and ζ in **9C**. The dihedrals for ϕ (A), ξ (B) and ζ (C) are shown in degrees whereas population is shown in arbitrary units. The raw data for the three simulations is shown rather than the best fit gaussian models for the simulations represented in Tables 7.10 and 7.11. The bin size for the three simulation histograms was every 1° . The definitions for models **I**, **II** and **III** are given in the text.....364

Figure 7.22. Bond Length, Relative Energy, and Dipole Hypersurfaces as a Function of ϕ , ξ and ζ for the *ccw* and *cw* H-bonding Conformers in **9C**. The dihedrals ϕ and ξ are measured in degrees. The global dipole is represented as the non-directional scalar quantity. Putative *ccw* and *cw* H-bonding geometries are depicted for the **II**

and **III** models and are indicated on the various hypersurfaces by open squares and circles respectively. The 0 kcal/mol global relative energy minimum used as a reference for panels C and G occurred in a slice through the ζ dimension not shown. The hydroxymethyl group, the O4 hydroxyl and all non-hydroxyl protons are omitted from the molecular models for clarity, except for H2 which was retained as a reference.....377

Figure 7.23. Radial Population Density Histograms about ϕ , ξ and ζ for **II** and **III** with Correlated Rotameric Transition Paths for the *ccw* and *cw* H-bonding Networks. Polar coordinates defined with arbitrary radial population units and a polar angle corresponding to the ϕ (A), ξ (B) or ζ (C) dihedrals measured in degrees. The annulus containing the nested Newman projections represents zero population. Black solid line represents model **II** and the grey line represents model **III**. Rotameric transition paths for the interconversion between *ccw* and *cw* conformers of **II** (D) and **III** (E) assume transitions proceed through adjacent rotameric regimes. Half arrows indicate transitional continuity about the entire rotational itinerary. The value for μ_ϕ used in the **III**_{*ccw*} model (E) is the weighted average of the μ_{ϕ_1} and μ_{ϕ_2} peaks from Table 7.11.....386

Figure 7.24. The Goodness of Fit Error for $^2J_{C1,C3}$ in **9C** as a Function of Dihedral. The definitions for the x and y values in panels B-D are found in panel A, where θ represents one of the dihedrals ζ , ϕ or ξ respectively. Data in panels B-D are color coded by virtue of the respective ξ_{DFT} , ϕ_{DFT} or ξ_{DFT} angles used to calculate $\Delta_{DFT-Fit}^2J_{C1,C3}$. The percentage of the 1728 DFT data points whose coupling magnitude was outside the range of the parameterization thereby precluding the calculation of $\theta_{\pm error}$ is indicated in parenthesis beside the panel label. Definitions for ζ , ϕ and ξ are found in the text.395

Figure 8.1. Atomic Numbering for I_E.....417

Figure 8.2. Northern Hemispheric Stereographical Projection of Pyranose Pseudorotational Itinerary with Ring Puckering Coordinates for I-V. ^aRing at the non-reducing end of the disaccharide II is glucopyranose instead of galactopyranose. The ring puckering parameters φ and θ are represented by angular and radial displacements about the ⁴C₁ origin (center point) respectively. The displacements for all crystal data in the main figure are based upon individual total puckering amplitudes, Q , while displacements in the inset are based upon an average amplitude for the depicted data ($Q_{avg} = 0.5734$). The inner and outer black circumferential rings represent the minimal and maximal Q values for the total data set, while the colored rings correspond to the Q values for the galacto and gluco rings of I_E and424

TABLES

Table 7.1. Comparison of Experimental $^2J_{C1,C3}$ Values in Hz for Selectively ^{13}C -Labeled Hexo-Pyranoses with Theory as a Function of Coupled Site Configuration.....	276
Table 7.2. Theoretical $^2J_{C1,C3}$ Values ^a as a Function of C2-O2 Rotamer Conformation Averaged Across All Staggered C1-O1 and C3-O3 Rotamer Conformations....	288
Table 7.3. Comparison between Calculated $^2J_{C1,C3}$ in 11C , 12C and 14C as a function of C2-C1-O1-H and C2-C3-O3-H dihedral.	292
Table 7.4. Statistical Comparison ^a between Calculated $^2J_{C1,C3}$ in 2C and $^2J_{C2,C4}$ in 9C	313
Table 7.5. Configuration dependence ^a of $J(L)$ in 2C , 1C and 5C	316
Table 7.6. Configuration dependence ^a of $J(\text{deloc})$ in 2C , 1C and 5C	320
Table 7.7. Comparison of the Difference Between Averaged H2-C2-O2-H <i>trans</i> and <i>gauche</i> $J(L)$ and $J(\text{deloc})$ with DFT Total $^2J_{C1,C3}$ as a Function of Coupled Site Configuration ^a	326
Table 7.8. ξ Conformational Dependence of O2 lp $\rightarrow\sigma_j^*$ as a Function of Coupled Site Configuration ^a	332

Table 7.9. Generalized Spin-Spin Coupling Equation ^a and Coefficients ^b for the Parameterization of Couplings Sensitive to C1 Through C3 Hydroxyl Orientation.	343
Table 7.10. Spin-spin Coupling Constant Comparisons between Experiment and Various Conformational Models.	362
Table 7.11. Conformational Model Parameters for ϕ , ξ and ζ in 9C	363
Table 8.1. Geometric Parameters for I _E and I _C (Å, °) with select comparisons to II-V.	419
Table 8.2. Hydrogen-bond geometry (Å, °) in Crystalline I _E	422
Table 8.3. Cremer-Pople Pyranose Ring Puckering Parameters ²⁵ for I-V.	426

SCHEMES

Scheme 7.1. Structures for model compounds 1C-18C (A) and experimental compounds 1E-7E , 11E-17E , and LE (B)	279
Scheme 7.2. Rotameric Definitions of the Dihedrals ζ , ξ and ϕ	285
Scheme 7.3. Explanation of $^2J_{C1,C3}$ Karplus Profile Shape Based on O2lp π Disposition	330
Scheme 7.4. Orbital Interpretations of Models II and III	379
Scheme 8.1. Coupling Reaction for the Synthesis of IE	428
Scheme 9.1. Complex <i>N</i> -Glycan.....	444
Scheme 9.2. Structure of <i>Campylobacter jejuni</i> Serotype O:19 Antigenic Lipopolysaccharide OH 4384-C1 and Related Human Neolacto Series Gangliosides. Residues in blue are structurally related portions of the antigenic terminus of the OH 4384-C1 LPS. Select abbreviations: LDHep, L- <i>glycero</i> -D- <i>manno</i> -heptose; PEA, <i>O</i> -phosphoethanolamine; KDO, 3-deoxy-D- <i>manno</i> -octulosonic acid.	445

CHAPTER 7:
HYDROXYL GROUP CONFIGURATION / CONFORMATION AND $^2J_{\text{CCC}}$ SPIN-
COUPLINGS IN SACCHARIDES: DFT, NJC AND NSA ANALYSES OF
EXPERIMENTAL DATA⁶

*“You are not misled by what you don’t know. The trouble comes from what you think
you know but that’s mistaken.”*

– Garron L. Klepach

“If it’s not true, it’s well invented.”

– Dante Alighieri

7.1. Abstract

Interpretation of experimental NMR spin-spin coupling constants (SSCC; J -couplings) using theoretically derived correlations between SSCCs and geometric parameters such as bond length, bond angle and dihedral angle has led to quantitative conformational analyses of saccharides in solution. Despite the success of these phenomenologic correlations, the underlying mechanisms of spin-density transfer and their relationship to molecular geometry remain largely unexplored in saccharide

⁶ I gratefully acknowledge my co-authors Hongqiu Zhao, Xiaosong Hu, Wenhui Zhang, Ian Carmichael and Anthony S. Serianni for their assistance in preparing this chapter, which is submitted to the Journal of the American Chemical Society.

systems. In our continuing efforts to develop NMR spin-couplings as experimental probes of carbohydrate structure, we sought to cross-validate experiment with theory for the dependence of intra-ring $^2J_{\text{CCC}}$ magnitude and sign on saccharide configuration at the coupled carbons (*J. Org. Chem.* **2007**, 72, 7511-7522). We also investigated the dependence of $^2J_{\text{CCC}}$ on hydroxyl group conformation (exocyclic C-O bond rotation). Since empirical correlations between experimental J -couplings and molecular geometry are to some extent incidental, with the former dependent on, and the latter dictating the underlying electronic structure of the molecule, an attempt was made to interpret theoretical findings within the context of geometry-dependent electron occupancy variations in a natural bonding orbital (NBO) analysis as correlated with specific spin-density transfer mechanisms described by a finite field perturbation natural J -coupling (NJC) analysis. Within the NBO paradigm, two distinct spin-density transfer mechanisms were identified for both the configurational and conformational dependencies, namely, destabilizing steric (through-space) and stabilizing hyperconjugative (through-bond) interactions. The merits and weaknesses of the NBO interpretation will be discussed.

7.1.1. Introduction

Carbohydrates are the most abundant class of organic biomolecules on earth, whose diverse functions extend far beyond central metabolism and energy storage. Carbohydrates assist in protein folding and processing, molecular recognition, cellular structural support, cell-cell adhesion, and hydration regulation^{1a}. Despite their importance, numerous fundamental questions^{1b-c} persist such as a detailed understanding

of the behavior of glucopyranose in aqueous solution. By virtue of their polyhydroxylated nature, a crucial component of understanding the structure-function relationship in carbohydrates is detailed knowledge of the solution behavior of the numerous C-O rotamers. The conformation of carbohydrate hydroxyl groups has been experimentally shown to significantly affect their reactivity as aglycones in glycosylation reactions⁸⁶. Additionally, hydroxyl group configuration is of central functional importance in molecular recognition events such as the axial O4 hydroxyl in the terminal galactose residue in the binding of *N*-acetyl-lactosamine to human galectin⁶⁹. Differential solvation properties of C4 epimers in hexopyranoses have been implicated in protein binding mechanisms⁶⁷. Furthermore, the overall hydration state of a saccharide and its effect on hydroxyl reorientation are key to the amphiphilic model of molecular recognition⁷⁰. For example, the binding of the bridgehead methyl-tri-mannoside from the high-mannose *N*-glycan to the concanavalin A lectin has been shown to crucially involve a single water molecule tightly bound to O2 of the central mannose residue^{70b}. Hydroxyl conformation of a particular residue can be highly context dependent and closely related to larger structural elements such as oligosaccharide tertiary structure, requiring a battery of tools for structural determination applicable to a broad range of potential molecular milieu. There is a wide array of complementary NMR observables such as chemical shift (δ), NOE measurements, indirect spin-spin coupling (J), and residual dipolar couplings (D) that return structural information⁷¹.

The dependence of a variety of proton based NMR spin-spin couplings ($J_{H,H}$) on molecular geometry is well recognized², however the advent of facile site specific ¹³C isotopic enrichment has led to the increased importance of carbon based couplings, $^nJ_{C,H}$

and ${}^nJ_{C,C}$ ³⁻⁴. A variety of these ${}^{13}\text{C}$ based couplings have been shown to be dependent on hydroxyl configuration and conformation^{4b,5-7,38}. Given the abundance of spin-spin coupling constant data systematically gleaned from an ever growing arsenal of strategically enriched saccharides, a host of empirical correlations between molecular configuration and J have been recognized such as the projection resultant¹¹ and the 1-3 di-axial effect⁷². Another example involves previously reported experimental ${}^2J_{C1,C3}$ values in a variety of hexo-pyranoses^{3a,10} that show a dependence on the configuration of the coupled carbons, specifically equatorial hydroxyls push the coupling towards more positive values (Table 7.1.)

The interpretation of experimental NMR spin-spin coupling constants (SSCC) via theoretically derived correlations between the pertinent SSCC and various geometric parameters such as bond length, bond angle, and dihedral has led to quantitative conformational analysis of saccharide structure³⁷. Despite the success of these phenomenologic correlations, the underlying mechanisms of spin density transfer and their relationship to molecular geometry remain largely unexplored in saccharide systems.

In the continuing effort to develop NMR spin-spin couplings as experimental probes of carbohydrate structure we sought to cross validate experiment with theory for the dependence of ${}^2J_{CC}$ magnitude on configuration at the coupled carbons. Further, we explored the potential dependence of ${}^2J_{CC}$ on the primary geometric parameter hydroxyl group conformation and the secondary parameters bond length and angle along the

TABLE 7.1. COMPARISON OF EXPERIMENTAL $^2J_{C1,C3}$ VALUES IN HZ FOR SELECTIVELY ^{13}C -LABELED HEXO-PYRANOSES WITH THEORY AS A FUNCTION OF COUPLED SITE CONFIGURATION

Group 1. (Equatorial / Equatorial)			Group 2. (Equatorial / Axial)			Group 3. (Axial / Axial)		
	DFT	Expt		DFT	Expt		DFT	Expt
β -Glc p (2)	(+) 4.6	(+) 4.5	α -Glc p (1)	(+) 0.1	<i>nc</i>	α -Allop (5)	(-) 2.0	(-) 2.4
β -Man p (4)	(+) 3.9	(+) 4.0	α -Man p (3)	(+) 0.1	<i>nc</i>	α -Altrop (7)	(-) 1.3	<i>nc</i>
			β -Allop (6)	(+) 0.3	<i>nc</i>			
			β -Altrop (8)	(+) 0.3				

In Hz \pm 0.1 Hz determined in D-glucose, D-mannose, methyl D-Allopyranoside, and methyl α -D-Altropyranoside selectively ^{13}C -Labeled at C1; in $^2\text{H}_2\text{O}$ at $\sim 25^\circ\text{C}$. An entry of *nc* implies that $J < 0.6$ Hz; coupling sign is in given in parenthesis. First header row denotes the axial / equatorial disposition of the hydroxyl groups for the coupled carbons, C1 and C3. Theoretical values were averaged across an ensemble of 27 conformations for a given configuration.

coupling pathway. As empirical correlations drawn between experimental J -couplings and molecular geometry are to a point incidental, the former being dependant upon, and the latter dictating the underlying electronic structure of the molecule, an attempt was made to understand the theoretical results within the context of geometry dependent electron occupancy variations in a natural bonding orbital (NBO) analysis as correlated with specific spin density transfer mechanisms described by a finite field perturbation natural J -coupling (NJC) analysis.

Finally, calculated $^2J_{C1,C3}$ is parameterized as a function of the conformational hypersurface for the C1-O1, C2-O2, and C3-O3 rotamers in β -D-glucopyranose by DFT along with the ensemble of $^3J_{HH}$, $^nJ_{CH}$, $^nJ_{CC}$, $^3J_{H(O)H}$ and $^nJ_{C(O)H}$ couplings sensitive to the mentioned dihedrals, and used in the conformational analysis of the C1-O1, C2-O2, and C3-O3 rotamers of *O*-methyl- β -D-glucopyranoside based on experimental couplings measured in DMSO. This includes the development of a generalized treatment for hydroxyl conformation and a thorough discussion of potential sources of error in the analysis.

7.2. Experimental

7.2.1. Experimental ^{13}C - ^{13}C Spin-Coupling Measurements in DMSO

The compounds used in this study are depicted in Scheme 7.1. The superscript ‘E’ designates the compound as one in which experimental measurements were conducted, whereas a superscript ‘C’ indicates a model compound used for theoretical calculations. When no superscript is indicated in the text, the compound number refers to both the

experimental compound and the theoretical analog. The various ^{13}C labeled reducing sugars were obtained from Omicron Biochemicals Incorporated, South Bend, Indiana. Methyl glycosides were prepared by standard Fischer glycosidation as described previously³⁸. Experimental ^1H - ^1H , ^{13}C - ^1H and ^{13}C - ^{13}C spin-couplings were measured in fresh $[\text{}^1\text{H}_6]$ -DMSO in 3 mm NMR tubes (Wilmad) at 25°C from 1-D spectra collected on a Varian Unity*Plus* 600 NMR spectrometer operating at 599.89 MHz for ^1H and 150.85 MHz for ^{13}C with a dual $^{13}\text{C}/^1\text{H}$ microprobe (Nalorac).

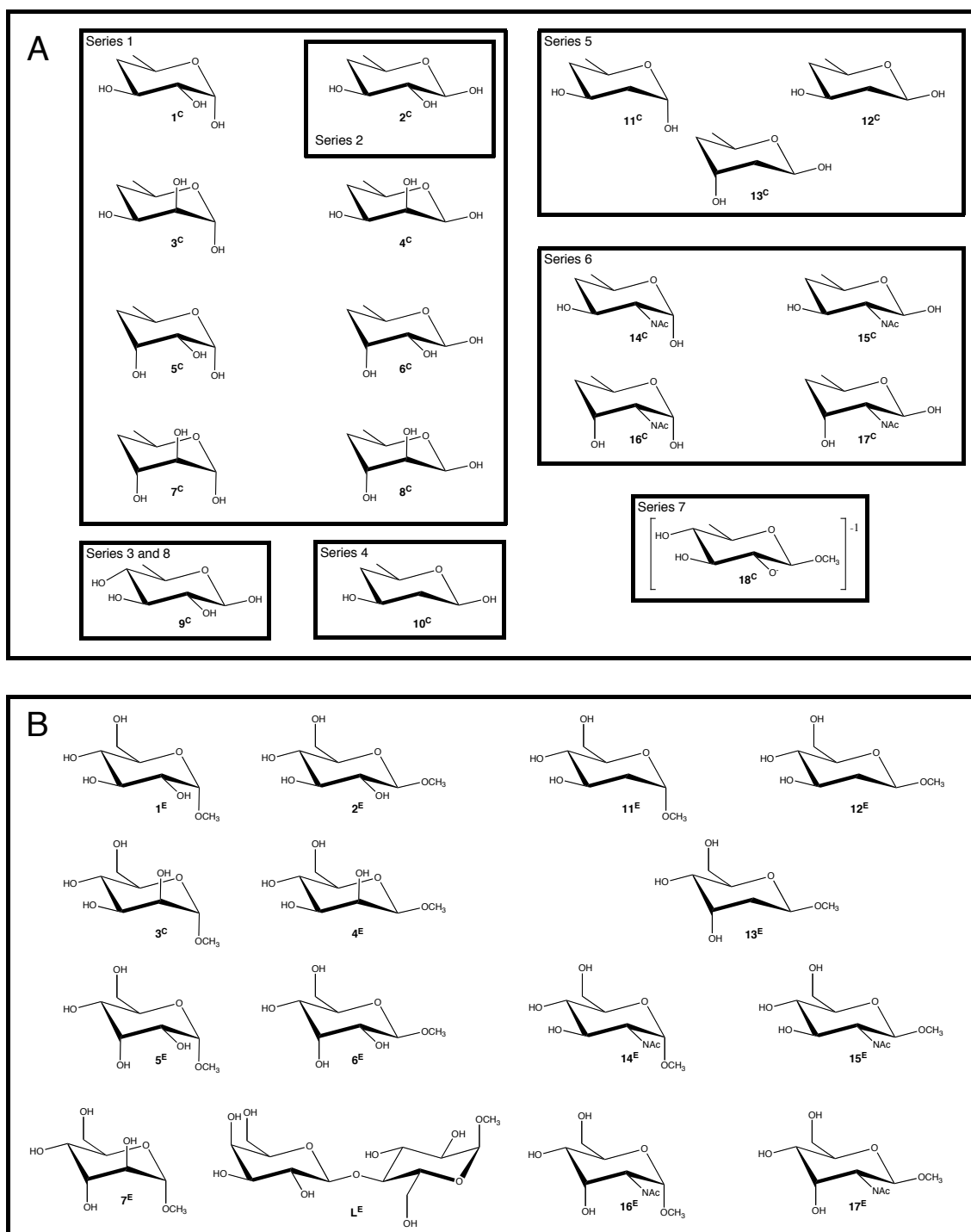
7.3. NMR titration of experimental $^2J_{\text{C1,C3}}$ by ionization of the O2 hydroxyl

Measurement of $^2J_{\text{C1,C3}}$ in a 300 mM aqueous solution (5% $^2\text{H}_2\text{O}$) of $[\text{}^{1-13}\text{C}]\text{-1E}$ was conducted at a series of pH values ranging from 9 to ~12.5. Measurements came from 1-D spectra collected on a Varian Unity*Plus* 600 NMR spectrometer operating at 150.85 MHz for ^{13}C with a dual $^{13}\text{C}/^1\text{H}$ microprobe (Nalorac).

7.4. Computational

7.4.1. Selection and Geometric Optimization of Model Compounds

Theoretical calculations of J_{CC} values were conducted in 8 distinct series (Scheme 7.1). All density functional theory (DFT) calculations were conducted within *Gaussian03*¹⁹ using the B3LYP functional.²⁰ The first, third through sixth and eighth series were conducted *in vacuo* with the 6-31G* basis set²¹ for geometric optimization as described previously.^{22,23} The second and seventh series were optimized in a polarizable



Scheme 7.1. Structures for model compounds 1^C-18^C (A) and experimental compounds 1^E-7^E, 11^E-17^E, and L^E (B)

solvent continuum model using the integral equation formalism¹⁶ in which the solute cavity is created via a series of overlapping spheres.^{17a-f} An extended double-zeta basis set ([5s2p1d|3s1p])¹⁸ was used for these series of calculations.

The structures in Series 1 consisted of the alpha and beta anomers of the 4,6-dideoxy-D-aldohexopyranoses: α -Glc_p (**1^C**), β -Glc_p (**2^C**), α -Man_p (**3^C**), β -Man_p (**4^C**), α -Allo_p (**5^C**), β -Allo_p (**6^C**), α -Alt_p (**7^C**), and β -Alt_p (**8^C**). These structures represent the eight possible C1 through C3 configurational permutations. Each of the eight model compounds was optimized in 27 unique conformations characterized by all possible combinations of the staggered geometries about the three dihedrals ϕ = C2-C1-O1-H, ξ = H2-C2-O2-H, and ζ = C2-C3-O3-H. These C-O rotamers were held in the appropriate fixed geometries during optimization.

The structures for the Series 2 calculations consisted of a much smaller subset of those used in the first series. All structures were in the β -Glc_p (**2^C**) configuration and all of the C-O rotamers mentioned above were fixed during the course of optimization. The ϕ dihedral was fixed at 180° for all structures. The ξ dihedral sampled all three staggered conformations for ζ values of 60° and 180°, yielding 6 structures in all.

The third series of geometric optimizations was conducted on 6-deoxy- β -D-glucopyranose (**9^C**) in which the ϕ dihedral was fixed at 180° for all structures. The 27 possible combinations of staggered C-O rotamers for the remaining hydroxyl groups were optimized with the dihedrals C3-C2-O2-H, H3-C3-O3-H, and C3-C4-O4-H all fixed at the appropriate values. All other geometric parameters were allowed to freely optimize. The Series 8 calculations were also conducted on **9^C**, however C3-C4-O4-H was held fixed at 180° while the three dihedrals ϕ , ξ and ζ were simultaneously scanned across

their rotational itineraries in 30° increments giving rise to 1728 unique energetically converged conformers.

The fifth series of calculations was conducted on **11^C**, **12^C** and **13^C**. These structures are analogous to **1^C**, **2^C** and **6^C** from Scheme 7.1A respectively except they are deoxygenated at C2. Geometries about ϕ and ζ were simultaneously scanned about their three respective staggered geometries while all other geometric parameters were allowed to relax for a total of 9 unique conformers.

The sixth series was conducted on structures **14^C** through **17^C** (Scheme 7.1A). These structures are C2 *N*-acetylated analogs of **1^C**, **2^C**, **5^C** and **6^C** from Scheme 7.1A respectively. Since the calculations were conducted to assess the effect of C2-N rotation on $^2J_{C1,C3}$, the ϕ and ζ dihedrals were frozen in the orientation of maximal coupling in Series 1 (180° for both dihedrals) while the C1-C2-N-C $_{\alpha}$ dihedral was iteratively scanned about its rotational itinerary in 15° increments and all other geometric parameters were allowed to relax during energetic minimization.

The seventh series of geometric optimizations was conducted on structure **18^C** (Scheme 7.1A). This structure is an analog of **9^C** derivatize with a methyl glycoside. Additionally the O2 hydroxyl proton has been removed creating a negative charge on O2, thereby mimicking the ionization state at high pH. The aglycone was held fixed anti to C2 in all calculations, while ζ was scanned about its three staggered geometries. All other geometric parameters were allowed to freely optimize.

7.4.2. Theoretical Calculations of ^{13}C - ^{13}C Spin-Coupling Constants

J -Coupling constants were calculated in all structures (Scheme 7.1A), using *Gaussian03*¹⁹ with the extended basis set ([5s2p1d|3s1p]).¹⁸ Both the Fermi and non-Fermi contact terms were recovered, and the reported values contain both contributions and are unscaled. The second and seventh series of structures had their J -coupling constants calculated *in solvo* using the solvent model formalism described above. All other SSCCs were calculated *in vacuo*.

7.4.3. Natural Bonding Orbital Calculations

A full Natural Bonding Orbital (NBO) analysis was conducted on all structures (Scheme 7.1A), using the NBOv3 package²⁴ native to *Gaussian03*^{19a} with the extended basis set ([5s2p1d|3s1p]).¹⁸ *Gaussview*4.12 was used to visualize all orbitals^{19b}. Natural J -Coupling (NJC) and Natural Steric Analysis (NSA) was conducted on the optimized geometries from **2C**, **1C** and **5C** (Scheme 7.1A), using the NBOv5.G package²⁴ as a sourced link within *Gaussian03*¹⁹ with the extended basis set ([5s2p1d|3s1p]).¹⁸

7.4.4. Theoretical SSCC Parameterization and Analysis of Experimental Couplings

Quantitative parameterization of a battery of 15 SSCCs sensitive to one or more of the ϕ , ξ and ζ dihedrals as a function of these C-O rotamers was based on the SSCCs calculated for the 1728 DFT optimized structures generated in Series 1. The parameterization was conducted by a combination of a least squares Monte Carlo fitting procedure, a “Robust” χ^2 based deterministic optimization and a variant of the deterministic Levenberg-Marquardt algorithm using *ProFit 6.1.6* (Quantum Soft, Zürich,

Switzerland). This procedure was conducted term-wise with all stochastic optimization steps utilizing > 1,000k iterations. The *3DPlotterGL 1.2.7* module within *ProFit 5.6.2* was used to generate the 3-dimensional images of the coupling profile and the 2-dimensional topographic bond length, dipole and relative energy hypersurfaces.

The analysis of experimental SSCCs was conducted using a combination of the Monte Carlo and Robust fitting procedures mentioned above in a plug-in to *ProFit 6.1.6* written in our lab called *GlyFit 2.0* which is freely available upon request. All statistical analyses were conducted using *SPSS 11.0.4* (SPSS Inc., Chicago, Illinois). The molecular model images were generated by *Spartan ST 1.1.3* (Wavefunction, Inc., Irvine, California).

7.5. Results and Discussion

7.5.1. The Dependence of $^2J_{C1,C3}$ on Configuration

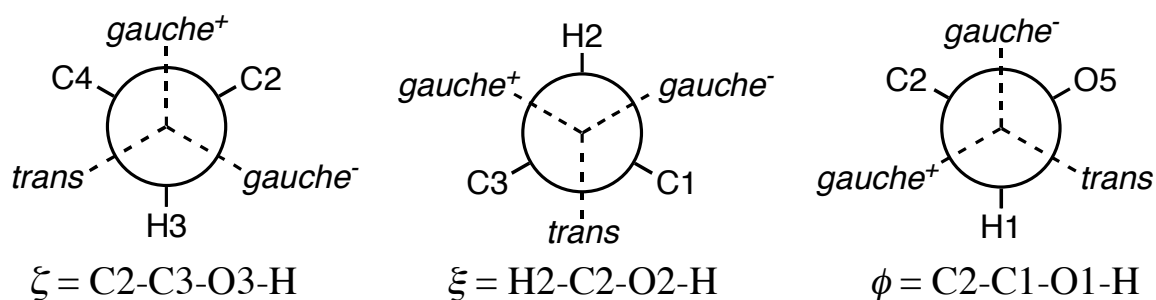
A comparison between the previously reported experimental $^2J_{C1,C3}^{3a,10}$ and analogous theoretical values determined by a linear averaging of the DFT values calculated for the 27 different combinations of staggered hydroxyl group conformations at C1 through C3 sampled for each configuration, models **1C-8C** (Series 1), can be found in Table 7.1. In all configurations for which the experimental coupling is available, the experimental and theoretical values are in good agreement, well within the uncertainty of the experimental measurements, except in the case of α -D-altropyranose. This is likely due to ring inter-conversion between 4C_1 and 1C_4 forms in solution affecting the

experimental value as discussed previously^{37c}. Data in Table 7.1 is organized into three groups distinguished by the combination of axially or equatorially oriented C1 and C3 hydroxyl substituents so that the following empirical trend emerges: $^2J_{C1,C3}$ values in which the hydroxyls at both coupled carbons are equatorial (Group 1) range between (+) 4-5 Hz; rings in which one of *either* of the coupled centers is equatorial and the other is axial (Group 2) show near zero $^2J_{C1,C3}$ values; and the di-axial case (Group 3) leads to values between (-) 2-3 Hz. Additionally, the theoretical data indicates that an axial O1 and equatorial O3 yield a $^2J_{C1,C3}$ value 0.2 Hz smaller than the contrary situation in which there is an axial O3 and equatorial O1. Any significance of this is attributable to the stereo-electronic differences between the anomeric carbon and C3. In contrast to the dependence of $^2J_{C1,C3}$ on configuration at the coupled carbons, configuration at the intervening carbon (C2) affected the magnitude of the coupling minimally and only in the di-equatorial or di-axial cases, **2C** exhibiting a ~0.6 Hz larger coupling than **4C** for both the experimental and DFT values, and **5C** a more negative coupling than **7C** by approximately 0.8 Hz.

7.5.2. The Dependence of $^2J_{C1,C3}$ on Conformation

When the $^2J_{C1,C3}$ values of all 27 combinations of staggered C1-O1, C2-O2 and C3-O3 rotamer conformations for a given model compound are sorted by the dihedrals ϕ (C2-C1-O1-H), and ζ (C2-C3-O3-H) and plotted as a function of ξ (H2-C2-O2-H), a strong dependence on the conformation of the C2 hydroxyl emerges. The rotameric definitions of these three dihedrals can be found in Scheme 7.2. Plots of $^2J_{C1,C3}$ versus ξ

for 2^C , 1^C , and 5^C representative of the three groups designated in Table 7.1 are found in Figure 7.1. The trend, which holds systematically throughout the entire data set (data for the remaining models omitted), exhibits equivalent couplings for the two gauche conformations ($\xi = 60^\circ$, -60°) and a coupling pushed towards more positive values when anti ($\xi = 180^\circ$). The spread in the data for any particular value of ξ is due to the dependence of ${}^2J_{C1,C3}$ upon the other two C-O rotamers belonging to the coupled carbons.



Scheme 7.2. Rotameric Definitions of the Dihedrals ξ , ξ and ϕ .

These secondary conformational dependencies are predominantly non-systematic and significantly smaller than the primary dependence of ${}^2J_{C1,C3}$ on the C2-O2 rotamer conformation. For any specific combination of C2-O2 and either C1-O1 or C3-O3 rotamer values, the secondary dependency of ${}^2J_{C1,C3}$ on the remaining either O3 or O1 hydroxyl conformation respectively is usually less than ~ 0.5 Hz in dynamic range for all combinations of C1 and C3 configurations except for the two Group 1 di-equatorial cases, 2^C and 4^C , in which instance the secondary dependence can vary by up to 1.3 Hz for geometries in which $H2-C2-O2-H = 180^\circ$. This demonstrates a theme repeated throughout the ensemble of configurations, namely that *the conformation of the C2*

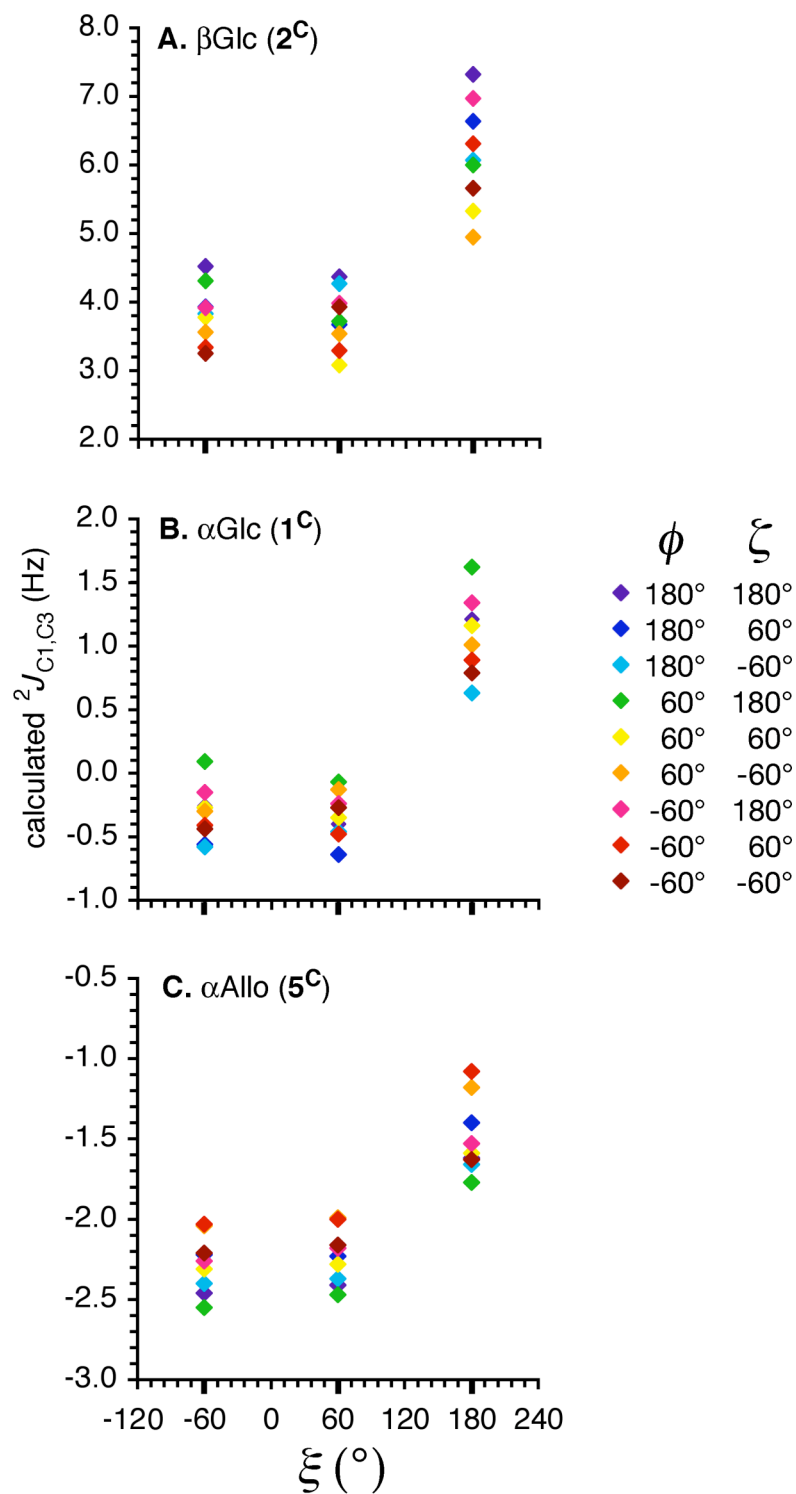


Figure 7.1. Conformational Dependence of ${}^2J_{C1,C3}$ on ξ . Definitions of ϕ , ξ and ζ are found in the text. Specific values for ϕ and ζ are represented by the coloration indicated in the legend.

hydroxyl group which orients the hydroxyl proton anti to the vicinal C-H bond potentiates the secondary conformational dependencies of the $^2J_{C1,C2}$ on the coupled carbon hydroxyl groups.

For example, the additive effects of variation in C1 and C3 hydroxyl conformation on the coupling can vary from ~0.5 Hz in range for gauche ξ in Group 3 structures **5^C** and **7^C**, and in the Group 2 structures such as **1^C**, **3^C**, **6^C**, and **8^C**, to upwards of ~2.4 Hz in **2^C** and ~2.7 Hz in **4^C** for geometries that have the O2 hydroxyl proton anti to H2. The nature of these secondary dependencies are highly context dependent; the dependency of either the O1 or O3 rotamers is dictated by the configuration of both O1 and O3, and in some cases by the conformation of the other C-O rotamer, presumably through a 1,3-di-axial type interaction.

It is worth noting that in the four model compounds in which O3 is equatorial, when the O3 hydroxyl proton is anti to C2 and in plane with the coupling pathway, the J is systematically larger by ~0.5 Hz. This appears to be true for O1 in the model compounds with the beta anomeric configurations as well, although the effect ϕ conformation is somewhat complicated by presence of a slight increase in the $^2J_{C1,C3}$ for O1 conformations that orient the hydroxyl proton anti to the ring oxygen. This latter effect of the O1 hydroxyl proton with respect to the ring oxygen is also observed in model compounds **1^C** and **3^C**, however the interpretation is again complicated by 1,3-di-axial interactions in **5^C** and **7^C**.

Table 7.2 contains average $^2J_{C1,C3}$ values for models **1^C**-**8^C** as a function of ξ conformation averaged across all values of the ϕ and ζ rotamers sampled for a given ξ

value. The data is grouped by configuration at the coupled sites according to the convention used in Table 7.1. Also contained in Table 7.2 are absolute values for the difference in ${}^2J_{C1,C3}$ between the average of the two O2 *gauche* conformations and the *anti* conformation (defined with respect to H2) for **1C-8C**. Again these values break up into the three categories defined by coupled carbon configuration, with the two di-

TABLE 7.2. THEORETICAL ${}^2J_{C1,C3}$ VALUES^a AS A FUNCTION OF C2-O2 ROTAMER CONFORMATION AVERAGED ACROSS ALL STAGGERED C1-O1 AND C3-O3 ROTAMER CONFORMATIONS

	H2-C2-O2-H (deg)			$\Delta_{t \rightarrow \bar{g}} {}^2J_{C1,C3}{}^b$
	180°	60°	-60°	
β -Glc p (2C)	6.1	3.8	3.8	2.4
β -Man p (4C)	5.6	3.1	3.1	2.5
α -Glc p (1C)	1.0	-0.3	-0.3	1.4
α -Man p (3C)	1.1	-0.3	-0.3	1.4
β -Allo p (6C)	1.2	-0.2	-0.2	1.4
β -Altrop (8C)	1.2	-0.2	-0.1	1.4
α -Allo p (5C)	-1.5	-2.2	-2.3	0.8
α -Altrop (7C)	-0.6	-1.6	-1.5	0.9

^aAll values reported in Hz. ^bAbsolute difference in ${}^2J_{C1,C3}$ between the nine averaged *trans* and the eighteen averaged *gauche*(+) and *gauche*(-) rotamers.

equatorial cases demonstrating a difference in the *gauche* versus *anti* O2 conformational dependence of $^2J_{C1,C3}$ of ~ 2.5 Hz; an average value of ~ 1.4 Hz in the four member mixed equatorial/axial configuration grouping; and ~ 0.9 Hz in the two cases that have both axial O1 and O3. This demonstrates an important principle, discussed in more detail below (*vide infra*, §I); the conformational dependence of $^2J_{C1,C3}$ on O2 hydroxyl conformation is attenuated by configuration at the coupled carbons; specifically the more axial C1 or C3 hydroxyls present, the less sensitive the coupling is to C2 hydroxyl conformation.

The effect of a solvent continuum on the geometric optimizations and all further calculated parameters was assessed for 6 staggered C-O conformational permutations in **2C** (Series 2) and found to be minimal when compared to *in vacuo* calculations. A comparison between the *in vacuo* and *in solvo* $^2J_{C1,C3}$ coupling data can be found in Figure 7.2. The 6 structures selected represent the extrema of the coupling profile for **2C**. The $^2J_{C1,C3}$ was attenuated by ~ 0.3 Hz in the conformation presenting the maximal coupling with the three dihedrals ϕ , ξ and ζ all equal to 180° , and the coupling was enhanced by ~ 0.2 Hz in the solvated calculations that had *gauche* C-O rotamer conformations that lead to less positive couplings. Taken together, this represents a reduction in the dynamic range of the coupling of ~ 0.5 Hz for the geometries explored in a coupling that shows a 4.2 Hz overall dynamic range.

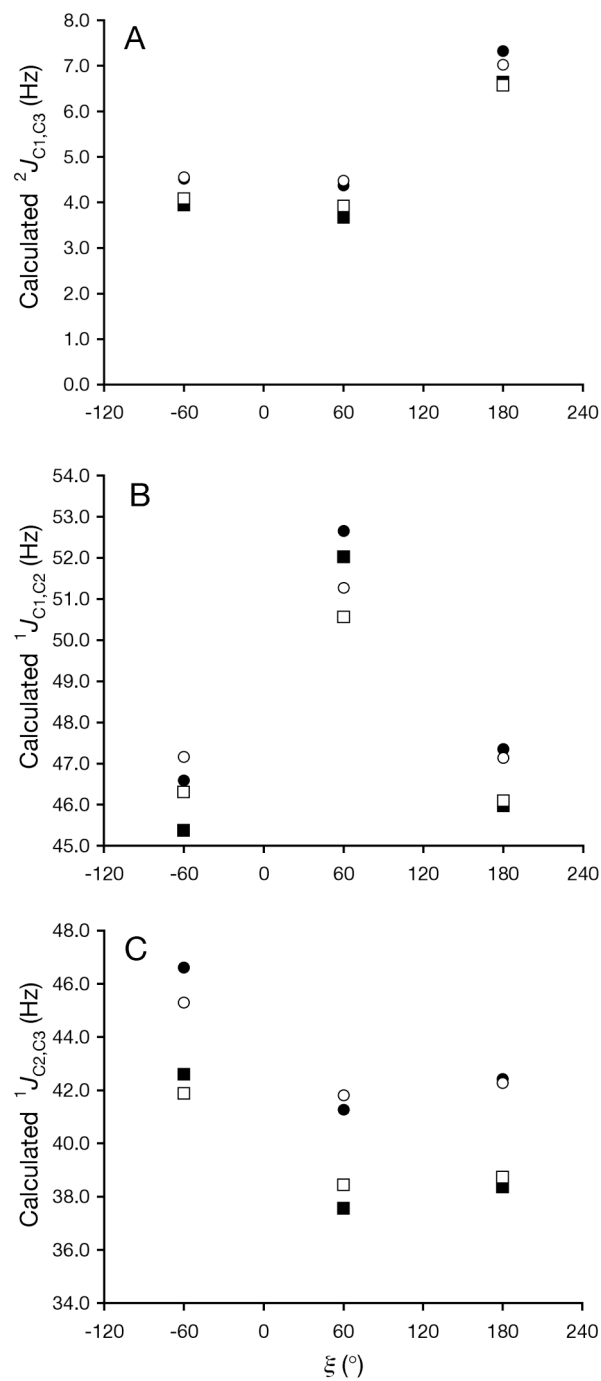


Figure 7.2. Comparison of Spin-Spin Couplings Calculated *In Vacuo* and *In Solvo* as a Function of ξ in $2C$. Data depicted with circles and squares are structures in which $C2-C3-O3-H$ (ξ) = 60° and 180° respectively. Solid and outlined shapes were data calculated *In Vacuo* and *In Solvo* respectively. Definition of ξ in the text.

7.5.3. The Effect of Deoxygenation on $^2J_{CC}$

Deoxygenation at the C2 position in hexo-pyranoses is linked to numerous biological activities. 2-Deoxy-D-glucose is an antimetabolite of glucose^{78a} with antiepileptic^{78b}, antiviral^{78c-d} and antitumor activity^{78e}. 2-Deoxy-D-glucose is also used as a tissue marker for glucose uptake and hexokinase activity due to this proteins ability to trap the saccharide in most tissue types. A 2-deoxy-D-glucose binding lectin has been identified in the symbiotic adherence of *Rhizobium trifolii* to clover root hairs²⁵. Deoxygenation at this position in a pyranose ring is also notably found in the saccharide sialic acid, an important 9 carbon α -keto acid.

The effect of C2 hydroxyl substitution on $^2J_{C1,C3}$ has been treated semi-quantitatively in the past through invoking the projection resultant relationship between the orientation of electronegative substituents with respect to the coupling path and $^2J_{CC}$ in aldopyranosyl rings¹¹. A comparison between experimental $^2J_{C1,C3}$ in **11^E**, **12^E** and **13^E** with DFT values in **11^C**, **12^C** and **13^C** (Series 5) as a function of C1-O1 and C3-O3 rotamer conformation can be found in Table 7.3. The Theoretical data in Table 7.3 cross validate the coupling trends in both the experimental data as a function of configuration and the couplings predicted by the projection resultant method as a function of electronegative substituent orientation. Deoxygenation at carbon 2 drives the coupling towards a more negative value when compared to the fully hydroxylated analogs for all combinations of configurations at C1, C2 and C3, in some instances by as much as 2.8 Hz. The effect of either C1-O1 or C3-O3 rotation in isolation on the dynamic range of the coupling for the 2-deoxy model compounds was always ≤ 0.7 Hz. The difference

between the experimental and theoretical values for $^2J_{C1,C3}$ in **12** is worthy of note; **12^E** displays a 1.8 Hz experimental coupling, however when the DFT coupling data for all 9

TABLE 7.3. COMPARISON BETWEEN CALCULATED $^2J_{C1,C3}$ IN **11^C**, **12^C** AND **14^C** AS A FUNCTION OF C2-C1-O1-H AND C2-C3-O3-H DIHEDRAL.

	$^2J_{C1,C3}$			$^1J_{C1,C2}$		
	12^C	11^C	14^C	12^C	11^C	14^C
C2-C1-O1-H = 180° C2-C3-O3-H						
180°	3.2	-1.5	-1.4	43.7	43.7	45.3
60°	2.6	-1.3	-1.1	42.6	40.2	46.0
-60°	2.5	-1.1	-1.0	43.0	40.7	45.3
C2-C1-O1-H = 60° C2-C3-O3-H						
180°	2.4	-1.5	-1.3	40.1	39.9	41.2
60°	1.9	-1.6	-1.7	38.8	39.2	40.7
-60°	1.7	-1.8	-1.5	39.2	42.9	41.4
C2-C1-O1-H = -60° C2-C3-O3-H						
180°	2.8	-1.7	-1.2	41.6	42.9	42.4
60°	2.2	-1.5	-1.6	40.3	39.1	42.4
-60°	2.0	-1.4	-1.3	40.7	39.8	43.0
Coupling Range	1.5	0.6	0.8	4.9	4.6	5.3
Average Coupling ^a	2.4 (1.8)	-1.5 (-2.3)	-1.4 (-1.8)	41.1 (40.3)	41.0 (41.4)	43.1 (43.0)

All values given in Hz. ^aValue in parentheses is experimental coupling for structural analog.

rotameric conformational combinations at C1-O1 and C3-O3 is averaged linearly, the value turns out to be 2.4 Hz. Given a dynamic range for the coupling of approximately 1.5 Hz with the minimal coupling value of 1.7 Hz, this ~0.6 Hz discrepancy between

theory and experiment seems to indicate that there is a potential skew in the populational profiles of the C2-C1-O1-H and the C2-C3-O3-H dihedrals away from 180°. This is surprising in light of the fact that C2-C1-O1-H = 180° is the preferred *exo*-anomeric configuration for β-D-glucose. This conclusion is supported by an analogous comparison between theory and experiment for $^1J_{C1,C2}$ in **12** (Table 7.3); the theoretical value for $^1J_{C1,C2}$ that results from an averaging of the couplings obtained for all 9 staggered conformational combinations of the C1-O1 and C3-O3 rotamers is 41.1 Hz, ~0.8 Hz larger than the experimental value of 40.3 Hz.

The relative effects on $^2J_{C1,C3}$ of deoxygenation at C2 versus C3 were explored by comparing both the theoretical and experimental data for this coupling as a function of the relevant hydroxyl substitutions and configuration (data contained in Chart 7.1). From this comparison it can be seen that deoxygenation at either C2 or C3 has approximately the same effect on the $^2J_{C1,C3}$ within Groups 1 or 3, however within the Group 2 compounds (compounds with opposing configurations at C1 and C3) there is a distinct difference between deoxygenation at C2 versus C3 as a function of the anomeric configuration. The beta anomers of 3-deoxy pyranoses display a ~3 Hz more positive $^2J_{C1,C3}$ than the corresponding alpha anomers. This difference is borne out by comparison to predictions obtained from the projection resultant¹¹. The correlation between both experimental and DFT $^2J_{C1,C3}$ values in **1 – 8** and the corresponding C2 and C3 deoxy derivatives with the appropriate calculated projection resultant values is linear (Figure 7.3). In the Group 2 compounds the relative contributions to the projection resultant and thereby the *J* from O1 versus O3 counter-balance each other, however when

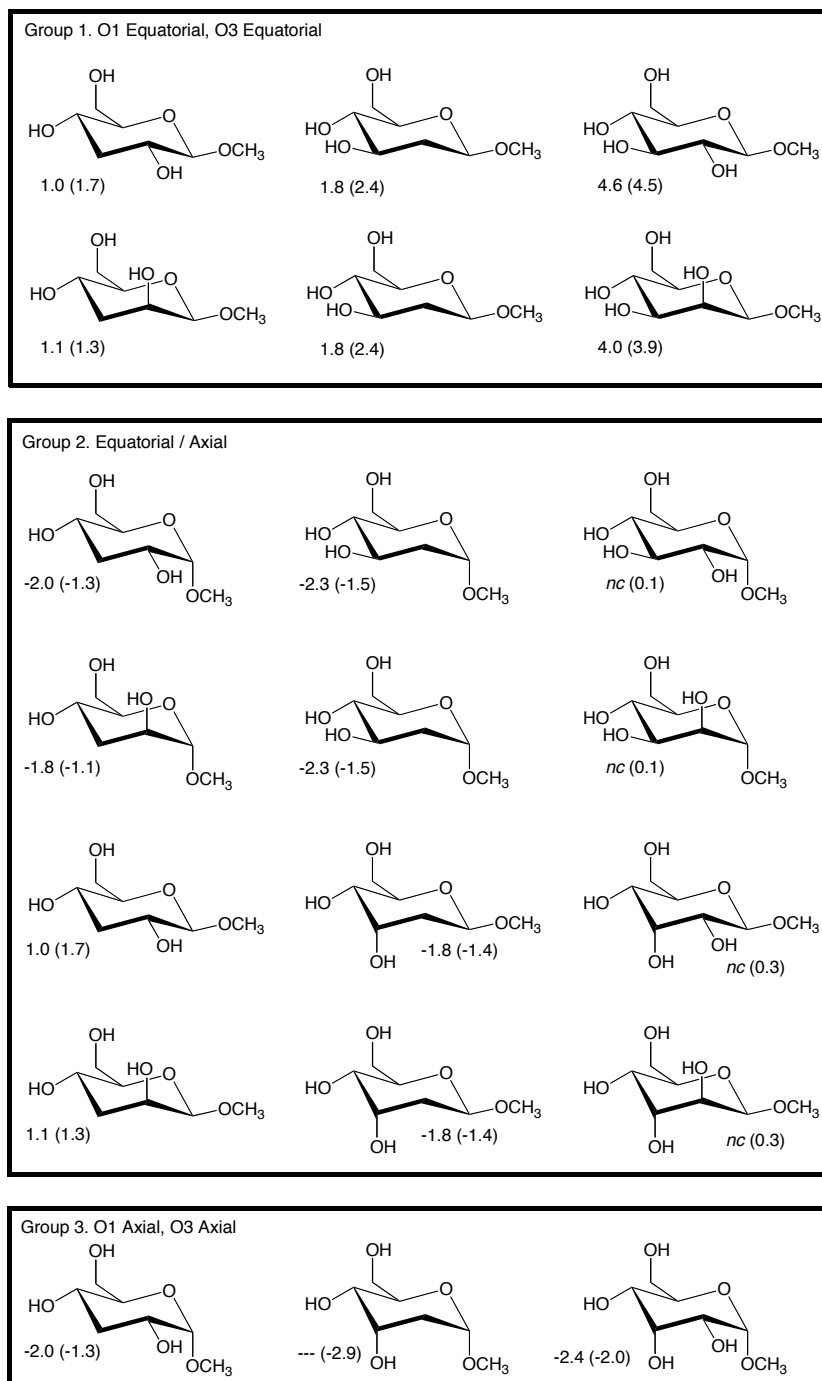


Chart 7.1. $^2J_{C1,C3}$ as a Function of C1 – C3 Configuration and C2 – C3 Substitution. All values in Hz. The first value is the experimental coupling and the value in parenthesis is the theoretical value averaged over all sampled conformers. Entries of *nc* or ‘---’ indicate respectively that either ‘no coupling’ was observed or that the coupling hasn’t been measured experimentally.

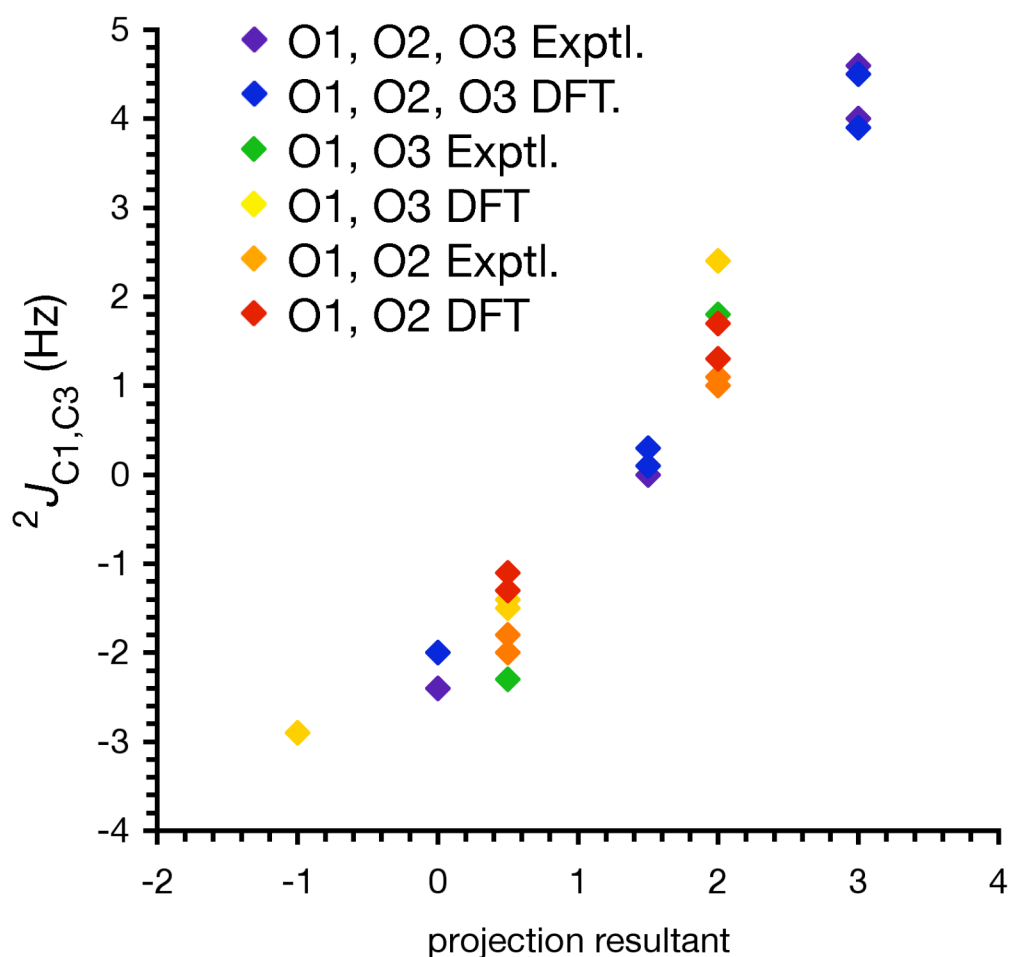


Figure 7.3. Projection Resultant for $^2J_{C1,C3}$ as a Function of C1, C2 and C3 Hydroxyl Substitution. The projection resultant is a unitless measure. C1 through C3 hydroxyl substitutions patterns are indicated by color where 'Exptl.' denotes an experimentally measured coupling and 'DFT' an averaged ab initio value. C1 through C3 hydroxyls not mentioned are absent (i.e. O1, O3 indicates a 2-deoxy structure). Data is not sorted by configuration.

the O3 hydroxyl is removed the anomeric configuration is allowed to dictate the magnitude of the coupling. The configuration of O2 does not demonstrate a similar effect on the coupling due to the symmetry of the axial and equatorial configurations of this hydroxyl with respect to the projection.

7.5.4. The Effect of *N*-acetylation on $^2J_{CC}$

A sense of the importance and ubiquity of sugar *N*-acetylation can be gained by noting that in the wwPDB there are currently >2,000 entries which contain at least one occurrence of a pyranose ring *N*-acetylated at the C2 position from amongst the >35,000 total entries²⁸. Despite this there were only 3 examples in which the *N*-acetyl group had the *manno*- configuration at C2. Therefore we explored the dependence of $^2J_{C1,C3}$ as a function of the C1-C2-N-C $_{\alpha}$ dihedral in models of the alpha and beta anomers of glucosamine and allosamine (**14^C** - **17^C**; Series 6, Scheme 7.1) with DFT and compared the average values with the appropriate experimental values (Figure 7.4). Both the experimental and averaged DFT $^2J_{C1,C3}$ values sorted as a function of C1 and C3 hydroxyl configuration in much the same way as the analogous compounds bearing a hydroxyl group at C2. The average DFT values for **15^C** and **16^C** are in good agreement with experiment, however the averaged theoretical values for **14^C** and **17^C** were both ~1 Hz more negative than the experimental values. This is likely due to the fact that all DFT calculations were conducted with the hydroxyl protons of O1 and O3 oriented anti to the coupling pathway, and as such the complete dynamic range of the coupling was probably not explored. The largest $^2J_{C1,C3}$ dynamic range, 5.2 Hz, was found in **15^C**, whereas a more modest dependence upon C1-C2-N-C $_{\alpha}$ was observed for **14^C**, **17^C**, and **16^C**; 2.3, 3.2 and 1.3 Hz respectively. When these ranges are compared respectively to $^2J_{C1,C3}$ ranges for **2^C**, **1^C**, **6^C** and **5^C** in which the O1 and O3 hydroxyl protons are both anti to C2, substitution of a *N*-acetyl has the effect of approximately doubling the dynamic range

as a function of C2 substituent conformation. A further analysis of the effect of *N*-acetyl side-chain conformation on spin-spin coupling can be found elsewhere⁷⁹.

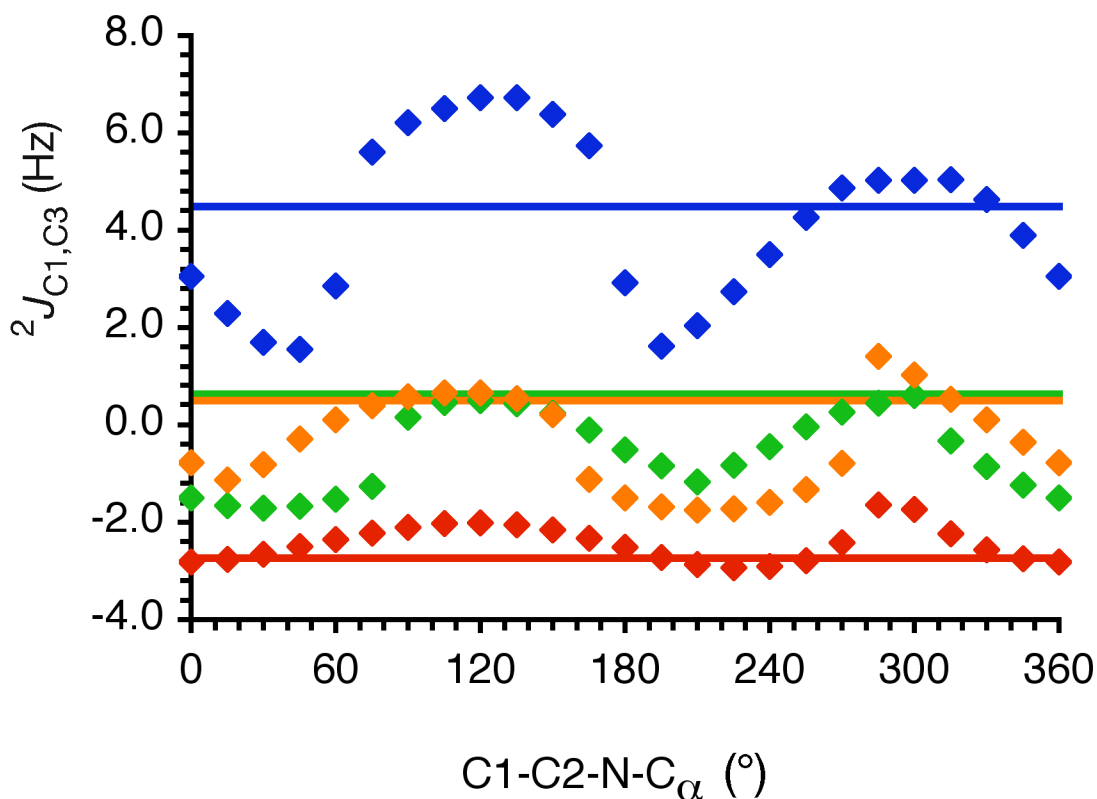


Figure 7.4. Comparison of Experimental and DFT ${}^2J_{C1,C3}$ Values in **14** - **17**. Diamonds represent *ab initio* coupling values from the Series 6 calculations, whereas the horizontal lines are experimental couplings. Color representations are blue: b-GlcNAc (**15**), green: a-GlcNAc (**14**), orange: b-AlloNAc (**17**), and red: a-AlloNAc (**16**).

7.5.5. C-C Bond Angle and Bond Length Considerations Regarding ${}^2J_{CC}$

The theoretical relationship between bond angle and ${}^2J_{CC}$ has been reviewed^{26, 27} and determined to be essentially invariant within the range of intra-annular C-C-C bond angles explored in this study. Indeed, from the comparison of ${}^2J_{C1,C3}$ versus C1-C2-C3 bond angle for **1C** – **8C** found in Figure 7.5A it can be seen that there is essentially no

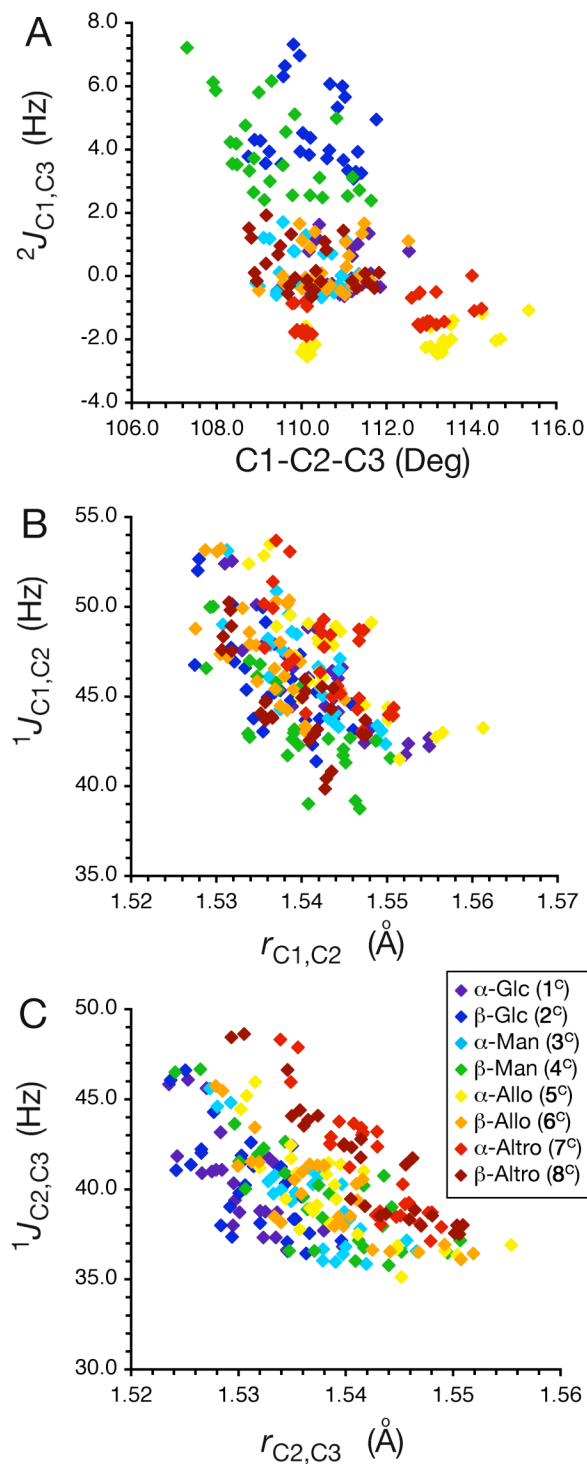


Figure 7.5. The Subtended Bond Angle and Length Dependence of $^2J_{C1,C3}$, $^1J_{C1,C2}$ and $^1J_{C2,C3}$ in 1^C - 8^C . Geometric parameters and spin-spin coupling constants derived from Series 1 calculations. Colors indicate C1 through C3 hydroxyl configuration as per legend.

correlation between these two parameters within the range of angles present ($107.3^\circ - 115.4^\circ$). However, one interesting feature of this plot is that the C1-C2-C3 bond angles for **5C** and **7C** separate into two distinct regions; the cluster around $\sim 110^\circ$ is characterized by the presence of a O1-H-O3 hydrogen bond beneath the ring, thereby compressing the bond angle, and a second cluster around $\sim 113.5^\circ$ in which the hydrogen bond is absent, allowing the bond angle to open up under the influence of the di-axial steric interactions. In panels B and C of Figure 7.5, the effect of carbon-carbon bond length on $^1J_{CC}$ is shown for **1C** – **8C**. A previously noted¹¹ trend showing a decrease in $^1J_{CC}$ as r_{CC} lengthens is qualitatively maintained across the data set, however the broad scatter in the data indicate that the correlation is not direct, and that both parameters are each independently correlated with the MO electron distribution. Put another way, *the distance between two directly bound carbons, and the scalar coupling between them are not necessarily equally subject to the same aspects of the molecular context*. For example, in **1C** – **8C** both $^1J_{C1,C2}$ and $^1J_{C2,C3}$ are accentuated when the hydroxyl proton of O2 is anti-periplanar to the respective coupling pathway, however the connection between O2 hydroxyl conformation and C1-C3 bond angle or either $r_{C1,C2}$ or $r_{C2,C3}$ is weak at best and dependent upon both the relative configuration of O2 with respect to O1 and O3 and upon the conformation of the other two vicinal hydroxyls (data not shown – or maybe supp info). This result is consistent with previous findings that a *cis* orientation of the lone pair electrons on a heteroatom directly attached to a coupled carbon add a positive contribution to the $^1J_{CC}$ whereas a *trans* orientation results in a negative contribution whose magnitude is dependent upon the identity and hybridization state of the

heteroatom^{4b,64c}. Gil and coworkers⁶⁵ provided an early account of how through space interactions between the lone pair electrons and carbon-carbon bonding or anti-bonding orbitals can lead to these positive or negative contributions respectively.

In an attempt to determine the dependence of $^2J_{CC}$ on subtended bond lengths, a comparison between $^2J_{C1,C3}$ in **1C**, **2C** and **5C** as a function of the sum of $r_{C1,C2}$ and $r_{C2,C3}$ can be found in Figure 7.6. The scatter in the data predominates and reveals nothing but a very rough correlation between larger J 's and an increasing sum of the subtended bond lengths. This finding is in accord with the conclusions of Cremer and coworkers based upon their J-OC-PSP method of theoretical spin-spin coupling analysis.⁴⁰

In a further attempt to uncover correlations between hydroxyl group conformation and C-C bond length and angle, a survey of the crystal structures in the Cambridge Structural Database (CSD) for all monosaccharide pyranoses that had one of the 8 possible configurations for C1 through C3 was undertaken. This study revealed 60 distinct structures characterized by a variety of derivatizations such as the presence of methyl-ethers or acetylation, however the chemical diversity of the compounds studied, the presence of distinctive crystal packing forces, and the effect of other geometric pressures unique to the unit cell of each individual crystal structure led to an intrinsic problem in using crystal data to draw clean correlations between C-O rotamer conformation and other structural parameters such as bond length.

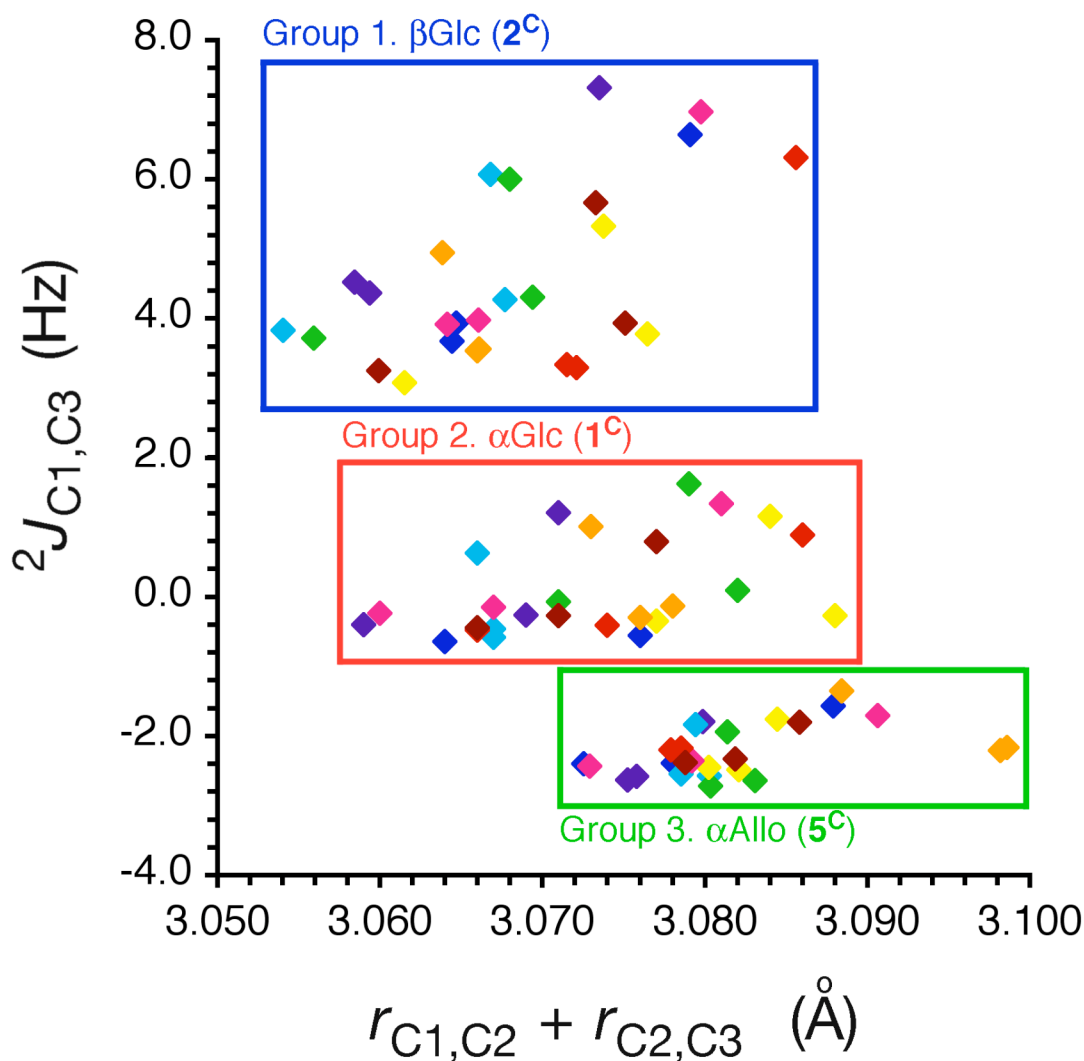


Figure 7.6. Dependence of $^2J_{C1,C3}$ on Additive Subtended Bond Lengths for 1^C , 2^C and 5^C . Color scheme for data is identical to Figure 7.1.

7.5.6. NBO Analysis: Lone Pair Donation into C-C σ^* Orbitals

Natural Bond Orbitals (NBOs) depict the Lewis-like molecular bonding pattern of electrons as a set of optimally condensed and orthonormal localized few-center orbitals^{28,29a}. In this system, the NBO analysis can be qualitatively verified by noting how well it describes the more conventional stereo-electronic phenomena in saccharides, the

endo- and *exo*-anomeric effects^{53b}. Let us start by examining the *exo*-anomeric effect as this will allow a comparison of electron densities between the presence and absence of the effect within the same molecule. The *exo*-anomeric effect is conventionally described in pyranose rings as the delocalization of density from one of the lone pair electrons on O1 into the C1-O5 anti-bonding orbital ($\sigma_{C1,O5}^*$) when the lone pair in question is aligned with this orbital and anti-periplanar to the C1-O5 bond; this is possible regardless of anomeric configuration. Indeed the NBO analyses for the 27 conformations of **1C** and **2C** demonstrate a tight sorting of the $\sigma_{C1,O5}^*$ orbital occupancy as a function of C2-C1-O1-H into low and high-density regions (Figure 7.7A-B). The C2-C1-O1-H value that orients the hydroxyl proton anti to the C1-O5 bond, -60° and 60° for **1C** and **2C** respectively, has an average occupancy of 0.020, whereas the other two rotamers in which one of the lone pairs is roughly aligned with the $\sigma_{C1,O5}^*$ orbital and the *exo*-anomeric effect is present have average occupancies of 0.030 (an orbital occupancy of 1.0 is defined as the orbital in question occupied by 1 electron).

The *endo*-anomeric effect, conventionally described as the donation of lone pair electron density from the ring oxygen into the $\sigma_{C1,O1}^*$ orbital for an axial O1 hydroxyl in 4C_1 pyranose rings, is equally well reproduced by the NBO analyses. The *endo*-anomeric effect is not possible for 4C_1 β -pyranoses, nor is it dependent upon O1 conformation. The $\sigma_{C1,O1}^*$ orbital occupancy for the 27 conformations of **1C** and **2C** do not vary significantly as a function of C2-C1-O1-H (Figure 7.7C-D), and cluster closely about the average values of 0.027 for **1C** in which the *endo*-anomeric effect is present and 0.17 for **2C** in which it is not. Thus, in accord with previously published results^{53a}, the NBO

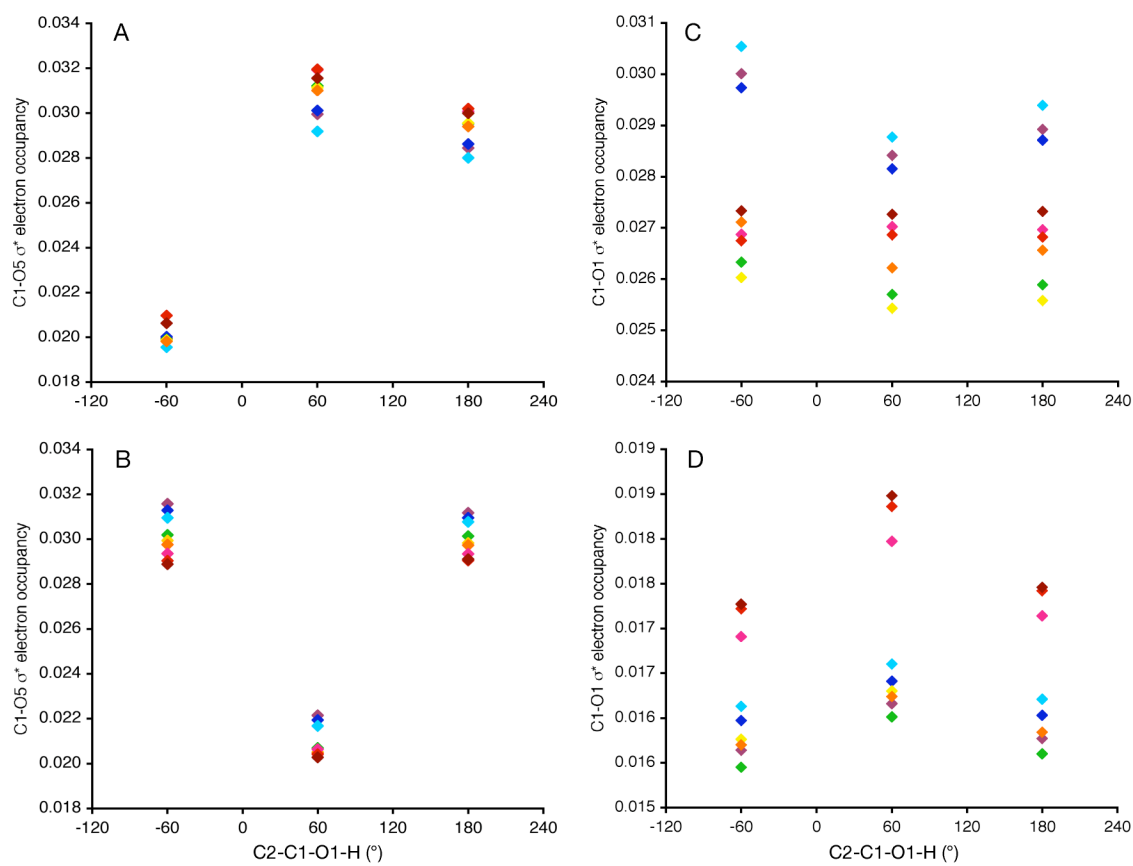


Figure 7.7. Electron Occupancy in C-O σ^* NBO's in **1C** and **2C** as a Function of C2-C1-O1-H Dihedral. Panels A and C depict data for **1C** and panels C and D contain data for **2C**. The labeling color scheme is the same as Figure 7.1. The clustering in panels A and B are indicative of the exoanomeric effect, whereas the relative ranges in the electron occupancy in panels C and D is indicative of the presence (panel C) or absence (panel D) of the endoanomeric effect.

analyses do a good job of representing the lone pair donations into acceptor σ^* orbitals inherent in the *endo*- and *exo*-anomeric effects, and should do an equally reasonable job of describing other electron delocalizations related to hydroxyl group conformation present in the model compounds.

As the effect of O2 conformation on $^2J_{C1,C3}$ was not mediated directly through any simple combination of bond angle or bond length effects, an attempt was made to

assess systematic shifts in localized electron densities as a function of hydroxyl conformation. Due to the high degree of auto correlation in the relatively large set of NBOs sensitive to the structural variations in this study, a proper orthogonal decomposition of the orbital occupancies into a maximally condensed set of principal component axes was employed to identify the salient orbital interactions. This is the principal component analysis (PCA) of hydroxyl conformation in the abstract orbital occupancy space. As a consequence of the size of the basis set used ([5s2p1d|3s1p]) there were 242 NBOs whose interaction as a function of conformation can be assessed from the correlation coefficients in the cross products matrix from the PCA. In all cases over 90% of the variance was described by the first six principal component axes. These axes were linearly regressed with respect to the direct DFT ${}^2J_{C1,C3}$ coupling values, and an iterative backward stepping process of eliminating non-significant NBOs from the PCA to achieve optimal regressions in a minimal set of PCA axes led to the identification of the orbital occupancies which in ensemble most directly correlated with the variation in ${}^2J_{C1,C3}$ (Figure 7.8). Panels A and B of Figure 7.8, which contain the first and second, and first and third principal component axes respectively for the NBO occupancies across the data set for 2C , indicate that the relative occupancy of the $O2p_{\sigma}$, $\sigma_{C1,C2}$ and $\sigma_{C2,C3}$ NBOs correlate strongly with *trans* geometries for ξ . Interestingly there is a nearly perfect inverse correlation with $O2p_{\pi}$ and the *trans* ξ geometries. In panel 8B the fact that the carbon-carbon σ^* and the $O2p_{\sigma}$ NBO principal component vectors are in the same quadrant as the *trans* ξ geometries is indicative of a correlation similar to that involving the σ_{CC} NBOs in panel A.

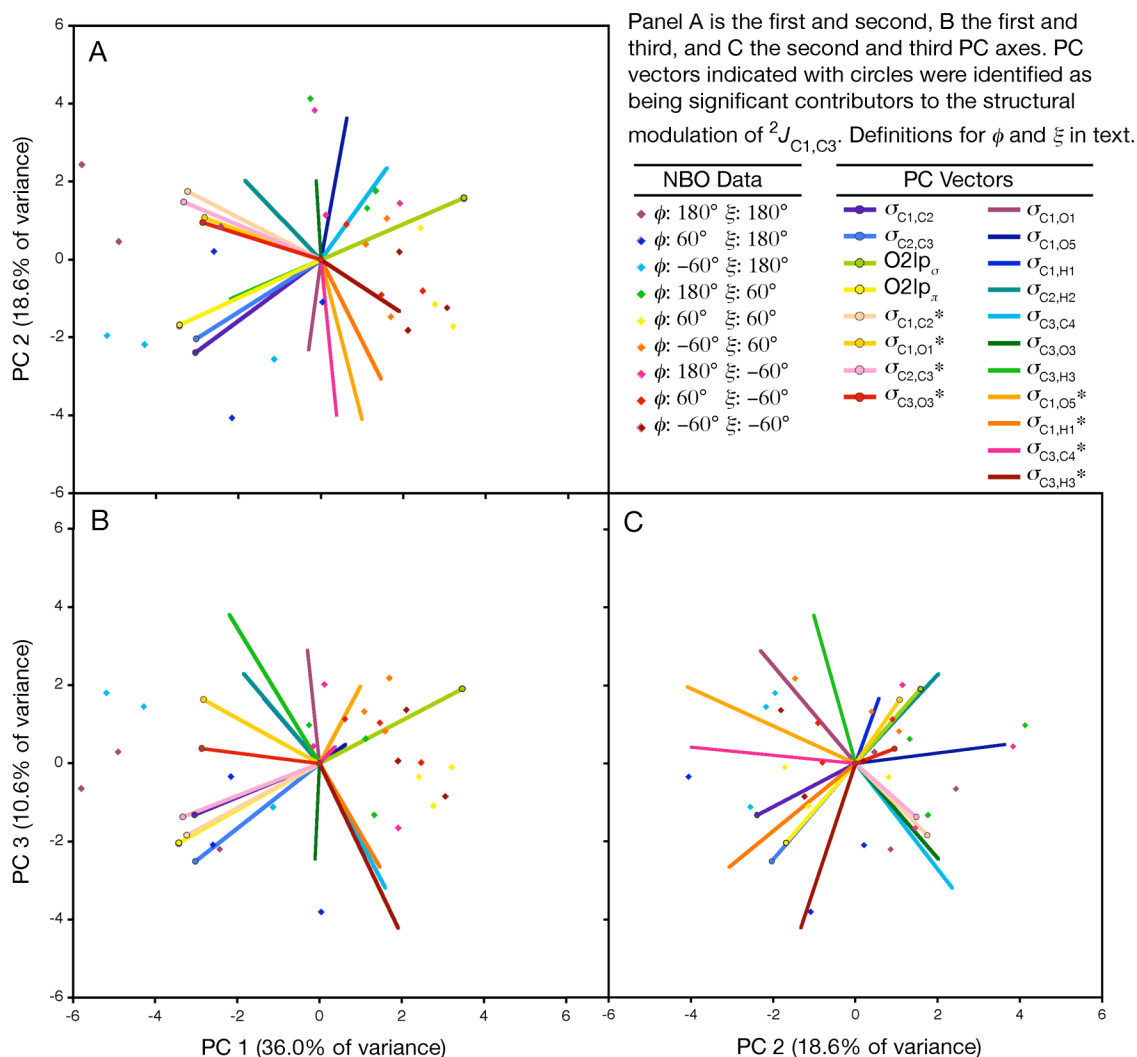


Figure 7.8. Principal Component Analysis (PCA) of NBO Occupancies in 2C

There have been numerous studies that have shown the effect of heteroatom lone pair electrons on proximal ^{13}C – ^{13}C couplings⁶⁴. These effects are known to arise from intramolecular interactions between these lone pair electrons and bonding or anti-bonding molecular orbitals^{4b}. It is not surprising that the NBO analysis of compounds **1C** – **8C** demonstrated the following trend: O2 lone pair electrons oriented anti-periplanar to a carbon-carbon bond led to increased electron occupancies in both the carbon-carbon

bonding (σ) and anti-bonding (σ^*) orbitals. The sum of either σ or σ^* C1-C2 and C2-C3 orbital occupancies scaled closely with ${}^2J_{\text{C1,C3}}$ as a function of O2 hydroxyl conformation. Furthermore the carbon-carbon bonding and anti-bonding orbitals correlate strongly in the same conformational space of the first principal component axis typified by an anti H2-C2-O2-H conformation. The necessity of C1-C2 and C2-C3 orbital summation can be rationalized based upon the symmetry of both the coupling path and those two bonds with respect to the O2 hydroxyl. Plots of the sum of the C1-C2 and C2-C3 σ and σ^* orbital occupancies versus H2-C2-O2-H for **2^C**, **1^C** and **5^C** can be found in Figure 7.9 and compared to the corresponding panels in Figure 7.1. In all cases, the H2-C2-O2-H conformation which orients the hydroxyl proton anti to H2 and a set of O2 lone pair electrons anti to both the C1-C2 and C2-C3 bonds demonstrates an accentuated electron occupancy in both carbon-carbon σ and σ^* orbitals simultaneously, which directly corresponds to a more positive ${}^2J_{\text{C1,C3}}$. The spread in the Figure 7.9 data is due to the variations in the C1 and C3 hydroxyls, which can also donate lone pair electron density into orbitals pertinent to the coupling pathway. The important distinction is that O2 is advantageously situated to significantly affect both the C1-C2 and C2-C3 orbitals in concert.

Despite the close correlation between ${}^2J_{\text{C1,C3}}$ and ξ dependent variations in electron densities in the subtended carbon-carbon σ and σ^* NBOs, the relationship alone cannot completely account for the magnitude of the ${}^2J_{\text{C1,C3}}$ dependence on the O2 dihedral. For example, in **2^C** the overlap in the ranges of electron occupancies as a function of O2 dihedral results in distinctly different J -values for identical orbital density

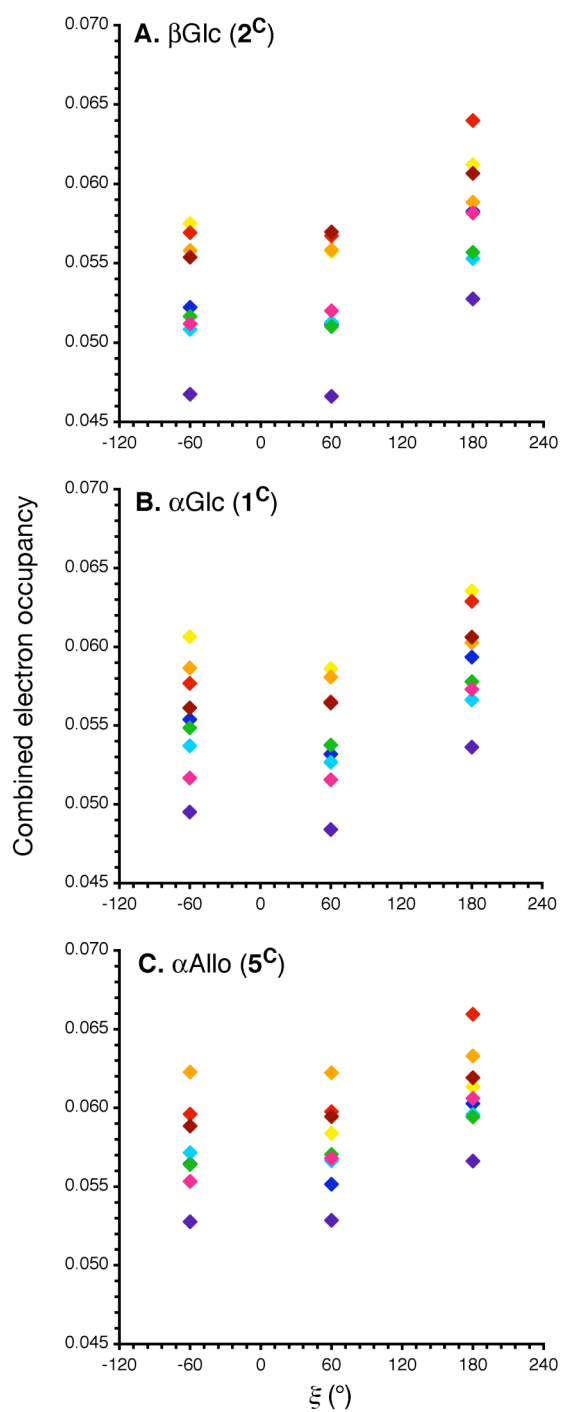


Figure 7.9. Combined NBO Electron Occupancy for 1^C , 2^C and 5^C as a Function of ξ . Combined electron occupancy is the sum of the individual occupancies for the $\sigma_{C1,C2}$, $\sigma_{C2,C3}$, $\sigma_{C1,C2}^*$ and $\sigma_{C2,C3}^*$ NBOs. Coloring scheme is identical to that in Figure 7.1. Dihedral ξ defined in text.

sums. This can be seen graphically in panel A of Figure 7.10, a normalized plot of $^2J_{C1,C3}$ versus the sum of C1-C2 and C2-C3 σ and σ^* orbital occupancies. An ideal correlation should have a slope of unity and an intercept of zero, however the data breaks up into three groupings, each with a similar near unity slope, but with distinctly different intercepts. Structurally these groupings are segregated by virtue of the combination of conformations for the O1 and O3 hydroxyls; structures with O1 and O3 hydroxyl protons anti to the coupling pathway have a positive intercept, those with O1 and O3 hydroxyl protons gauche to the coupling pathway exhibit a negative intercept, and those structures with only one of the two terminal hydroxyl groups anti to the pathway show a near zero intercept. These groupings spread the data neatly with an inverse correlation (a slope of roughly -1).

Ultimately, a set of NBOs is an attempt to parse the at times diffuse electron densities described by canonical MOs into more localized Lewis bonding, core, Rydberg, and anti-bonding orbitals. Outside of any meaning ascribed to a particular NBO, the actual quantity that correlates the conformational variation of electronegative substituents along the coupling pathway with $^2J_{C1,C3}$ is the corresponding deformation of electron density and resultant modulation of spin density transfer between the coupled nuclei. Ideally this deformation is described by the total change in the electron density across the ensemble of relevant NBOs.

Noting that the original correlation is due to vicinal lone pair delocalizations into bonding and anti-bonding orbitals involving C2 that are symmetric about the H2-C2-O2 molecular plane, an investigation of the electron occupancy of the σ and σ^* orbitals of all bonds vicinal to the O1, O2 or O3 lone pairs as a function of hydroxyl conformation

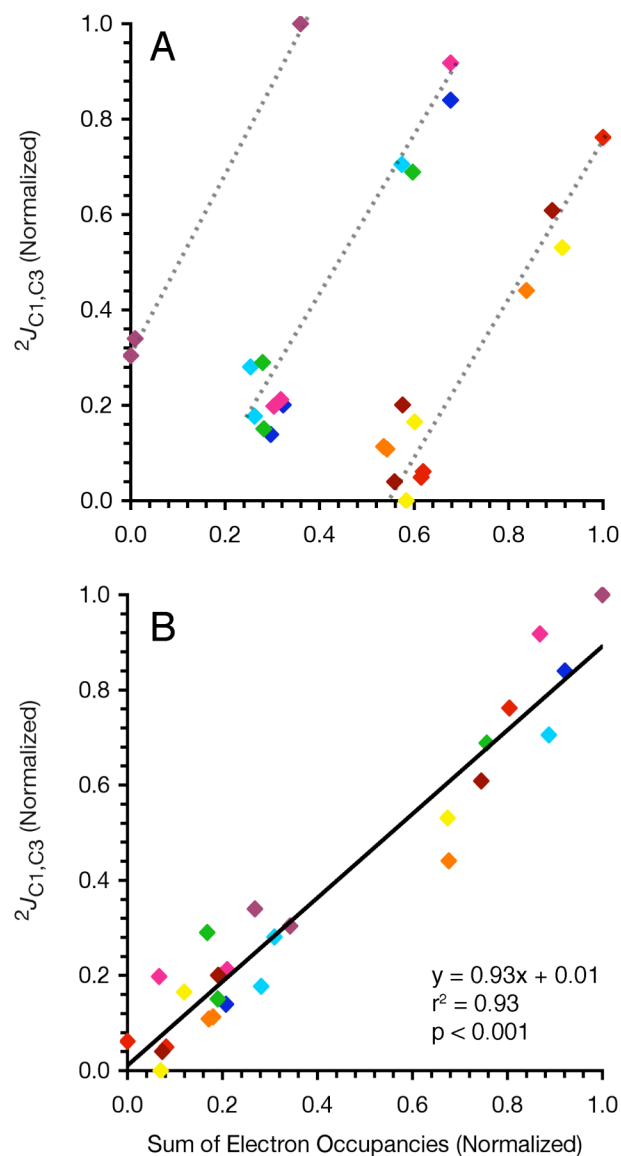


Figure 7.10. Normalized Plots of $2J_{C1,C3}$ vs. Sums of Selected NBO Occupancies in 2C. The conformational color scheme is identical to Figure 1. (A) The sC1,C2, sC2,C3, sC1,C2* and sC2,C3* NBO occupancies were summed for a given conformer. The grey dotted lines indicate geometrically related conformers. (B) The s and s* NBO occupancies for the C1-O5, C1-C2, C2-C3, C3-C4, C1-H1, C3-H3 bonds were summed along with sC2,H2 for the 27 conformers of 2C.

reveals an inverse correlation of orbital occupancy in the C2-H2 σ and σ^* orbitals with $^2J_{C1,C3}$ as a function of ξ and direct correlations between the C1-O5, C3-C4 and C1-H1, C3-H3 σ and σ^* orbital occupancies with $^2J_{C1,C3}$ as a function of proximal hydroxyl conformation (data not shown). A normalized plot of the sum of σ and σ^* electron occupancy as a function of ξ for the C1-O5, C1-C2, C2-C3, C3-C4, C1-H1, C2-H2, and C3-H3 bonds (the complete set of bonds vicinal to the O1, O2 or O3 lone pairs) maps closely onto a similarly normalized plot of $^2J_{C1,C3}$. A regression of the relationship yields a slope of 0.9, y-intercept -0.03, and $r^2 = 0.79$. This relationship can be improved to a slope of 0.9, y-intercept 0.01, and $r^2 = 0.93$ by the removal of the $\sigma_{C2,H2}^*$ orbital from the comparison (Figure 7.10B). This anti-bonding orbital could be reasonably removed from the comparison without the concomitant removal of any other orbital terms due to the placement of this orbital within the H2-C2-O2 plane of coupling pathway symmetry. In other words, $\sigma_{C2,H2}^*$ is minimally populated and does not have any direct contact with either of the coupled carbons, and thus is not an efficient vehicle for the transfer of spin density as mediated by O2 orbitals associated with variation in ξ conformation. This is not to say that this NBO is not involved in any $^2J_{C1,C3}$ coupling mechanisms, just that the ξ dependence of the coupling is not likely to involve this orbital.

7.5.7. The Influence of C-O rotation on $^2J_{C2,C4}$ in β -Glc

The effect of coupled carbon configuration on experimental $^2J_{C2,C4}$ values has been reported^{37c} and is analogous in trend to that of $^2J_{C1,C3}$ if somewhat attenuated in

magnitude. While a complete set of calculations seeking to reproduce this effect for the set of all eight hydroxylated configurations at C2 through C4 was not undertaken, a study of the behavior of $^2J_{C2,C4}$ in **9C** as a function of C2 through C4 hydroxyl rotamer conformation (Series 3) showed that when the $^2J_{C2,C4}$ for the 27 conformations generated were averaged the resultant value of 2.5 Hz was in good agreement with the experimental value of 2.8 Hz. This model with equatorial configurations at C2 through C4 was chosen due to both biological relevance and by way of comparison to $^2J_{C1,C3}$ found in **2C**, which exhibits the maximal potential coupling and maximal dependence on the central hydroxyl group conformation. Thus, the plot of $^2J_{C2,C4}$ versus H3-C3-O3-H dihedral for **9C** (Figure 7.11) is analogous to Figure 7.1A. As was the case for $^2J_{C1,C3}$, there is a strong dependency on the conformation of the hydroxyl group intermediate to the coupling pathway. Here again the spread in the coupling data for any particular H3-C3-O3-H dihedral value is due to the conformational dependency of the coupled carbon hydroxyl groups. The coupling magnitude is maximized for each of the three staggered H3-C3-O3-H rotamers when both of the hydroxyl groups appended to the coupled carbons are in plane with the coupling pathway; that is when the hydroxyl protons of both O2 and O4 are anti to C3. Table 7.4 contains a statistical comparison between $^2J_{C1,C3}$ and $^2J_{C2,C4}$ as a function of conformation of the hydroxyl group intermediate to the coupling pathway. The overall dynamic range of both couplings is identical at ~4.2 Hz. While the magnitude of the average coupling for each of the three staggered rotamers about the central hydroxyl group is approximately 2 Hz smaller for $^2J_{C2,C4}$ than $^2J_{C1,C3}$,

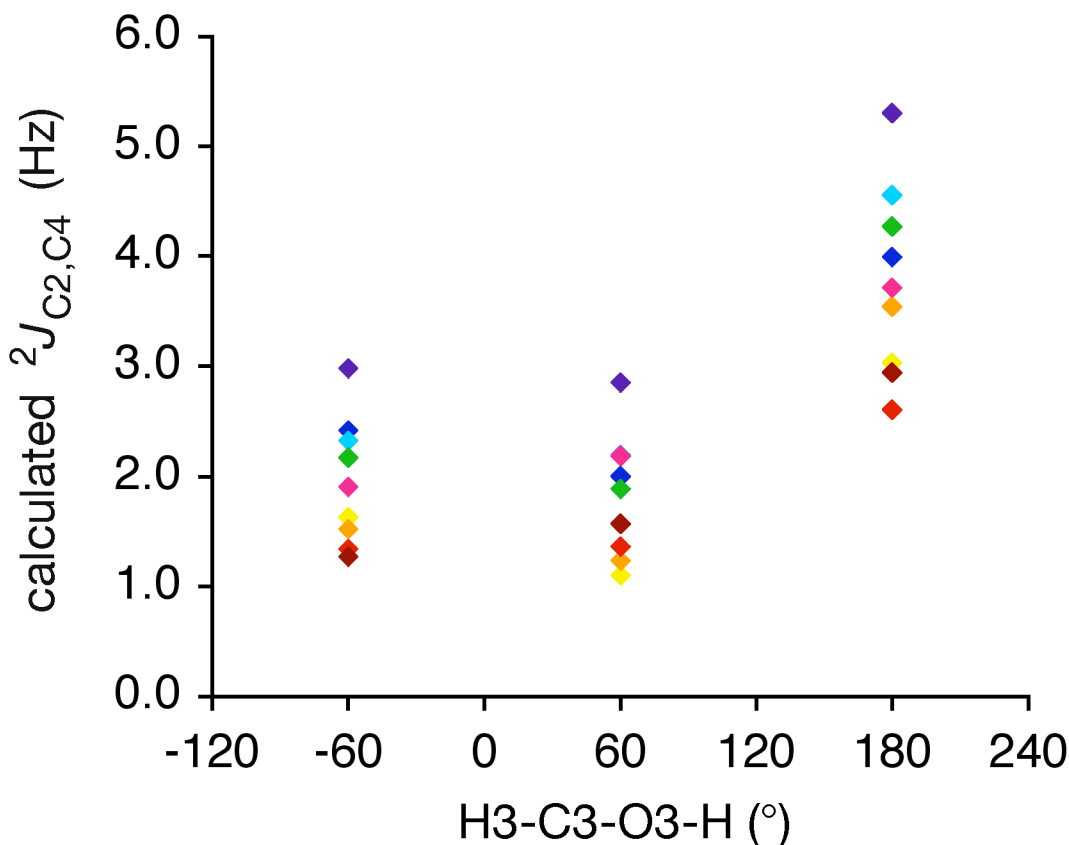


Figure 7.11. Calculated ${}^2J_{C2,C4}$ in $\mathbf{9^C}$ as a function of the H3-C3-O3-H dihedral. Color scheme is analogous to that found in Fig. 7.1, except shifted with respect to the coupling path. For example, the C1-O1 torsion in Fig. 7.1 is the C2-O2 torsion here, etc.

the coupling ranges for these three rotamers are similar between ${}^2J_{C1,C3}$ and ${}^2J_{C2,C4}$, indicating comparable secondary dependencies of these two couplings.

A linear regression of normalized ${}^2J_{C2,C4}$ data in $\mathbf{9^C}$ versus the normalized sum of the σ and σ^* orbital occupancies for all bonds vicinal to the O2, O3 or O4 lone pairs shows a correlation similar to that of ${}^2J_{C1,C3}$ in $\mathbf{2^C}$, with a slope of 0.8, y-intercept = 0.02, and $r^2 = 0.73$. Also analogous is the improvement of the correlation upon exclusion of the C3-H3 σ^* orbital from the fit (slope = 0.8, y-intercept = 0.05, and $r^2 = 0.82$).

TABLE 7.4. STATISTICAL COMPARISON^A BETWEEN CALCULATED $^2J_{C1,C3}$ IN 2^C AND $^2J_{C2,C4}$ IN 9^C .

	$^2J_{C1,C3}$ in 2^C	$^2J_{C2,C4}$ in 9^C
Overall Coupling Range	4.2	4.2
Overall Average Coupling ^b	4.5 (4.6)	2.5 (2.8)
	H2-C2-O2-H = 180°	H3-C3-O3-H = 180°
Coupling Range	2.4	2.7
Average Coupling	6.1	3.8
(O1/O2) ^c Average Coupling Range	1.3	1.5
(O3/O4) ^d Average Coupling Range	1.2	1.2
	H2-C2-O2-H = 60°	H3-C3-O3-H = 60°
Coupling Range	1.3	1.7
Average Coupling	3.8	1.8
(O1/O2) ^c Average Coupling Range	0.7	0.9
(O3/O4) ^d Average Coupling Range	0.7	0.8
	H2-C2-O2-H = -60°	H3-C3-O3-H = -60°
Coupling Range	1.3	1.7
Average Coupling	3.8	2.0
(O1/O2) ^c Average Coupling Range	0.6	1.1
(O3/O4) ^d Average Coupling Range	0.7	0.6

^aAll values given in Hz. ^bValue in parentheses is experimental coupling. ^cAverage Coupling Range in column one (data for 2^C) is the average of the three coupling ranges for C2-C1-O1-H variation given one of the three staggered C2-C3-O3-H conformations, and in column two (data for 9^C) is the average of the three coupling ranges for C3-C2-O2-H variation given one of the three staggered C2-C3-O3-H conformations. ^dAverage Coupling Range in column one (data for 2^C) is the average of the three coupling ranges for C2-C3-O3-H variation given one of the three staggered C2-C1-O1-H conformations, and in column two (data for 9^C) is the average of the three coupling ranges for C3-C4-O4-H variation given one of the three staggered C2-C1-O1-H conformations.

These results indicate the generalizability of utilizing ${}^2J_{\text{CC}}$ not only for the determination of hydroxyl configuration at the coupled carbons, but as a conformational probe of the hydroxyl intermediate to the coupling pathway. This raises the remaining question as to the relative importance that a particular NBO plays in the spin-spin coupling mechanism as measured by its individual contribution to the J . Natural J -Coupling (NJC) analysis is a well-established technique for relating the molecular electronic structure with J -coupling mechanisms to yield an intuitive understanding of coupling sign and magnitude as a function of molecular conformation.^{29c}

7.5.8. Natural J -coupling Analysis I: Configurational Dependence of ${}^2J_{\text{C1,C3}}$

The natural J -coupling (NJC) method is an analytical technique for deconvoluting the Fermi contact term of NMR spin-spin coupling into individual and donor→acceptor contributions by way of finite perturbation theory as applied to *ab initio* DFT wave functions. NJC analysis parses the total J into three distinct components, the Lewis intrabond coupling contribution ($J^{(\text{L})}$), an intra-bond repolarization contribution ($J^{(\text{repol})}$), and an inter-bond spin delocalization contribution ($J^{(\text{deloc})}$).^{29c} The $J^{(\text{L})}$ component represents that portion of the total spin density consequent with the slight perturbation dependent shift in the natural Lewis structures of the a and b spin orbitals. This can be considered a “through space” or steric transfer of spin information between bonding orbitals s_i and s_j as a result of their requisite mutual orthogonality descending from the Pauli exclusion principle.³⁹ The intrabond repolarization contribution ($J^{(\text{repol})}$) is as the name implies, a repolarization of the electron density contained within the parent s_i orbital, typically as a result of strong electron correlation phenomena involving s_i .^{29c} The

molecular structures explored in this study are typical of organic sp^3 hybridized compounds in so far as the $J^{(\text{repol})}$ term is small with respect to the $J^{(\text{L})}$ contribution. The third spin mechanism defined in NJC analysis, $J^{(\text{deloc})}$, operates via the transfer of spin density between acceptor-donor pairs of orbitals, typically $\sigma_i \rightarrow \sigma_j^*$ interactions. This is often called the “through bond” spin-coupling mechanism, a term in this context synonymous with hyper-conjugation.

In all instances the value for the FC portion of ${}^2J_{\text{C1,C3}}$ calculated by the finite field single perturbation method in NJC was within ± 0.1 Hz of the value derived by the directly calculated DFT method (the sum of the remaining three Ramsey terms was in all cases much less than ± 0.1 Hz, and as such, negligible). The perturbing field was applied to nucleus C1 and observed at C3 for all calculations, however when the field was applied to C3 and the response observed at C1 in the 27 **2C** structures, the corresponding values were identical within the range of experimental error (data not shown).

NJC analysis reveals that the difference in the magnitude and sign of ${}^2J_{\text{C1,C3}}$ as a function of coupled carbon configuration is conferred roughly equally through both the $J^{(\text{L})}$ and the $J^{(\text{deloc})}$ coupling mechanisms. Averages of the Lewis contributions to the total ${}^2J_{\text{C1,C3}}$ from the ensemble of NLMO's across the 27 conformations of **2C**, **1C** and **5C** exhibit a distinct pattern (Table 7.5). When considering the configurational shift from an equatorial O1 in **2C** to an axial O1 in **1C**, there is a 2.7 Hz and 0.9 Hz less positive contribution to the $J^{(\text{L})}$ from the $\sigma_{\text{C1,H1}}$ and $\sigma_{\text{C2,C3}}$ orbitals respectively. This can be symbolized as $\Delta_{\beta\bar{\text{G}} \rightarrow \alpha\bar{\text{G}}} {}^2J_{\text{C1,C3}}^{(\sigma_{\text{C1,H1}})} = 2.7 \text{ Hz}$ and $\Delta_{\beta\bar{\text{G}} \rightarrow \alpha\bar{\text{G}}} {}^2J_{\text{C1,C3}}^{(\sigma_{\text{C2,C3}})} = 0.9 \text{ Hz}$, where $\Delta_{\beta\bar{\text{G}} \rightarrow \alpha\bar{\text{G}}}$ represents the difference in the indicated Lewis coupling term between the

TABLE 7.5. CONFIGURATION DEPENDENCE^A OF $J^{(L)}$ IN **2^C**, **1^C** AND **5^C**.

NBO	$\Delta_{\beta G \rightarrow \alpha G} J^{(L)b}$	NBO	$\Delta_{\alpha G \rightarrow \alpha A} J^{(L)}$
$\sigma_{C1,H1}$	-2.7	$\sigma_{C3,H3}$	-1.3
$\sigma_{C2,C3}$	-0.9	$\sigma_{C1,C2}$	-0.8
$\sigma_{C3,H3}$	-0.3	$\sigma_{C1,O5}$	-0.2
$\sigma_{C3,C4}$	0.0	$\sigma_{C1,O1}$	-0.1
$\sigma_{C1,C2}$	0.0	$\sigma_{C3,C4}$	-0.1
$\sigma_{C2,H2}$	0.1	$\sigma_{C2,H2}$	0.0
$\sigma_{C1,O5}$	0.1	$\sigma_{C2,C3}$	0.0
$\sigma_{C3,O3}$	0.5	$\sigma_{C3,O3}$	0.4
$\sigma_{C1,O1}$	1.3	$\sigma_{C1,H1}$	0.7

^aAll values given in Hz. ^bValues calculated as the average of the $J^{(L)}$ from the Lewis bonding portion of the particular NBO across the 27 conformations in **1^C** minus the analogous value in **2^C** for column two, and **5^C** minus **1^C** for column four.

ensemble average of **2^C** and **1^C**. A structurally analogous trend emerges for the $\sigma_{C3,H3}$ and $\sigma_{C1,C2}$ orbitals when considering the configurational shift from an equatorial O3 in **1^C** to an axial O3 in **5^C**; $\sigma_{C3,H3}$ and $\sigma_{C1,C2}$ contribute 1.3 Hz and 0.9 Hz less positive value to the ${}^2J_{C1,C3}^{(L)}$ in **5^C** when compared to **1^C**. Given a negative magnitude shift in the total ${}^2J_{C1,C3}$ of 4.5 Hz and 2.1 Hz for the comparisons between **2^C** to **1^C**, and **1^C** to

5C, and to avoid over emphasizing the importance of these $J^{(L)}$ shifts, it is crucial to note in Table 7.5 that there are offsetting shifts to the ${}^2J_{C1,C3}^{(L)}$ which are contrary to the trend in the overall ${}^2J_{C1,C3}$. The important point is that the above-indicated orbitals are the most significant directly correlated contributors to the net change in the Lewis coupling mechanism portion of the total J .

The ${}^2J_{C1,C3}^{(L)}$ contribution to the total spin-coupling can be thought of as arising from Pauli exchange antisymmetry contributions to the molecular potential energy^{29c,30-31}. This steric exchange energy can be quantitatively approximated within the NBO framework as the energy difference between the preorthogonal and the fully orthogonalized NBOs. This steric energy has been shown to scale linearly with the corresponding Lewis coupling terms. When a full natural steric analysis (NSA) was conducted on the 27 conformers of **2C**, **1C** and **5C**, a complex series of variation emerged in the pairwise steric exchange energies ($\Delta E_{i,j}$) involving $\sigma_{C1,H1}$ or $\sigma_{C3,H3}$ with other neighboring NBOs given the respective changes in configuration at either C1 or C3. The largest contribution to this variation was due to interactions with vicinal and 1,3-diaxial $\sigma_{C,H}$ or O5lp NBOs (data not shown). In contrast, $\Delta_{\beta\bar{G} \rightarrow \alpha\bar{G}} {}^2J_{C1,C3}^{(\sigma_{C2,C3})}$ and $\Delta_{\alpha\bar{G} \rightarrow \alpha\bar{A}} {}^2J_{C1,C3}^{(\sigma_{C1,C2})}$ track closely with $\Delta_{\beta\bar{G} \rightarrow \alpha\bar{G}} E_{\sigma_{C2,C3}, \sigma_{C1,O1}} = -1.0$ kcal/mol and $\Delta_{\alpha\bar{G} \rightarrow \alpha\bar{A}} E_{\sigma_{C1,C2}, \sigma_{C3,O3}} = -1.7$ kcal/mol respectively. This can be visualized through a comparative inspection of the pre-orthogonal NBO (PNBO) interactions between $\sigma_{C2,C3}$ and $\sigma_{C1,O1}$ for **2C** and **1C** (Figure 7.12A) and between $\sigma_{C1,C2}$ and $\sigma_{C3,O3}$ for **1C** and **5C** (Figure 7.12B).

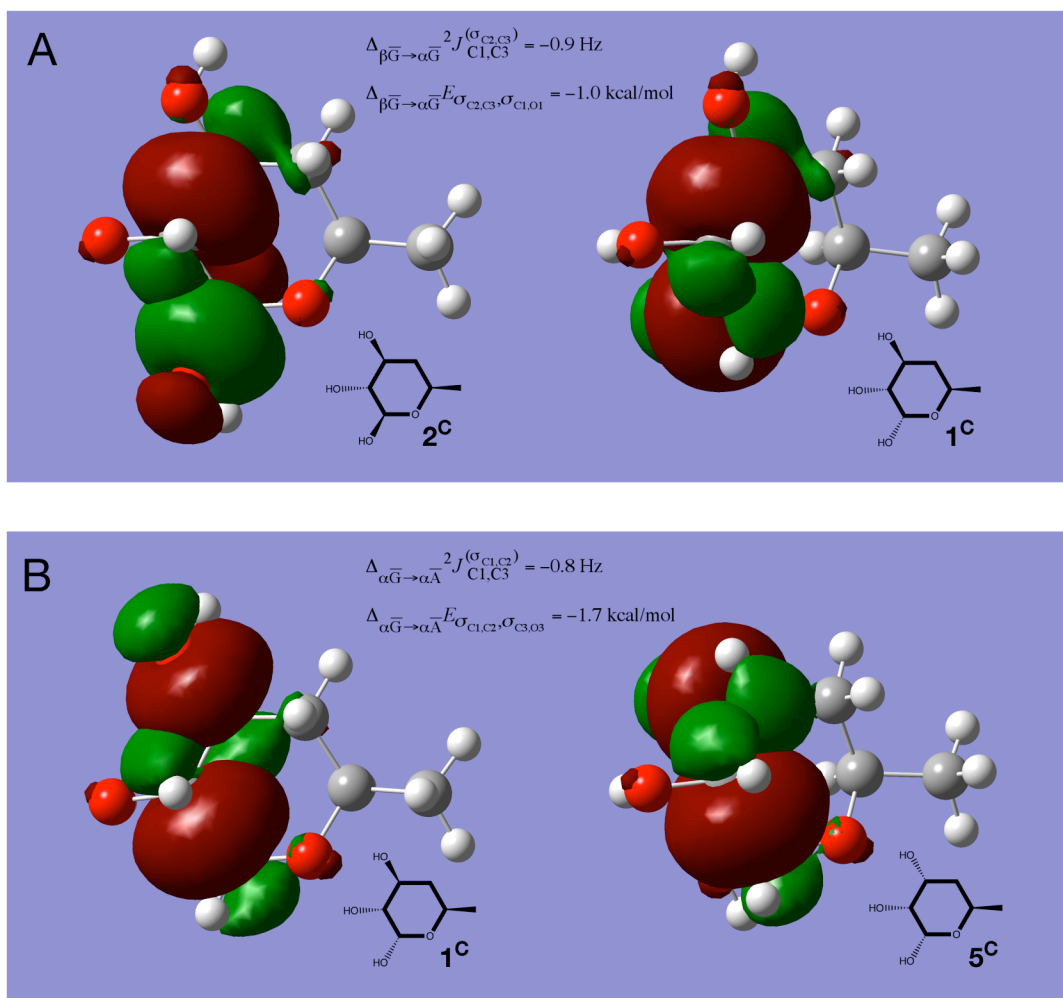


Figure 7.12 Configuration Dependent Variation in the Vicinal $J(\sigma_{C,C})$. The differences in the pairwise steric exchange energy between $\sigma_{C2,C3}$ and $\sigma_{C1,O1}$ (A) and $\sigma_{C1,C2}$ and $\sigma_{C3,O3}$ (B) are significant contributors to the configuration dependent changes in the magnitude of $2J_{C1,C3}$ between the epimeric pairs **2C** and **1C**, and **1C** and **5C** respectively. The orbitals depicted are pre-orthogonal NBOs. The bonding (σ) orbitals have three lobes with a node at each of the bonded carbon or oxygen nuclei. The coloration denotes orbital phasing and is consistent only within an individual PNBO. The schematic diagrams show the relative hydroxyl orientations obscured by the presence of orbitals.

The hyperconjugative $\sigma_i \rightarrow \sigma_j^*$ or “through bond” coupling contribution ($J^{(\text{deloc})}$) is the other significant mechanism for the transfer of spin density in this system. This mechanism involves pair-wise donor \rightarrow acceptor orbital interactions. In this system donors consist of NBOs that include core carbon or lone pair orbitals, while the acceptor orbitals were all antibonding orbitals (as opposed to long range Rydberg (γ^*) orbitals). A summary of the significant donor \rightarrow acceptor interactions as a function of C1 and C3 configuration can be found in Table 7.6. Upon going from beta (**2C**) to alpha Glucose (**1C**), $\sigma_{\text{C1},\text{O1}}^*$ contributed an average of 3.5 Hz less (more negative value) to the $J^{(\text{deloc})}$ portion of the total coupling. The significantly more positive coupling contribution from $\sigma_{\text{C1},\text{O1}}^*$ in **2C** versus **1C** is due to a $J_{\text{vic}}^{(\text{deloc})}$ coupling mechanism. The subscript ‘vic’ emphasizes the vicinal aspect of the delocalization. This was created by an overlap favorable to the hyperconjugative interaction $\sigma_{\text{C2},\text{C3}} \rightarrow \sigma_{\text{C1},\text{O1}}^*$, which, as a result of the axial orientation of the O1 hydroxyl in α -Glc, is present in **2C** and absent **1C** (Figure 7.13A). The structurally analogous hyperconjugation $\sigma_{\text{C1},\text{C2}} \rightarrow \sigma_{\text{C3},\text{O3}}^*$ preferentially facilitated in **1C** by the equatorial O3 over **5C** is responsible for 1.8 Hz of the 2.7 Hz difference in the total contributions to $J^{(\text{deloc})}$ from $\sigma_{\text{C3},\text{O3}}^*$ between these two structures (Figure 7.13B). These differences can be represented symbolically as $\Delta_{\beta\bar{\text{G}} \rightarrow \alpha\bar{\text{G}}}^2 J_{\text{C1},\text{C3}}^{(\sigma_{\text{C2},\text{C3}} \rightarrow \sigma_{\text{C1},\text{O1}}^*)} = -1.9 \text{ Hz}$ for the most important shift in donor \rightarrow acceptor mediated transfer of spin density associated with C1 configuration and $\Delta_{\alpha\bar{\text{G}} \rightarrow \alpha\bar{\text{A}}}^2 J_{\text{C1},\text{C3}}^{(\sigma_{\text{C1},\text{C2}} \rightarrow \sigma_{\text{C3},\text{O3}}^*)} = -1.8 \text{ Hz}$ likewise for C3 configuration (Figure 7.13). The

bar over the letters representing the saccharide identities indicates that the difference is between values that were averaged over the respective ensembles of 27 conformers.

TABLE 7.6. CONFIGURATION DEPENDENCE^a OF $J^{(\text{DELOC})}$ IN **2C**, **1C** AND **5C**.

Acceptor	Donor ^b	$\Delta_{\beta\text{G}\rightarrow\alpha\text{G}} J^{(\text{deloc})\text{c}}$	Acceptor	Donor	$\Delta_{\alpha\text{G}\rightarrow\alpha\text{A}} J^{(\text{deloc})}$
$\sigma_{\text{C1},\text{O1}}^*$	Total	-3.5	$\sigma_{\text{C3},\text{O3}}^*$	Total	-2.7
	$\sigma_{\text{C2},\text{C3}}$	-1.9		$\sigma_{\text{C1},\text{C2}}$	-1.8
	$\sigma_{\text{C1},\text{C2}}$	-0.7		$\sigma_{\text{C2},\text{C3}}$	-0.5
	C1_{cr}	-1.0		C3_{cr}	-0.8
	O2_{lp}	-0.2		O2_{lp}	-0.2
$\sigma_{\text{C3},\text{O3}}^*$	Total	-1.1	$\sigma_{\text{C1},\text{C2}}^*$	Total	-0.5
	$\sigma_{\text{C1},\text{H1}}$	-0.5		$\sigma_{\text{C1},\text{H1}}$	-0.2
	$\sigma_{\text{C1},\text{C2}}$	-0.3		O2_{lp}	-0.1
	C3_{cr}	-0.4			
$\sigma_{\text{C2},\text{C3}}^*$	Total	-0.5	$\sigma_{\text{C1},\text{H1}}^*$	Total	-0.4
	C3_{cr}	-0.3		$\sigma_{\text{C2},\text{C3}}$	-0.2
$\sigma_{\text{C3},\text{H3}}^*$	Total	0.1	$\sigma_{\text{C3},\text{C4}}^*$	Total	0.2
$\sigma_{\text{C1},\text{O5}}^*$	Total	0.1	$\sigma_{\text{C1},\text{O1}}^*$	Total	0.3
$\sigma_{\text{C3},\text{C4}}^*$	Total	0.3	$\sigma_{\text{C1},\text{O5}}^*$	Total	0.3
$\sigma_{\text{C1},\text{C2}}^*$	Total	0.8	$\sigma_{\text{C2},\text{C3}}^*$	Total	0.8
$\sigma_{\text{C1},\text{H1}}^*$	Total	2.0	$\sigma_{\text{C3},\text{H3}}^*$	Total	1.7

^aAll values given in Hz. ^bOnly donor orbital contributions of $\leq |0.1|$ Hz that directly correlated with the overall configurational trend in $^2J_{\text{C1},\text{C3}}$ were included. ^cValues calculated as the average of the particular $\sigma_i \rightarrow \sigma_j^*$ delocalization across the 27 conformations in **1C** minus the analogous value in **2C** for column three, and **5C** minus **1C** for column six. Indicated ‘Total’ values are similar differences for $\sigma_{\text{total}} \rightarrow \sigma_j^*$.

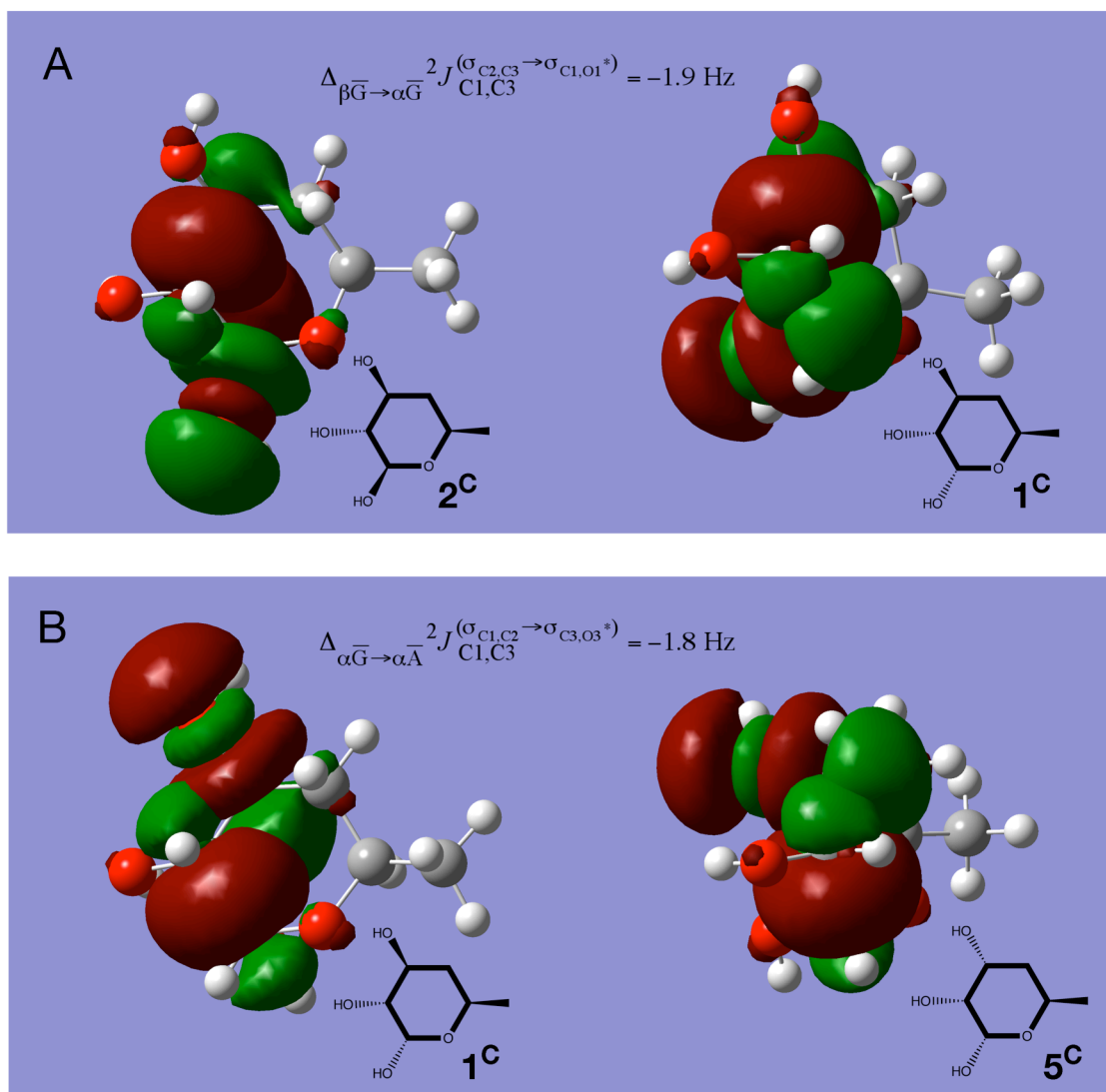


Figure 7.13. Configuration Dependence of ${}^2J_{C1,C3}$ as Mediated by Vicinal J_{deloc} . The differences in orbital interactions $\sigma_{C2,C3} \rightarrow \sigma_{C1,O1}^*$ (A) and $\sigma_{C1,C2} \rightarrow \sigma_{C3,O3}^*$ (B) are the largest contributors to the configuration dependent changes in the magnitude of ${}^2J_{C1,C3}$ between the epimeric pairs **2C** and **1C**, and **1C** and **5C** respectively. The bonding (σ) orbitals have three lobes with a node at each of the bonded carbon nuclei, and the antibonding (σ^*) orbitals have four lobes with a node each at the bonded carbon and oxygen nuclei, as well as a third node between them. The coloration denotes orbital phasing and is consistent only within an individual PNBO. The schematic diagrams show the relative hydroxyl orientations obscured by the presence of orbitals.

Any spin-coupling mechanism can be thought of as a channel for the transfer of spin density information between coupled nuclei, in this case carbons 1 and 3. The more robust the “electronic path” between the nuclei, the more efficient is the transfer of spin density via that mechanism. The orbital depictions of the vicinal $^2J_{C1,C3}^{(\sigma_{CC} \rightarrow \sigma_{CO}^*)}$ mechanisms in Figure 7.13 provide an intuitive picture of how information about the spin state of one coupled carbon may reach the other, and how this mechanism is acutely sensitive to the relative spatial dispositions of the involved orbitals. For example, in the left of panel 13A the central lobe of the bonding C2-C3 orbital in **2C** has a large overlap with the back lobe of the C1-O1 anti-bonding orbital, whereas these two lobes have a much more limited interaction in **1C** due to the axial orientation of the C1-O1 bond vector. There is an interaction between the back lobe of $\sigma_{C2,C3}$ with the back lobe of $\sigma_{C1,O1}^*$ in **1C**, however there is a much more direct link between the coupled nuclei in **2C** via the central lobe of $\sigma_{C2,C3}$ and the back lobe of $\sigma_{C1,O1}^*$. There is an additional interaction in **2C** between the back lobe of $\sigma_{C2,C3}$ with the larger of the two central lobes of $\sigma_{C1,O1}^*$, which certainly contributes to the observed effect. An analogous argument involving $\sigma_{C1,C2}$ and $\sigma_{C3,O3}^*$ applies to the comparison of **1C** and **5C** in panel 13B.

While the steric $\mathcal{J}^{(L)}$ contribution to the total coupling can be related to destabilizing or repulsive interactions between orbitals, the $\mathcal{J}^{(deloc)}$ contribution can be similarly related to energetic stabilization connected specifically with $\sigma_i \rightarrow \sigma_j^*$ delocalization^{29c,31}. The magnitude of this energy lowering delocalization ($E_{i \rightarrow j^*}$) can be accurately approximated by a second order perturbation theory analysis of the Fock matrix in the NBO basis³¹. The only second order perturbation energies involving carbon-

carbon bonding orbitals as donors that tracked systematically with the configurational dependence of ${}^2J_{\text{C1,C3}}^{(\sigma_{\text{C,C}} \rightarrow \sigma_{\text{C,O}}^*)}$ were $\Delta_{\beta\bar{\text{G}} \rightarrow \alpha\bar{\text{G}}} E_{\sigma_{\text{C2,C3}} \rightarrow \sigma_{\text{C1,O1}}^*} = -1.4$ kcal/mol and $\Delta_{\alpha\bar{\text{G}} \rightarrow \alpha\bar{\text{A}}} E_{\sigma_{\text{C1,C2}} \rightarrow \sigma_{\text{C3,O3}}^*} = -1.7$ kcal/mol, thus validating the findings of the NJC analysis.

Other less important spin-coupling mechanisms involving these two respective $\sigma_{\text{C,O}}^*$ acceptor orbitals are *a)* the geminal delocalizations from either $\sigma_{\text{C1,C2}}$ as a function of C1 configuration or $\sigma_{\text{C2,C3}}$ for C3 configuration, and *b)* donation from the core NAO of the carbon whose configuration is in question into that carbon's $\sigma_{\text{C,O}}^*$. Additionally there is a differential interaction between the π -shaped O2 lone pair orbital and either $\sigma_{\text{C1,O1}}^*$ or $\sigma_{\text{C3,O3}}^*$ as a function of C1 or C3 configuration respectively, however this interaction is highly C2-O2 conformation dependent (*vide infra*, §I). Other $\mathcal{J}^{(\text{deloc})}$ mechanisms directly correlated with the overall trend in ${}^2J_{\text{C1,C3}}$ are significantly smaller and not clearly systematic. Analogous to the Lewis coupling mechanisms (*vide supra*) there is a net effect on the $\mathcal{J}^{(\text{deloc})}$ portion of ${}^2J_{\text{C1,C3}}$ from offsetting $\sigma_i \rightarrow \sigma_j^*$ hyperconjugations which inversely correlate with the overall trend in ${}^2J_{\text{C1,C3}}$, the most significant of which are $\sigma_{\text{total}} \rightarrow \sigma_{\text{C1,H1}}^*$ for changes in configuration at C1 and $\sigma_{\text{total}} \rightarrow \sigma_{\text{C3,H3}}^*$ for C3 (Table 7.6), however the magnitude of these effects are smaller than those involving the C-O anti-bonding NBOs, hence the overall coupling behavior observed.

7.5.9. Natural J-coupling Analysis II: Conformational Dependence of ${}^2J_{\text{C1,C3}}$

The C2-O2 conformational dependence of ${}^2J_{\text{C1,C3}}$ resolves into distinct $\mathcal{J}^{(\text{L})}$ and $\mathcal{J}^{(\text{deloc})}$ portions in much the same manner as the configurational dependence. The

symbology $\Delta_{t \rightarrow \bar{g}} {}^2J_{C1,C3}^{(L)}$ will be used to denote the difference between the average $J^{(L)}$ portion of ${}^2J_{C1,C3}$ for the nine C1-O1/C3-O3 conformers with H2-C2-O2-H = 180° (*trans*) and the similar average of the eighteen C1-O1/C3-O3 conformers with H2-C2-O2-H = 60°/-60° (*gauche*). An analogous symbology will be used for $J^{(\text{deloc})}$. The ‘L’ or ‘deloc’ superscripts can be replaced with a σ_i or $\sigma_i \rightarrow \sigma_j^*$ respectively to denote a specific coupling mechanism. A summary of the major C2-O2 conformation dependent $J^{(L)}$ and $J^{(\text{deloc})}$ mechanisms along with a comparison to the trends in the overall DFT data can be found in Table 7.7.

The most significant contributions to $J^{(L)}$ that directly correlate with the overall C2-O2 conformation dependent trend in ${}^2J_{C1,C3}$ come from $\sigma_{C1,C2}$ and $\sigma_{C2,C3}$, which taken together are responsible for 1.0 Hz of the observed dependence regardless of configuration at the coupled carbons. The only pairwise steric exchange energies ($\Delta E_{i,j}$) involving $\sigma_{C1,C2}$ or $\sigma_{C2,C3}$ that vary as a function of ξ conformation exclusively involve the O2 lone pair orbitals. These lone pair orbitals have distinct pseudo-s and pseudo- π shapes (Figure 7.14) and are symbolically represented as O2lp $_{\sigma}$ and O2lp $_{\pi}$. Of the two, O2lp $_{\pi}$ had the most significant variation in the steric interactions with $\sigma_{C1,C2}$ or $\sigma_{C2,C3}$ as a function of ξ conformation.

It should be noted that the conventional tetrahedral depiction of the total lone pair electron density around the oxygen nucleus is related to the actual electronic distribution only by virtue of a shared symmetry about the C-O-H plane. Considered separately, O2lp $_{\sigma}$ has a bilateral plane of symmetry coincident with the C2-O2-H plane, whereas the

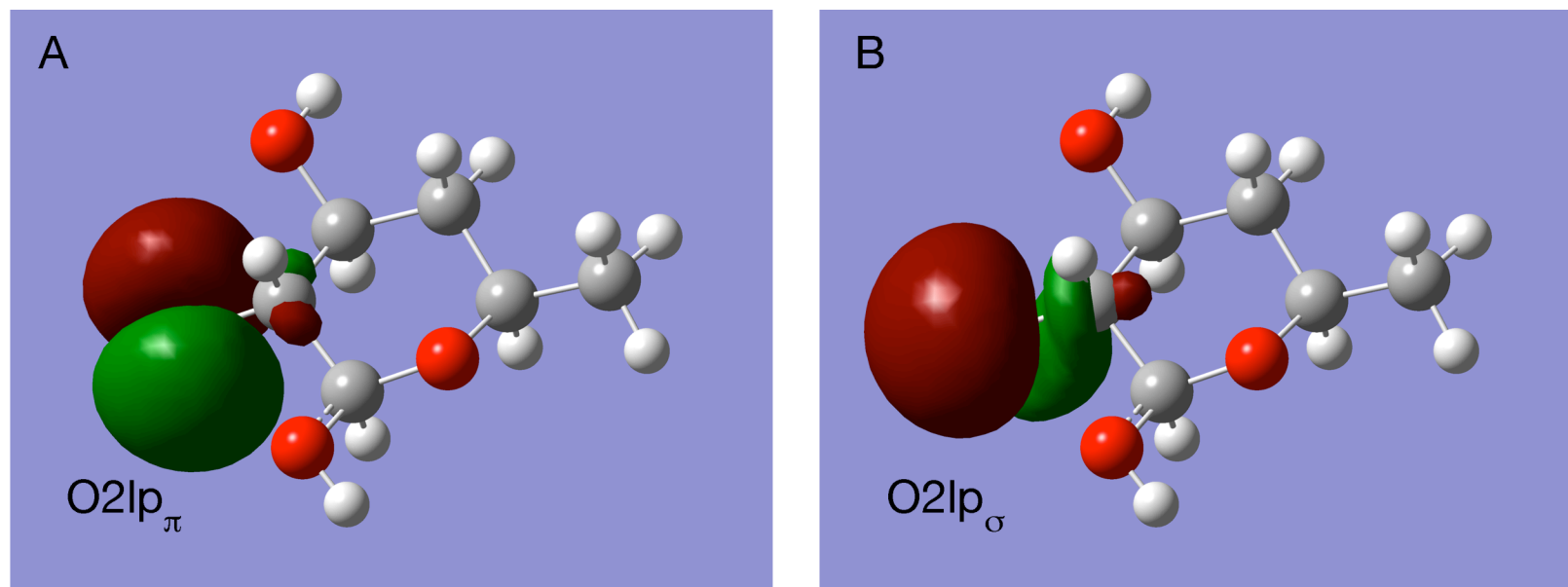


Figure 7.14. O2 Lone Pair Orbitals from NBO Analysis of 2^C . Pseudo π shaped orbital (A) and pseudo σ shaped orbital (B) on O2 are the largest contributors to the ξ conformation dependent changes in the magnitude of ${}^2J_{C1,C3}$ between the two gauche and the trans conformers in 2^C . The dihedrals ϕ , ξ and $\zeta = 180^\circ$ for both structures depicted. The coloration denotes orbital phasing.

TABLE 7.7. COMPARISON OF THE DIFFERENCE BETWEEN AVERAGED H2-C2-O2-H *TRANS* AND *GAUCHE* $J^{(L)}$ AND $J^{(DELOC)}$ WITH DFT TOTAL ${}^2J_{C1,C3}$ AS A FUNCTION OF COUPLED SITE CONFIGURATION^A

NBO	$\Delta_{t \rightarrow \bar{g}} {}^2J_{C1,C3}^{(L) \text{ } b}$		
	2C (β -Glc)	1C (α -Glc)	5C (α -Allo)
$\sigma_{C1,C2}$	-0.5 (0.0) ^c	-0.6 (0.0)	-0.6 (0.0)
$\sigma_{C2,C3}$	-0.5 (0.0)	-0.4 (0.0)	-0.4 (0.0)
Net $J^{(L)}$	-1.0	-1.0	-1.0
Donor NBO	$\Delta_{t \rightarrow \bar{g}} {}^2J_{C1,C3}^{(deloc) \text{ } d}$		
O2 lp_{σ}^e	2.1 (0.1)	1.8 (0.2)	1.7 (0.1)
O2 lp_{π}^e	-3.2 (0.0)	-2.4 (-0.1)	-2.0 (0.0)
Net O2 $lp \text{ } J^{(deloc)}$	-1.1	-0.6	-0.4
NJC $J^{(L)} + J^{(deloc)}$	-2.1	-1.6	-1.4
DFT $\Delta_{t \rightarrow \bar{g}} {}^2J_{C1,C3}^f$	-2.4	-1.4	-0.8

^aAll values reported in Hz. ^bDifference between the average $J^{(L)}$ for the indicated NBO of the nine C1-O1/C3-O3 conformers of H2-C2-O2-H = 180° (*trans*) and the eighteen C1-O1/C3-O3 conformers of H2-C2-O2-H = 60°/-60° (*gauche*). ^cValue in parenthesis is the difference between the average across the nine C1-O1/C3-O3 conformers each of H2-C2-O2-H = -60° and 60°. ^dDifference between the average $J^{(deloc)}$ for the indicated donor NAO of the nine C1-O1/C3-O3 conformers of H2-C2-O2-H = 180° (*trans*) and the eighteen C1-O1/C3-O3 conformers of H2-C2-O2-H = 60°/-60° (*gauche*). ^eThe symbol lp_{σ} denotes the O2 pseudo- σ shaped and lp_{π} the pseudo-p shaped lone pair natural atomic orbitals. ^fValues taken from Table 2.

symmetry of O2 lp_{π} is a bilateral plane running along the C2-O2 bond vector and perpendicular to the C2-O2-H plane. This nuanced description of the geometry of lone

pair density is essential to a detailed understanding of the ξ dependent ${}^2J_{C1,C3}$ Karplus profile.

Figure 7.15 contains the pairwise steric exchange energies between $\sigma_{C1,C2}$ or $\sigma_{C2,C3}$ and $O2lp_\pi$ as averages for the 3 conformers of ξ in **2C**, along with depictions of the overlap between the corresponding pre-orthogonal NBOs. The ξ dependent variation in the pairwise steric exchange energies in **2C** are $\Delta_{t \rightarrow \bar{g}} E_{\sigma_{C1,C2}, O2lp_\pi} = -1.6$ kcal/mol and $\Delta_{t \rightarrow \bar{g}} E_{\sigma_{C2,C3}, O2lp_\pi} = -1.4$ kcal/mol, with comparable values for the analogous energy differences in **1C** and **5C**. From Figure 7.15 it is of note that $E_{\sigma_{C,C}, O2lp_\pi}$ tracks along with the angle between the $O2lp_\pi$ plane of symmetry and the corresponding C-C bond vector (and thus the axis of symmetry for the corresponding $\sigma_{C,C}$ NBO), whereas either ${}^2J_{C1,C3}^{(\sigma_{C1,C2})}$ or ${}^2J_{C1,C3}^{(\sigma_{C2,C3})}$ track with the *sum* of the two $E_{\sigma_{C,C}, O2lp_\pi}$ as a function of ξ . This provides an indication as to the underlying coupling mechanism at work and insight into the shape of the ξ dependent ${}^2J_{C1,C3}$ Karplus profile (*vide infra*, §L).

When $\xi = 180^\circ$, there is the maximal simultaneous steric interaction between $O2lp_\pi$ and both of the $\sigma_{C,C}$ NBOs. When $\xi = \pm 60^\circ$, one of the steric interactions between $O2lp_\pi$ and a $\sigma_{C,C}$ is roughly identical to $\xi = 180^\circ$ while the other is significantly reduced. The NBO paradigm considers pairwise coupling mechanisms between bonding orbitals sterically interacting or between donor \rightarrow acceptor NBOs engaged in hyperconjugation, however the complexities of alternate possible coupling mechanisms have been rigorously treated⁴⁰. The transfer of spin information from C1 to C3 as mediated by $O2lp_\pi$, $\sigma_{C1,C2}$ and $\sigma_{C2,C3}$ is likely well represented by the

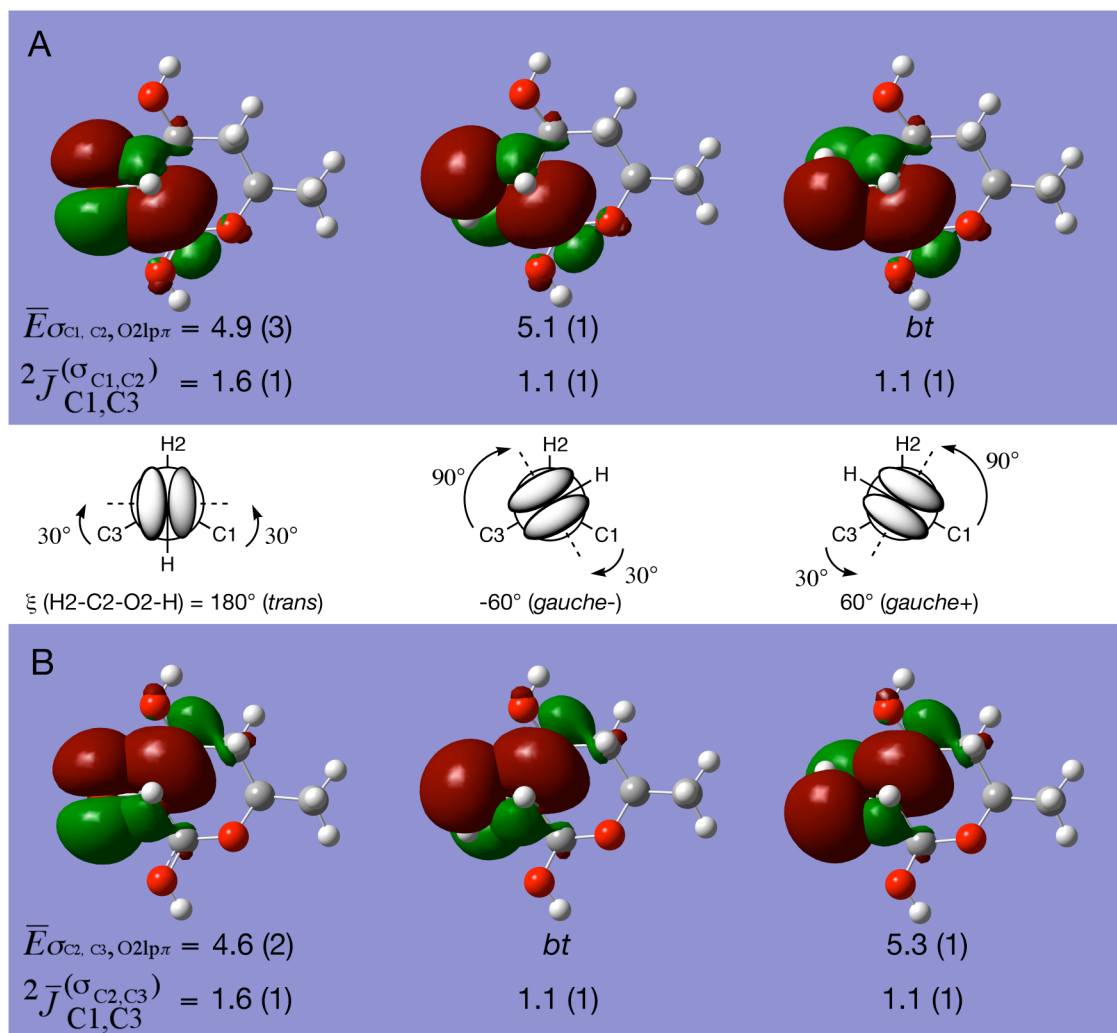
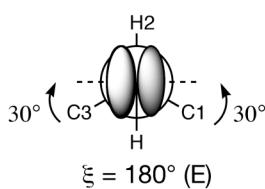
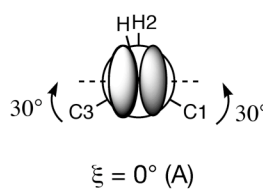
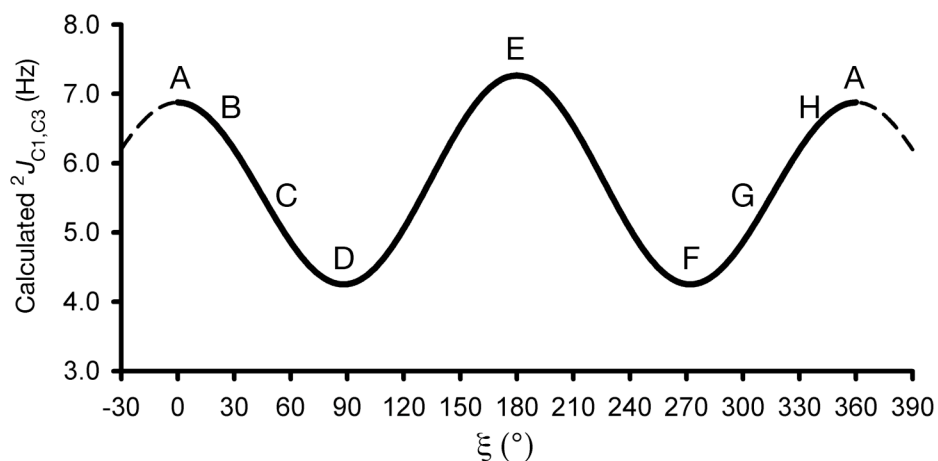


Figure 7.15. The ξ Dependent Variation in the Pairwise Steric Exchange Energy Between $O2lp\pi$ and $\sigma_{C1, C2}$ or $\sigma_{C2, C3}$. The overlap in the $2C$ PNBOs for $O2lp\pi$ and $\sigma_{C1, C2}$ (A) or $O2lp\pi$ and $\sigma_{C2, C3}$ (B) are a reflection of the angle between the bilateral plane of symmetry for $O2lp\pi$ (dotted line in central scheme) and either the $C1-C2$ or $C2-C3$ bond vectors respectively. This in turn is directly related to the pairwise steric exchange energies (E_{ij}) which are listed in kcal/mol. An entry of *bt* indicates that the E_{ij} was below the 0.2 kcal/mol threshold. The corresponding $J^{(L)}$ contributions are listed in Hz for comparison. The bar above the E and the J indicates that the reported value is averaged over the 9 conformers of $C1-O1$ and $C3-O3$. The number in the parenthesis is the standard deviation in the last reported digit. The orbital coloration denotes phasing and is consistent only within an individual PNBO. The ϕ and ξ conformations depicted (definitions in text) all = 180° , whereas ξ conformation is consistent within a column.

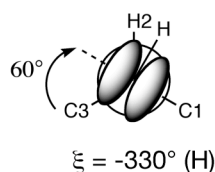
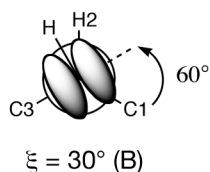
(B)→*k*→*m*→*l*→(A) mechanism outlined by Cremer, *et al.*⁴⁰, where (B) and (A) are coupled nuclei, in this case C1 and C3, and *k*, *m* and *l* are intermediate molecular orbitals, here $\sigma_{C1,C2}$, $O2lp_{\pi}$ and $\sigma_{C2,C3}$ respectively. Thus, while ${}^2J_{C1,C3}^{(\sigma_{C1,C2})}$ and ${}^2J_{C1,C3}^{(\sigma_{C2,C3})}$ represent equal parts of a concerted coupling mechanism, the two $E_{\sigma_{C,C},O2lp_{\pi}}$ are discreet phenomena.

This analysis can be applied to understanding the bi-modality exhibited by the ξ dependent ${}^2J_{C1,C3}$ Karplus profile (Scheme 7.3). The ${}^2J_{C1,C3}$ is the most positive in value when $\xi = 0^\circ$ and 180° due to maximization of the simultaneous steric interactions between $O2lp_{\pi}$ and both of the $\sigma_{C,C}$ NBOs, whereas the coupling is the most negative in value when there is the greatest simultaneous minimization of the same interactions, which happens at $\xi = 90^\circ$ and 270° . The small difference between the magnitude of $\xi = 0^\circ$ and 180° is due in part to a minor variation in the steric interactions between $O2lp_{\pi}$ and the $\sigma_{C,C}$ NBOs as a result of the slight asymmetry of $O2lp_{\pi}$ across its plane of symmetry induced by the presence of the hydroxyl proton.

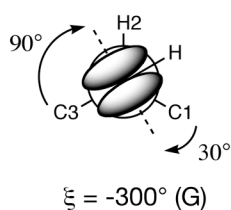
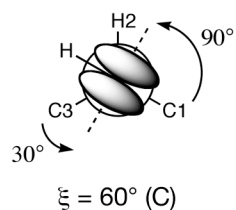
There are no significant $J^{(L)}$ mechanisms that were explicitly identified as inversely correlated with the overall coupling trends in ${}^2J_{C1,C3}$, however it is likely that the ξ dependent variation in the ${}^2J_{C1,C3}^{(\sigma_{C1,C2})}$ and ${}^2J_{C1,C3}^{(\sigma_{C2,C3})}$ terms was slightly attenuated through the interaction of $O2lp_{\sigma}$ with $\sigma_{C1,C2}$ and $\sigma_{C2,C3}$. This is based upon the observation that the additive pairwise steric exchange energy between $O2lp_{\sigma}$ and both of the $\sigma_{C,C}$ NBOs was minimized when $\xi = 180^\circ$ in direct contrast to $O2lp_{\pi}$. As noted above, the variation in the pairwise steric exchange energy between $O2lp_{\sigma}$ and the $\sigma_{C,C}$



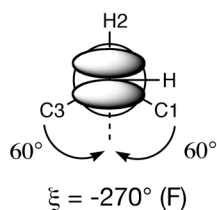
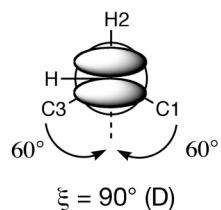
Maximum simultaneous alignment between $O2lp_\pi$ and both $\sigma_{C,C}$ NBOs.



Maximum alignment between $O2lp_\pi$ and one $\sigma_{C,C}$ NBO, intermediate alignment between $O2lp_\pi$ and other $\sigma_{C,C}$ NBO.



Near maximum alignment between $O2lp_\pi$ and one $\sigma_{C,C}$ NBO, minimum alignment between $O2lp_\pi$ and other $\sigma_{C,C}$ NBO.



Minimum simultaneous alignment between $O2lp_\pi$ and both $\sigma_{C,C}$ NBOs.

Scheme 7.3. Explanation of $^2J_{C1,C3}$ Karplus Profile Shape Based on $O2lp_\pi$ Disposition

NBOs as a function of ξ is much smaller than that of $O2lp_\pi$; there is a 1.7 kcal/mol greater $E_{\sigma_{C,C},O2lp_\sigma}$ when the O2 hydroxyl proton and thereby the belly of the central lobe of $O2lp_\sigma$ are anti to the respective C-C bond and closely interacting with the back lobe of the $\sigma_{C,C}$ compared to either gauche orientation.

In terms of other inverse $J^{(L)}$ correlations with ξ dependent ${}^2J_{C1,C3}$ behaviour, there are the minor observations that $\sigma_{C1,H1}$ gives rise to $\Delta_{t \rightarrow \bar{g}} {}^2J_{C1,C3}^{(\sigma_{C1,H1})} = -0.2$ Hz in **1C**, and that $\sigma_{C1,H1}$ and $\sigma_{C3,H3}$ are each responsible for a -0.3 Hz change in $\Delta_{t \rightarrow \bar{g}} {}^2J_{C1,C3}^{(L)}$ for **5C**. This effect, presumably arising from the change in the disposition of these bonding orbitals with respect to both the coupling path and the O2 hydroxyl, contributes to the differential sensitivity of ${}^2J_{C1,C3}$ to the C2-O2 conformation between the three structural groups in Table 7.2.

The only donor orbital contributions to $\Delta_{t \rightarrow \bar{g}} {}^2J_{C1,C3}^{(deloc)}$ that demonstrate significant ξ conformation dependence arise from the O2 lone pair orbitals. From Table 7.7 it can be seen that $\Delta_{t \rightarrow \bar{g}} {}^2J_{C1,C3}^{(O2lp_\pi \rightarrow \sigma_{j^*})}$ is directly correlated while $\Delta_{t \rightarrow \bar{g}} {}^2J_{C1,C3}^{(O2lp_\sigma \rightarrow \sigma_{j^*})}$ is inversely correlated with the overall C2-O2 conformational dependence of ${}^2J_{C1,C3}$. The net effect of C2-O2 rotation on ${}^2J_{C1,C3}^{(deloc)}$ is essentially the combined and counterbalancing contributions from the $O2lp_\sigma \rightarrow \sigma_{j^*}$ and $O2lp_\pi \rightarrow \sigma_{j^*}$ mechanisms. In all cases $\Delta_{t \rightarrow \bar{g}} {}^2J_{C1,C3}^{(O2lp_\pi \rightarrow \sigma_{j^*})}$ is of a greater absolute magnitude than $\Delta_{t \rightarrow \bar{g}} {}^2J_{C1,C3}^{(O2lp_\sigma \rightarrow \sigma_{j^*})}$, however the difference between these two sets of mechanisms

decreases in going from Group 1 to Group 3 structures. This is the ${}^2J_{C1,C3}^{(\text{deloc})}$ contribution to the differential sensitivity of ${}^2J_{C1,C3}$ to the C2-O2 conformation as a function of coupled carbon configuration alluded to previously (*vide supra*, §B).

There are two primary and two secondary acceptor orbitals that are responsible for the transfer of spin density from O2lp_σ and O2lp_π (Table 7.8) Independent of coupled carbon configuration, σ_{C1,C2}^{*} and σ_{C2,C3}^{*} are the primary acceptor orbitals and when taken together are responsible for ~90% of the spin density transfer as mediated by O2lp_σ and O2lp_π (Figure 7.16).

TABLE 7.8. ξ CONFORMATIONAL DEPENDENCE OF O2 LP→σ_j^{*} AS A
FUNCTION OF COUPLED SITE CONFIGURATION^A

Acceptor NBO	$\Delta_{t \rightarrow \bar{g}} {}^2J_{C1,C3}^{(O2\text{ lp}\sigma \rightarrow \sigma_j^*)b}$			$\Delta_{t \rightarrow \bar{g}} {}^2J_{C1,C3}^{(O2\text{ lp}\pi \rightarrow \sigma_j^*)b}$		
	2C(β-Glc)	1C(α-Glc)	5C(α-Allo)	2C(β-Glc)	1C(α-Glc)	5C(α-Allo)
σ _{C1,C2} [*]	0.9	0.9	0.8	-1.3	-1.3	-1.0
σ _{C2,C3} [*]	1.0	0.8	0.9	-1.3	-0.9	-1.0
σ _{C1,O1} [*]	0.0	-0.1	-0.1	-0.2	0.3	0.3
σ _{C3,O3} [*]	-0.1	0.0	-0.2	-0.2	-0.1	0.3
σ _{C1,H1} [*]	0.1	0.2	-0.2	0.1	-0.2	-0.2
σ _{C3,H3} [*]	0.0	-0.1	0.0	0.0	0.0	0.0

^aAll values reported in Hz. ^bDifference between the average O2 lp→σ_j^{*} for the indicated acceptor NBO (σ_j^{*}) of the nine C1-O1/C3-O3 conformers of H2-C2-O2-H = 180° (*trans*) and the eighteen C1-O1/C3-O3 conformers of H2-C2-O2-H = 60°/-60° (*gauche*). The symbol lp_σ denotes the O2 pseudo-σ shaped and lp_π the pseudo-p shaped lone pair natural atomic orbitals.

This result is consistent with previous finding that lone pair interactions with carbon-carbon anti-bonding orbitals can significantly affect spin-spin coupling sign and

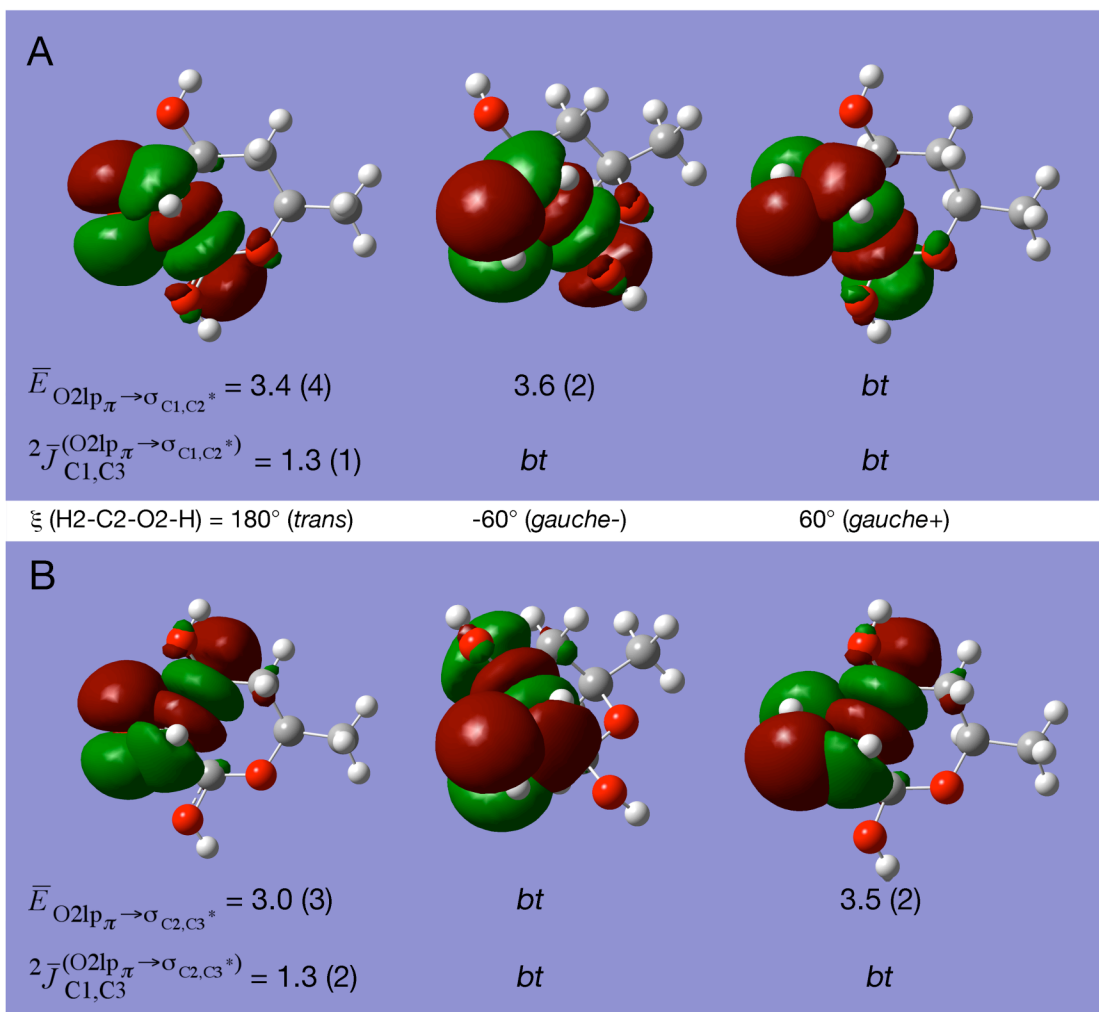


Figure 7.16. The ξ Dependant Variation in the Second Order Perturbation Energy Between O2lp π and $\sigma_{\text{C1,C2}^*}$ or $\sigma_{\text{C2,C3}^*}$ in 2C. The overlap in the 2C PNBOs for O2lp π and $\sigma_{\text{C1,C2}^*}$ (A) or O2lp π and $\sigma_{\text{C2,C3}^*}$ (B) vary as a function of ξ . This in turn is directly related to the 2nd-order perturbation energy ($E_{i,j}$) which are listed in kcal/mol. The corresponding $J^{(\text{deloc})}$ is listed in Hz for comparison. An entry of *bt* indicates that the $E_{i,j}$ or the $J^{(\text{deloc})}$ were below the respective 0.2 kcal/mol or 0.1 Hz thresholds. The bar above the E and the J indicates that the reported value is averaged over the 9 conformers of C1-O1 and C3-O3. The number in the parenthesis is the standard deviation in the last digit. The antibonding ($\sigma_{\text{C,C}^*}$) orbitals have four lobes with a node at each of the bonded carbon nuclei, as well as a third node between them. The ϕ and ζ conformations depicted (definitions in text) are all = 180°, whereas ξ conformation is consistent within a column.

magnitude^{4b,64-65}. From inspection of the $O2lp_\pi$ and $\sigma_{C,C}^*$ PNBOs in Figure 7.16 it can be seen that when $H2-C2-O2-H = 180^\circ$, there is a significant well aligned overlap between $O2lp_\pi$ and both the back lobe and the first central lobe of either $\sigma_{C,C}^*$. This is reflected in the $E_{O2lp_\pi \rightarrow \sigma_{C,C}^*}$ and the corresponding ${}^2J_{C1,C3}^{(O2lp_\pi \rightarrow \sigma_{C,C}^*)}$ values for this geometry. In conformations where $\xi = \pm 60^\circ$, there is an optimal alignment between $O2lp_\pi$ and one of the $\sigma_{C,C}^*$ NBOs similar to the $\xi = 180^\circ$ situation, but a poor alignment between $O2lp_\pi$ and the other $\sigma_{C,C}^*$ NBO. The aligned orbital pair show a 2nd-order perturbation energy similar in magnitude to the $\xi = 180^\circ$ orientation, whereas the $E_{O2lp_\pi \rightarrow \sigma_{C,C}^*}$ for misaligned pair of orbitals is <0.2 kcal/mol. In contrast, when $\xi = \pm 60^\circ$ the ${}^2J_{C1,C3}^{(O2lp_\pi \rightarrow \sigma_{C,C}^*)}$ contribution for either $\sigma_{C,C}^*$ is <0.1 Hz, indicating that the ${}^2J_{C1,C3}^{(O2lp_\pi \rightarrow \sigma_{C,C}^*)}$ coupling mechanism is likely a concerted $(B) \rightarrow k \rightarrow m \rightarrow l \rightarrow (A)$ type analogous to the ${}^2J_{C1,C3}^{(\sigma_{C,C})}$ Lewis coupling mechanism except that in this case k and l are C-C anti-bonding rather than bonding orbitals.

The correlation either with or against ξ dependent ${}^2J_{C1,C3}$ coupling behavior from the respective ${}^2J_{C1,C3}^{(O2lp_\pi \rightarrow \sigma_{C,C}^*)}$ or ${}^2J_{C1,C3}^{(O2lp_\sigma \rightarrow \sigma_{C,C}^*)}$ coupling mechanisms is due to a pattern of $O2lp_\sigma$ and $O2lp_\pi$ orbital interaction with the two $\sigma_{C,C}^*$ in a manner analogous to their mediation of the $J^{(L)}$ contributions from the two bonding $\sigma_{C,C}$ NBOs. This is reflected in the second order perturbation energies for the hyperconjugations between $O2lp_\sigma$ or $O2lp_\pi$ with either of the two $\sigma_{C,C}^*$. Analogous to the relative magnitude of the

effects that $O2lp_{\sigma}$ versus $O2lp_{\pi}$ had on the two $E_{\sigma_{C,C}}$ and by inference on the $J^{(L)}$ contributions from the two bonding $\sigma_{C,C}$ NBOs, the observed $J^{(deloc)}$ contribution to the total coupling is a counterbalancing combination of effects from the two O2 lone pairs, where the effect due to $O2lp_{\pi}$ is the larger of the two due to its increased potential for interaction with the two $\sigma_{C,C}^*$.

The identity of the pertinent acceptor orbitals of secondary importance is contingent upon configuration at the coupled carbon in question, and as such are either the coupled carbon $\sigma_{C,O}^*$ or $\sigma_{C,H}^*$ orbitals. The difference in the degree of orbital overlap and thereby the efficiency of spin density transfer via the $O2lp_{\pi} \rightarrow \sigma_{C,O}^*$ versus $O2lp_{\pi} \rightarrow \sigma_{C,H}^*$ mechanisms is at the root of the small C1/C3 configuration dependent attenuation of the response in ${}^2J_{C1,C3}^{(deloc)}$ to ξ conformation.

The manner in which C-O rotamer conformation dependent variation in ${}^2J_{C1,C3}$ is mediated via a small ensemble of spin transfer mechanism is well visualized through a normalized plot of DFT calculated ${}^2J_{C1,C3}$ versus the sum of ${}^2J_{C1,C3}^{(L)}$ for the $\sigma_{C1,C2}$ and $\sigma_{C2,C3}$ terms and ${}^2J_{C1,C3}^{(deloc)}$ for the $\sigma_{C1,C2}^*$ and $\sigma_{C2,C3}^*$ terms, with a C1/C3 configuration dependent inclusion of $\sigma_{C1,O1}^*$ and/or $\sigma_{C3,O3}^*$ (Figure 7.17). For **2**^C the normalized sum of all six of the above mentioned terms describes ~99% of the variation in ${}^2J_{C1,C3}$, with a remarkably good linear fit that has a slope of 1.03 and an intercept of -0.01. The analogous fit for **1**^C with its axial O1 hydroxyl, is improved by the removal of the $\sigma_{C1,O1}^*$ term ($r^2 = 0.96$) and roughly insensitive to the inclusion of the $\sigma_{C3,O3}^*$ term

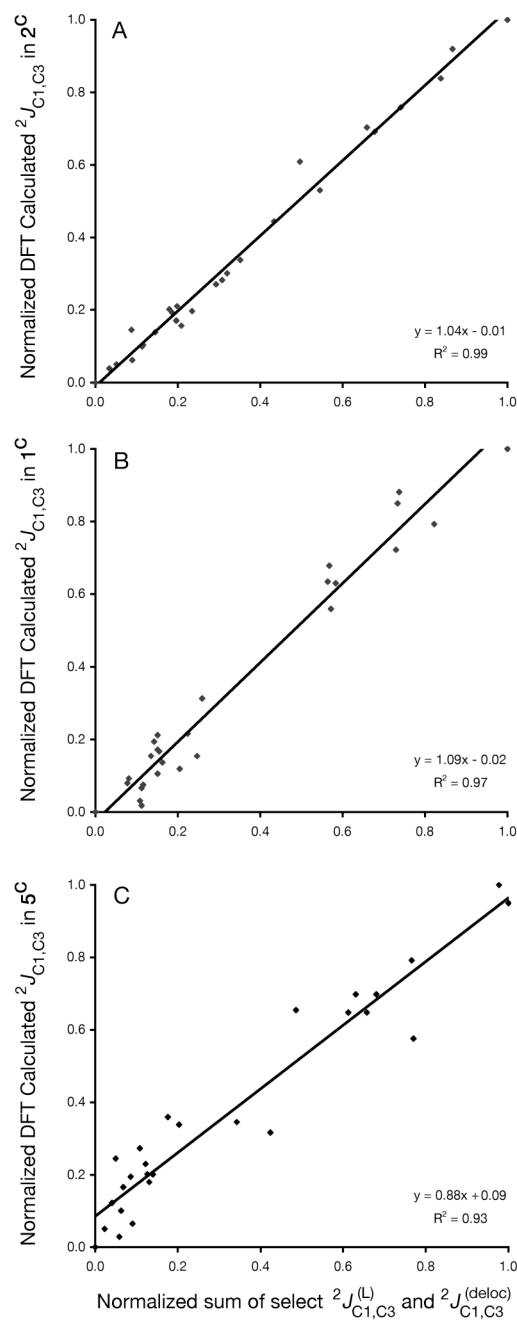


Figure 7.17. The Normalized Relationship Between ${}^2J_{C1,C3}$ and the Sum of Select ${}^2J_{C1,C3}^{(L)}$ and ${}^2J_{C1,C3}^{(deloc)}$. The select ${}^2J_{C1,C3}^{(L)}$ and ${}^2J_{C1,C3}^{(deloc)}$ terms summed and normalized for the 27 conformations of **2C** (A) are $\sigma_{C1,C2}$, $\sigma_{C2,C3}$ and $\sigma_{C1,C2}^*$, $\sigma_{C2,C3}^*$, $\sigma_{C1,O1}^*$, $\sigma_{C3,O3}^*$ respectively with the removal of $\sigma_{C1,O1}^*$ for **1C** (B) and the removal of both $\sigma_{C,O}^*$ terms for **5C** (C)

($r^2 = 0.97$) whereas the optimal fit for ^5C that accounts for ~92% of the variation in $^2J_{\text{C1,C3}}$ is obtained through the sum of $^2J_{\text{C1,C3}}^{(\text{L})}$ for the $\sigma_{\text{C1,C2}}$ and $\sigma_{\text{C2,C3}}$ terms and $^2J_{\text{C1,C3}}^{(\text{deloc})}$ for the $\sigma_{\text{C1,C2}}^*$ and $\sigma_{\text{C2,C3}}^*$ terms alone.

7.5.10. Coupling Trend Validation in Conformationally Determined Lacotside System

Even given this study's coarse sampling of conformations upon which $^2J_{\text{CC}}$ coupling structure relationships are based, a qualitative validation of this parameter's power to probe C-O rotamer conformation can be found in the independently conformationally determined system of α -D-lactoside (L^{E} , Scheme 7.1). The $^2J_{\text{C2,C4}}$ value in 2^{E} of 2.8 Hz is identical to the room temperature value measured in the glucose residue of L^{E} in 10% $^2\text{H}_2\text{O}$, however the same coupling is found to be 2.3 Hz when measured in $\text{H}_2\text{O}/\text{acetone-}d_6$ at -5 °C, and 1.9 Hz in $\text{DMSO-}d_6$ at 25 °C, conditions both shown to favor an inter-residue $\text{O3}'\text{-H}\cdots\text{O5}$ hydrogen bond. Secondary isotope effects on the chemical shift of C3' in L^{E} from the deuteration of the O3' hydroxyl in $\text{DMSO-}d_6$ at 25 °C have been shown to be indicative of the persistent involvement of O3' in an intramolecular H-bond^{90a}. Such a hydrogen bond requires a $\text{H3}'\text{-C3}'\text{-O3}'\text{-H}$ *gauche*+ conformation with rotamer populations estimated at 85% in $\text{H}_2\text{O}/\text{acetone-}d_6$ (-5 °C) and 90% in $\text{DMSO-}d_6$ (25 °C).^{32a}

When the three experimentally measured trans-glycoside spin-spin couplings sensitive to psi are analyzed in the light of DFT parameterizations a single state average value of C1-O-C4'-C3' between 0° and -15° results for each of the three sample

conditions.³³ From this study it is also known that C1-O-C4'-C3' dependent variation in $^2J_{C2',C4'}$ is unimodal with a minimum at 0° and a range of < 0.5 Hz between -60° and 60°. A qualitative back-calculation of $^2J_{C2',C4'}$ in **L^E** is possible utilizing *i.*) the Series 3 DFT data in which C3-C4-O4-H = -60° based upon the preceding argument, *ii.*) the published estimates for H3'-C3'-O3'-H *gauche*+ populations in the three experimental conditions, and assuming *iii.*) a conformationally averaged C2-O2 rotamer.

Despite the use of the coarsely sampled relationships between C-O rotamer conformation and $^2J_{C2,C4}$ in this study, this technique reproduces the trend in the experimental data remarkably well by predicting couplings of 2.2, 1.7, and 1.6 Hz for 10% ²H₂O (25 °C), H₂O/acetone-*d*₆ (-20 °C) and DMSO-*d*₆ (25 °C) respectively. This makes intuitive sense by noting that the presence of the O3'-H...O5 hydrogen bond pulls the H3'-C3'-O3'-H rotamer out of the anti conformation typical of maximal coupling.

7.5.11. O2 Ionization and $^2J_{C1,C3}$

Thinking about the system simplistically, we predicted that if $^2J_{C1,C3}$ became more positive in value when the O2 hydroxy proton is *gauche* to both carbon-carbon bonds of the coupling pathway, then the *J* should also become more positive in value when the comparatively acidic O2 hydroxyl proton⁸⁰ is titrated off at high pH. That is to say, $^2J_{C1,C3}$ should be more sensitive to ionization at O2 than O1 or O3 due to the heightened ability of O2 to interact with orbitals relevant to the coupling path. The magnitude of $^2J_{C1,C3}$ in [1-¹³C]-Methyl-β-D-glucoside became more positive in value by approximately 0.3 Hz when measured between the pH range of 9.8 and 12.4. Considering

that the aglycone is stable to hydrolysis throughout this pH range, and that the O2 hydroxyl is known to be more acidic than O3⁸⁰, this shift in J is likely due to ionization of the O2 hydroxyl. A pH dependent shift in C-O rotamer conformation is difficult to absolutely rule out without more extensive labeling, although $^1J_{C1,C2}$ varied by only about 0.1 Hz, indicating that the C1-O1 rotamer did not significantly reorient as a function of pH.

This trend was reproduced by *in solvo* DFT calculations for **18C** (Series 7) in which the methyl aglycone was held fixed anti to C2, the O2 hydroxyl was ionized and the O3 hydroxyl sampled its three rotamers. The average $^2J_{C1,C3}$ from the three geometries sampled was 9.8 Hz. The calculation indicates that the O2 hydroxyl proton was likely incompletely titrated off, although our prediction was qualitatively verified.

7.5.12. The Conformational Analysis of ϕ , ξ and ζ in Methyl- β -D-glucopyranose

The Parameterization of spin-spin couplings sensitive to ϕ , ξ and ζ in 9C — In light of the significant conformational dependence of $^2J_{C1,C3}$ on the three hydroxyl substituents, an attempt was made to quantitatively parameterize the coupling as a function of C-O rotamer in **9C**. A $12 \times 12 \times 12$ grid of B3LYP/631G* optimized structures for **9C** in which ϕ , ξ and ζ conformation was varied in 30° increments (C3-C4-O4-H = 180° for all structures) gave rise to a total of 1,728 converged structures. The total ensemble of spin-spin couplings for each of the 1,728 structures was calculated *in vacuo* with the extended [5s2p1d|3s1p]¹⁸ basis set.

A plot of the calculated $^2J_{C1,C3}$ as a function of each of the three varied dihedrals is found in Figure 7.18. A distinctive characteristic of these plots is the strong bi-modal dependence of the coupling on ξ conformation and the much weaker secondary dependence upon the terminal C-O rotamers. This pattern of coupling behavior in which *i*) the primary bi-modal C-O conformational dependence of the geminal coupling is dictated by the hydroxyl appended to a carbon intermediate to the coupling path and *ii*) the secondary uni-modal dependence arises from the conformation of a hydroxyl appended to a coupled carbon is analogous to the pattern of conformational dependence reported previously in $^2J_{CCH}$ ³⁸.

The parameterization of the coupling as a function of C-O rotamer was conducted iteratively using a large number of trigonometric terms in the initial equation. Initial coefficients were determined individually and subsequently optimized in pairs and triads by a least squares Monte Carlo fitting procedure. Following several rounds of stochastic optimization, a deterministic minimization algorithm was applied to the optimization of all parameters. This method minimized a mean deviation, χ_R , of the type

$$\chi_R = \sum_i [R_\theta(d_{\theta_i}) + R_J(d_{J_i})] \quad (7.1)$$

in which the distances d_{θ_i} and d_{J_i} are the difference between the DFT data and the value of the function at that coordinate. The deviation function $R(d)$ is the mean square deviation between the DFT data points and the function. The parameterization process explored a wide variety of trigonometric terms, however only terms whose coefficients

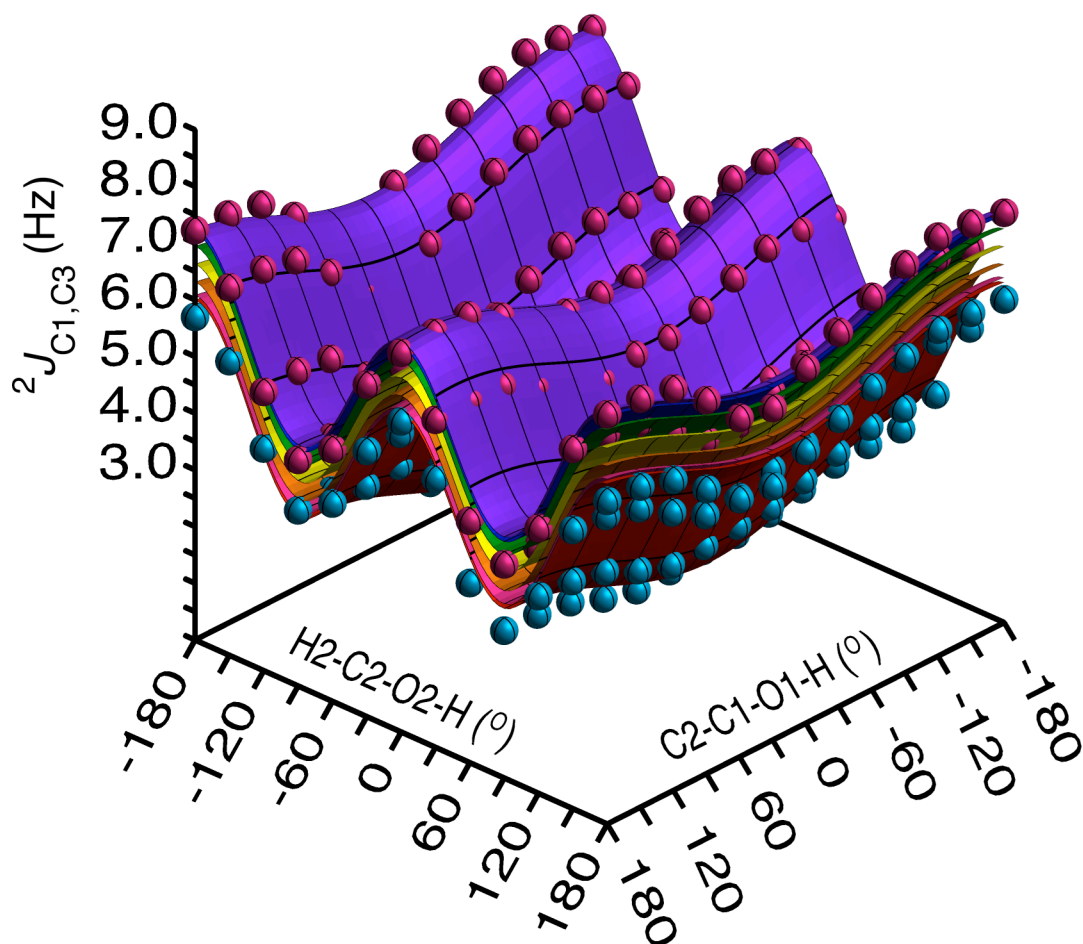


Figure 7.18. Calculated ${}^2J_{C1,C3}$ as a Function of ϕ , ξ and ζ with Parameterization Overlays. The ξ (H2-C2-O2-H) and ϕ (C2-C1-O1-H) dimensions are explicitly indicated while the ζ (C2-C3-O3-H) dimension is implicitly indicated by the vertical stacking of the hypersurfaces. Magenta and teal spheres indicate ${}^2J_{C1,C3}$ calculated by DFT for $\zeta = \pm 180^\circ$ and 0° respectively. Spheres have a 0.5 Hz diameter. DFT data for intermediate ζ values omitted for clarity. Solid surfaces represent best fit to the DFT data for $\zeta = \pm 180^\circ$ (indigo), $\pm 150^\circ$ (blue), $\pm 120^\circ$ (green), $\pm 90^\circ$ (yellow), $\pm 60^\circ$ (orange), $\pm 30^\circ$ (fuchsia) and 0° (red).

were ≥ 0.1 were included in the final form of the equation for the sake of simplicity and due to the minimal contribution to the overall χ^2 between the DFT coupling values and the fit from these minor terms. When a term was identified for elimination, its coefficient was set to zero and the other parameter coefficients were re-optimized by way of a round

of Monte Carlo followed by the deterministic minimization algorithm. The total equation used for parameterization that includes all trigonometric terms along with the optimized coefficient values for ${}^2J_{C1,C3}$ along with the parameterizations for a number of other couplings sensitive to one or more of the C1 through C3 hydroxyl rotamers can be found in Table 7.9. The coefficient values for ${}^2J_{C1,C3}$ gave rise to a simple $\chi^2 = 106.2$ ($N = 1,728$), and $\text{rms} = 0.25$ Hz. There was a 5.6 Hz range with a 7.6 Hz maximum and a 1.9 Hz minimum coupling in the ensemble of DFT data, whereas the parameterization has a 7.3 Hz maximum and a 2.0 Hz minimum coupling to yield a 5.2 Hz range. This indicates that there is a small ensemble of possible couplings that are not reproduced by the parameterization. As expected the full dynamic range of the coupling was not sampled during the qualitative Series 1 set of calculations in which only the staggered rotamers about ϕ , ξ and ζ were explored. The mean value for ${}^2J_{C1,C3}$ averaged across all 1,728 conformations was 4.6 Hz (theoretically represented by the constant term from the leading coefficient A in Table 7.9), a value identical to the experimentally measured coupling in water. Figure 7.18 is a plot of the ξ versus ϕ hypersurface of parameterized ${}^2J_{C1,C3}$ as a series of 30° increment slices through the ζ dimension. The superimposed 0.5 Hz diameter spheres represent the DFT data for $\zeta = 180^\circ$ (magenta) and 0° (teal). The surfaces for a pair of ζ (or ϕ) values symmetric about 180° are superimposable due to the intrinsic symmetry of the parameterization in this dimension. It is notable that dissymmetries particularly in the ϕ and ζ dimensions of the DFT coupling profile give rise to the majority of the rms deviation of the theoretical parameterization from the DFT data. This is particularly true at the extrema, a fact reflected in the less negative kurtosis

TABLE 7.9. GENERALIZED SPIN-SPIN COUPLING EQUATION^A AND COEFFICIENTS^B FOR THE PARAMETERIZATION OF COUPLINGS SENSITIVE TO C1 THROUGH C3 HYDROXYL ORIENTATION.

$${}^nJ = A + B \cos(\xi) + C \cos(2\xi) + D \cos(\phi) + E \cos(2\phi) + F \cos(\xi)\sin(\phi) + G \cos(\xi)\sin(\zeta) + H \cos(\zeta) + I \cos(2\zeta) + J \sin(\zeta) + K \sin(2\zeta) + L \sin(\phi) + M \sin(2\phi) + N \sin(\xi) + O \sin(2\xi) + P \cos(3\xi) + Q \cos(3\phi) + R \sin^2(\xi) + S \sin^2(\phi)$$

Coefficients	${}^1J_{C1C2}$	${}^1J_{C2C3}$	${}^1J_{C1H1}$	${}^1J_{C2H2}$	${}^2J_{C1C3}$	${}^2J_{C1H2}$	${}^2J_{C2H1}$	${}^3J_{H1H2}$	${}^2J_{C2OH2}$	${}^2J_{C3OH3}$	${}^3J_{C1OH2}$	${}^3J_{C2OH3}$	${}^3J_{C3OH2}$	${}^3J_{H2OH2}$	${}^3J_{H3OH3}$
A	46.037	40.327	161.040	147.926	4.554	-5.376	2.429	7.392	-2.310	-2.544	3.988	3.267	3.249	5.830	5.730
B	1.714	1.271	3.777	-4.418	-0.195	0.197	-0.214		-0.056	0.015	1.299	0.035	0.821	-2.761	-0.036
C	0.120	0.195	-1.082	-0.918	1.409	-0.427	0.108		-0.120	-0.088	-2.555	0.080	-1.895	6.779	0.005
D	-3.016	-0.048	2.553	-1.963	-0.474	0.376	0.625		0.154	0.010	-0.206	0.023	-0.082	0.035	0.001
E	-0.718	0.519	2.537	-0.229	-0.079	0.069	-0.532		0.019	0.001	-0.121	0.059	-0.019	0.045	0.008
F	0.308		-0.486		0.165	0.172	0.034	0.269	-0.007					0.006	
G		-0.200								0.011	0.045	-0.017	0.006		-0.120
H	0.061	-3.078	0.672	-2.254	-0.633	-0.189	0.302	-0.169	0.147	0.114	-0.121	-1.217	-0.166	0.047	1.207
I	0.439	-0.231	0.532	-0.610	-0.081	-0.051	0.043	-0.145	0.020	-0.002	-0.008	3.899	-0.064	0.036	-3.592
J			0.620	-1.985											2.392
K			-0.104	-1.678											5.507
L	0.148	0.170	-5.185	2.201			-0.176								
M	-0.907	-0.006	0.994				1.387	-0.209							
N	2.896	-2.337	1.491		-0.043	-0.409	-0.182	-0.173			2.034		-1.289		
O	-0.393	0.166	-1.522	1.843	0.005	-0.587					4.184		-3.406		
P			-0.056	0.528											
Q			0.485												
R			-0.195					-0.140							
S			0.347	0.081				-0.097							
χ^2	1110.0	614.8	1131.3	309.2	106.2	276.6	135.4	298.8	182.5	108.0	237.0	184.8	127.4	270.8	117.3

^a Equations parameterized as a function of the dihedrals ϕ = C2-C1-O1-H, ξ = H2-C2-O2-H and ζ = C2-C3-O3-H. ^b Coefficients determined by a two step Monte Carlo and robust deterministic optimization protocol.

of the DFT data set (-0.7) when compared to that of a similar ensemble of back calculated values (-0.9).

Development of a Generalized Conformational Model — The objective of a conformational analysis based on a NMR spectral property is to impose geometric constraints upon a structural parameter, typically a dihedral angle in the case of indirect spin-spin coupling, by interpreting the experimental measurement through some form of structure-coupling relationship, in this case theoretically derived. Two methods for the determination of the continuous population distribution about a rotational axis from indirect spin-spin coupling constants predominate in the literature^{2c}.

The first, proposed by Lin^{88a-c} and extended by Hägele^{88d} utilized a Boltzmann weighting of the standard Karplus equation to analyze $^3J_{\text{HH}}$ in a variety of asymmetrically [1,2] and [1,1,2] di- and tri-substituted ethanes. This Boltzmann distribution assumes the modified Gwin-Pitzer rotational energy function⁸⁹ describing both the energy difference between rotamers as well as the barriers separating rotamers. The major drawbacks of this method arise from the general inapplicability of assumptions intrinsic to the potential function, as well as the difficulty in determining the requisite individual rotameric energy barriers and the maximal dipole interaction energies specific to the system.

An alternative approach to rotamer analysis is the continuous probability distribution (CUPID) method and was put forth by Džakula in 1992⁹⁰. The probability distribution $\varphi(\theta)$ is a periodic function of the dihedral θ , and as such can be expanded as the Fourier series:

$$\varphi(\theta) = \varphi_0 + \sum_{n=1}^N \varphi_n \cos(n\theta) + \sum_{n=1}^N \sigma_n \sin(n\theta) \quad (7.2)$$

where the $2N$ number of φ_n and σ_n coefficients can be calculated from the Karplus equation dependent analysis of experimental couplings ($\varphi_0 = 1/2\pi$). An advantage of this approach is that *a priori* knowledge regarding the various rotamers is not necessary, in contrast to the Lin method. Another advantage is that given more than $2N$ experimental couplings of n^{th} order profiles, the Fourier coefficients can be obtained through basic linear regression techniques. This is a powerful method if the number of experimental probes is high; as N approaches infinity, the Fourier expansion approaches the exact solution. The obvious weakness is that there are not an infinite number of spin-spin couplings available, however reasonably precise results can be obtained with a relatively low order partial Fourier series. A more serious drawback of this approach is that since typical Karplus-like equations are second order at most, only the first- and second-order Fourier terms can be determined.

An alternative to the Fourier expansion method is to describe the probability distribution as a set of normal population distributions described by a central value (μ) and a standard deviation peak width (σ), both quantities measured in degrees for each of the dimensions in the conformational space.⁴² Using a Gaussian distribution, the normal population density φ as a function of the 360° conformational range θ is described by Equation 7.3,

$$\varphi_{\mu,\sigma^2}(\theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\theta-\mu)^2}{2\sigma^2}\right) \quad (7.3)$$

in which the standard deviation $\sigma > 0$. The relationship between σ and the population width at half height, Γ , is given by Equation 7.4.

$$\Gamma = 2\sqrt{2\ln 2}\sigma \approx 2.355\sigma \quad (7.4)$$

This conformational profile could take the form of a single state model, typified by a distribution with a single maximum, or a multi-state model that may have a number of maxima. These extremes form a continuum with intermediate profiles typified by broad flat regions of distributed population.

Given the theoretical parameterizations for a set of known experimental couplings within a certain conformational space, the logical first step is to consider the simple limiting case in which the peak width is essentially zero and the central value of the populational profile becomes the conformation of one or more static structures depending upon the number of conformational states being modeled. In other words, when $\sigma = 0$ there is no librational averaging about σ . For the sake of simplicity I will consider single state models at first, although a multi-state model would simply involve a population based weighted averaging between states.

Since the process of conformational analysis involves translating a measurement in Hz into a set of geometric conformations, an isosurface plot of ${}^2J_{C1,C3}$ in Hz within the (in this instance 3-dimensional) conformational space derived from the

parameterization is particularly useful (Figure 7.19). For example, in Figure 7.19, given an experimental coupling of ~ 4.6 Hz and assuming a single state static structure, the surface described by the grey spheres represents all conformational combinations of ϕ , ξ and ζ (within a 5° resolution for ϕ and ξ and a 30° resolution for ζ) that have a static structure with a coupling of 4.6 Hz. This ensemble is symbolized by $\{\mu_\phi, \mu_\xi, \mu_\zeta\}$. To further refine the analysis, another experimental coupling would need to be analyzed in the same conformational space, thereby limiting the ensemble of allowable conformations by the intersection of the hypersurfaces from the two couplings.

The dimensionality, n , of a conformational space modeled by normal population density functions is dictated by five factors: *a*) the number of dihedrals considered in the analysis, *b*) the number of conformational states about each of the dihedrals considered, *c*) the unique values μ and σ describing the shape of each distinct conformational state, *d*) the relative populations of the various states about a particular dihedral, and *e*) multiplicities due to the correlation of conformational components between conformationally interdependent dihedrals. In general, the dimensionality of the surface describing the ensemble of conformations which simultaneously satisfy x number of unique experimental couplings within n dimensional conformational space is between n and $n_{+1}-x$ or 0, whichever is greater (zero dimensionality implies a single unique solution). The number of unique couplings sensitive to the specific conformational space that are necessary to find a unique solution ranges from 1 to n dependent upon the experimental coupling and the topology of the accompanying coupling profile. A single experimental coupling at the extreme of the theoretical coupling profile can profoundly restrict the potential solutions in the conformational space, however the occurrence of

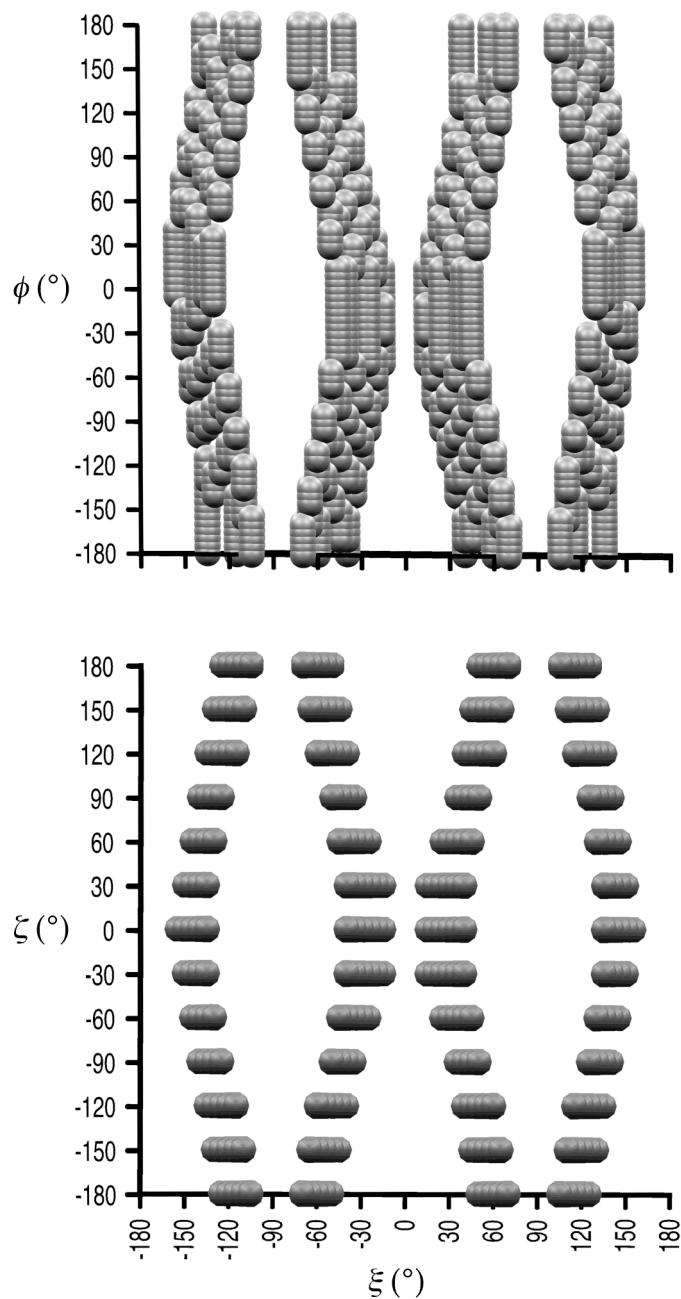


Figure 7.19. Isosurface of ${}^2J_{C1,C3}$ in the ϕ , ξ and ζ conformational space. Grey spheres are combinations of ϕ , ξ and ζ (definitions in text) that lead to a ${}^2J_{C1,C3}$ of 4.6 Hz as determined by DFT. The spheres have a 15° radius and are digitized every 5° in the ϕ and ξ dimensions and every 30° in the ζ dimension.

this is highly unlikely because it implies a static structure devoid of conformational averaging that fortuitously exists in the geometry of maximal coupling.

It is possible that the solution hypersurfaces from two separate couplings superimposed within the same model conformational space will not have any points of intersection. This is a result of the intrinsic assumptions of the model, the most likely of which being either that $\sigma = 0$ and/or presuming a single state model. If this assumption is relaxed so that σ can be searched across a specified range, the additional degree of freedom will expand the potential for overlap between the two coupling profile solution surfaces. Alternatively, if there is a significant intersection between the solution hypersurfaces of two coupling profiles, setting a finite range for σ with some reasonable minimum and maximum values can significantly refine the conformational ensemble of this intersection.

When considering the possibility of librational averaging, the n -dimensional conformational space is expanded by the additional conformational parameter σ , and as a result there is an analogous increase in the dimensionality of the solution topology. This parameter σ can be independently applied to each of the other conformational dimensions, in this case ϕ , ξ and ζ , such that the dimensionality n of the conformational space is now 6 and the set of parameters describing the solution is $\{\mu_\phi, \mu_\xi, \mu_\zeta, \sigma_\phi, \sigma_\xi, \sigma_\zeta\}$, potentially requiring up to 6 independent experimental couplings which each report on some aspect of the conformational space. The value of n is equivalent to the number of elements in the parameter set for a given conformational space model. When more than a single state model is considered for one or more of the dihedrals, the degree of n increases yet again. For example consider a model in which the populational profile about

ϕ is described by three states, while the other two dihedrals are both still single state. In this case, the dimensionality n of the conformational space is 10 and the set of parameters describing the solution is $\{\mu_{\phi_1}, \mu_{\phi_2}, \mu_{\phi_3}, \mu_{\xi}, \mu_{\zeta}, \sigma_{\phi_1}, \sigma_{\phi_2}, \sigma_{\phi_3}, \sigma_{\xi}, \sigma_{\zeta}\}$. It is common to utilize a three-state model of rotational isomerism in the treatment of rapid reorientation around an sp^3 hybridized dihedral in NMR studies^{2c}. If the assumption that the three states about ϕ are equally populated is dropped, the parameter set increases to 13 with the inclusion of w_{ϕ_1} , w_{ϕ_2} , and w_{ϕ_3} , the individual weightings for each of the three states. The sum of w_{ϕ_1} , w_{ϕ_2} , and w_{ϕ_3} in this case or N number of weightings for N total conformational states about a single dihedral is always unity, thus providing an additional constraint upon the available degrees of freedom in $\varphi_{\mu,\sigma}(\phi)$. Figure 7.20 shows the relationship of μ , σ , and w to the shape of the population distribution. An expansion of Equation 7.3 to include N number of weightings w about a dihedral θ is found in Equation 7.5.

$$\varphi_{w,\mu,\sigma^2}(\theta) = \sum_{i=1}^N \frac{w_i}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(\theta - \mu_i)^2}{2\sigma_i^2}\right) \quad (7.5)$$

Any sp^3 -hybridized dihedral modeled thus with one conformational distribution per rotamer has 8 undetermined conformational elements (9 minus 1 as a result of applying the equation: $1 = w_1 + w_2 + w_3$). Accordingly, a conformational analysis of the C1 through C3 hydroxyls using the above convention involves a solution set with 24 elements.

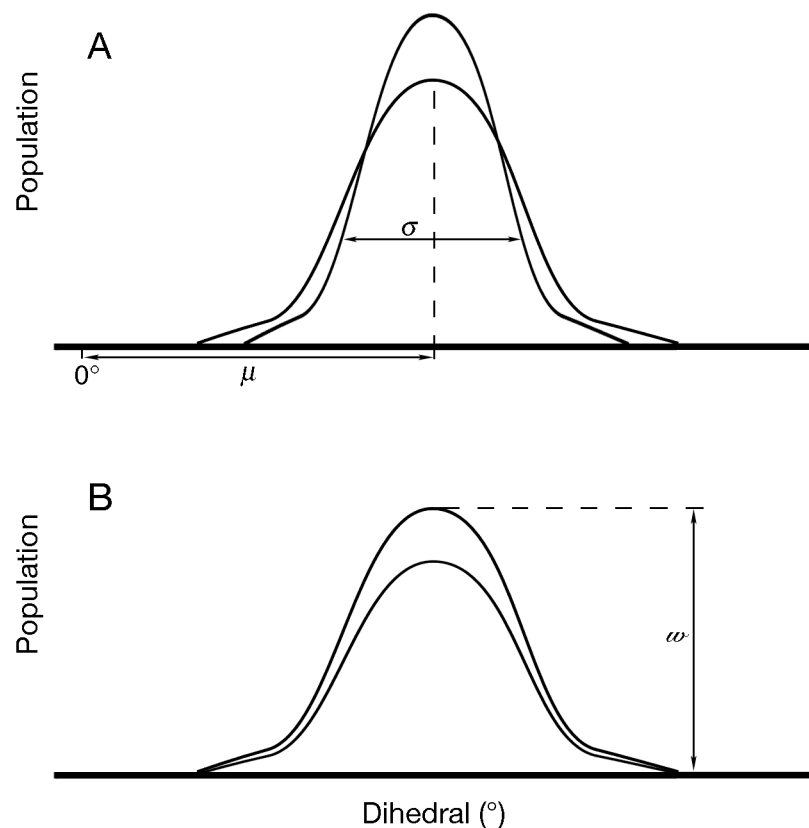


Figure 7.20. The Effect of μ , σ and w on Population Distribution. The parameter μ controls the location of the peak maximum across the rotational itinerary whereas σ dictates the peak width (A). The parameter w dictates the fractional weighting of a peak with respect to others about the same rotational itinerary (B). The two curves in (B) bound areas of different size whereas the two curves in (A) have the same area.

It can be appreciated how specific decisions relating to the dimensionality of a model conformational space can rapidly increase the number of couplings required to uniquely interpret an experimental coupling within the context of that space. Due to the unique chemical environment occupied by any two non-symmetrically disposed nuclei in a molecular framework, it is practicably impossible to find even two unique couplings whose total structural sensitivity is dependent upon an identical set of geometric parameters. There are multiple couplings with both overlapping and non-overlapping

conformational dependencies across a spectrum of structural elements. For example, there has been a detailed study of the dependence of direct ($^1J_{C1,H2}$), geminal ($^2J_{CH}$, $^2J_{HH}$), and vicinal ($^3J_{CH}$, $^3J_{HH}$) couplings on the C1-O1 conformation in methyl- α -D-xylopyranoside^{62a} and methyl- β -D-xylopyranoside^{62b} although it is known that a number of these couplings exhibit dual dependencies on other hydroxyl C-O rotamers¹¹ in a manner analogous to that observed for $^1J_{C1,C2}$ ³. Consequently, when trying to employ the maximum number of available couplings that report either in whole or in part on a particular geometric feature, it is often the case that information relating to additional geometric features becomes necessary.

As the number of dihedrals considered and thus the total conformational dimensionality n increase, so do the number of couplings required to uniquely define the conformational space. For example attempting to define the populational profiles about ϕ and ξ by using $^2J_{C1,C3}$ as one of the structural constraints requires some consideration of ζ if a precision greater than the dependence of $^2J_{C1,C3}$ on ζ is desired in the ϕ and ξ profiles. However many of the available couplings that report on the C3 hydroxyl conformation are also dependent upon the conformation of the C4 hydroxy (*vide supra*, §G, *re* $^2J_{C2,C4}$). Modeling the total conformational space of an aldohexopyranose for the 5 hydroxyl C-O rotamers and the exocyclic C5-C6 hydroxy methyl rotamer using the above conventions and neglecting correlation effects would require a minimum of eight couplings per dihedral for maximum total of 48 couplings. This number could be considerably reduced if some couplings returned information about multiple dihedrals such as the triple dependence of $^2J_{C1,C3}$ on the C1 through C3 hydroxyls. Nevertheless

there are 50 possible ${}^nJ_{CC}$, ${}^nJ_{CH}$, and ${}^nJ_{HH}$ ($1 \leq n \leq 3$) throughout a typical aldohexopyranose that could be potentially used in such an analysis. Karplus equations for ${}^3J_{HCOH}$ ³² and ${}^3J_{CCOH}$ ^{32a} have been reported, while ${}^2J_{COH}$ are readily measurable and parameterizable. This number grows to 68 if couplings involving hydroxyl protons are included (measurements must be made in DMSO or in low temperature water/acetone mixtures to avoid signal broadening due to exchange).

If a fully crossed exploration of the 6 dihedral conformational space were undertaken utilizing 30° increments similar to that conducted in this 3 dihedral study, 12⁶ or ~2.9 million DFT structures would result. Since only a small handful of couplings are expected to show significant dependence upon more than three dihedrals, a modular approach to coupling parameterization can be applied to overcome this computationally infeasible proposition. For example ${}^2J_{C1,C3}$ is not expected to show any direct dependence upon the C6-O6 dihedral. The hexopyranose ring can be artificially divided up into three coupling regions described by *i*) the dihedrals explored in this study, namely C1-O1, C2-O2, C3-O3, and two other regions *ii*) C2-O2, C3-O3, C4-O4, and *iii*) C4-O4, C5-C6, C6-O6. This would limit the number of DFT structure calculations to 3×12³ or 5,184, a third of which are presented here. This is very tractable considering the potential for extensive parallel computation.

These considerations allude to the high degree of conformation dependent correlation present throughout the molecular spin-network. The cross correlation between spin-spin coupling profiles is due to the reliance of distinct couplings upon similar coupling mechanisms which utilize identical molecular orbitals for the transfer of spin density information. This form of correlation is electronic in nature, whereas the final

facet of non-temporal structural complexity is realized through a consideration of spatial nuclear distribution in terms of conformational correlation. Suppose a conformational distribution in the rotameric regime ϕ_1 is itself a function of the conformation about the neighboring dihedral ξ . That is to say that the set of μ_{ϕ_1} , σ_{ϕ_1} , and w_{ϕ_1} are different for each rotameric regime ξ_1 , ξ_2 and ξ_3 . Applying the above modeling conventions to the cross correlation of these two dihedrals, $\varphi_{w,\mu,\sigma}(\phi) \star \varphi_{w,\mu,\sigma}(\xi)^9$, yields 25 undetermined conformational elements in the solution set (that is, 3 elements, μ , σ and w , per rotamer, times 3 rotamers per dihedral, times 2 dihedrals, minus 2 due to the use of the weighting equation: $1 = w_1 + \dots + w_{27}$, for each dihedral). Given a complete set of experimental couplings and corresponding conformationally comprehensive coupling parameterizations, an integrated picture of the correlated populational ensemble of molecular geometries is in principle obtainable.

Optimization Protocols for the Conformational Analysis of ϕ , ξ and ζ in Methyl- β -D-glucopyranose from Multiple Experimental Couplings — Falling short of a total conformational analysis of the 6 conformationally mobile dihedrals present in β -D-glucose, the question becomes whether it is possible to find a subset of the total couplings exclusively sensitive to the structural element(s) of interest that is *a*) experimentally accessible, and *b*) large enough to uniquely constrain the said structural element. Minimally it is possible to define a reduced dimensionality solution surface in the conformational space that can be further constrained through the use of other, non-scalar coupling probes of conformation. Alternatively, assumptions about the nature of some aspects of the model conformational space can reduce its dimensionality and allow for

the unique solution using the available experimental couplings. The total expression for the calculation of the ϕ , ξ and ζ dependent ${}^2J_{C1,C3}$ using the above outlined modeling conventions including correlation effects is

$${}^2J_{C1,C3} = \iiint_{2\pi} \left(\varphi_{w,\mu,\sigma^2}(\phi) \star \varphi_{w,\mu,\sigma^2}(\xi) \star \varphi_{w,\mu,\sigma^2}(\zeta) \right) f(\phi, \xi, \zeta) d\phi d\xi d\zeta \quad (7.6)$$

where $f(\phi, \xi, \zeta)$ is the parameterized coupling profile for ${}^2J_{C1,C3}$. A similar expression could be written for any coupling sensitive to one or more of the dihedrals ϕ , ξ and ζ . To find the solution set of conformational elements that satisfies all experimental couplings sensitive to ϕ , ξ , and ζ , one would simply need to vary the ensemble of μ , σ and w values for each rotamer simultaneously across the set of expressions such as (5) for which a coupling was known following an optimization protocol similar to that described previously for parameter refinement.

This can be put into practice for a model with three conformational states per dihedral each defined by a unique set of μ , σ , and w values using the set of experimental values for the couplings listed in Table 7.9. These couplings were measured at 600 MHz and room temperature in a 2H_6 -DMSO solvent on Methyl- β -D-Glucoside labeled with ${}^{13}C$ at either C1, C2, or C3.

The difference between values for a particular coupling made in aqueous conditions versus in DMSO are likely to be predominantly due to a difference in conformational profiles about the reorienting dihedrals between the different solvent systems, although some dependence upon the solvent dielectric is possible. The

conformational shift can be rationalized by recognizing that there is a change in the H-bonding inter- versus intra-molecular H-bonding potential for saccharides in DMSO when compared to an aqueous environment⁶⁸ (*vide infra*). There have been several experimental NMR studies probing intramolecular H-bonding in oligosaccharides through the use of a DMSO solvent^{32a,91}. In terms of the applicability of the DFT calculations to couplings measured in DMSO, values calculated for a particular structure *in vacuo* are expected to be slightly closer to the coupling for a similar geometry in DMSO ($\epsilon_r = 47.2$) than in water ($\epsilon_r = 80$).

Theoretical parameterizations for the couplings listed in Table 7.9 were conducted in a manner analogous to $^2J_{C1,C3}$. There are more couplings than are listed in Table 7.9 that are also sensitive to one or more of the dihedrals ϕ , ξ , and ζ , however these profiles were sufficiently dependent upon other conformational parameters such as the C4-O4 dihedral to preclude their use in this study.

Since there are only 15 and 10 experimental couplings used in the initial and final analyses respectively, a number of simplifying assumptions must be made to achieve a unique solution. The first assumption is the neglect of correlated conformational contributions. The result of this is that the solution will represent an averaging of the correlated contributions to a particular rotameric state. The specific envelope of the averaging is dependent upon the topology of the relevant coupling profiles and is not necessarily expected to be linear in nature. When cross-correlation is removed the populational weighting term in Equation 7.6 becomes the simple product of the three normal population distribution functions. Even with this assumption there remain 24 undetermined elements in the conformational solution set (μ , σ , and w values for each of

the three rotamers, times three rotamers, times three dihedrals, minus one w value per dihedral due to the use of the weighting equation: $1 = w_1 + w_2 + w_3$, for each dihedral). The second assumption that the standard deviation for all states is a reasonable arbitrary value, in this case 20° , further reduces the number of undetermined elements in the conformational solution set to 15. The remaining degrees of freedom (μ and w for the three states about each dihedral) are varied in an attempt to minimize the χ^2 between the couplings calculated for the particular μ and w values by integration across the full 360° itineraries of ϕ , ξ , and ζ and the corresponding experimental couplings. Either stochastic, deterministic or a combination of the two optimization protocols can be used, with stochastic methods used initially and deterministic algorithms used for refinement.

Since we are trying to simultaneously solve 10 equations in 15 unknowns the solution set is potentially a 15 dimensional hypersurface, however due to the multiple conformational dependencies of the Karplus profiles employed and the specific nature of the experimental couplings used for the fit it may be possible to find a single exact solution. If this method does not produce an exact solution (meaning that $\chi^2 > 0$), as previously indicated, it is a result of one of the assumptions (or due to one or more intrinsic sources of error, *vide infra*). If this is the case the restrictions on the assumed values of σ can be relaxed and optimization can proceed from a set of reasonable starting values in which a revolving subset of conformational elements is allowed to optimize iteratively. If there is a good degree of confidence in the initial values for the various μ , σ , and w values (as derived from molecular dynamics simulations for example), a manual form of rationally guided optimization utilizing alternating stochastic and

deterministic optimization algorithms is preferred to prevent the fitting trajectory from potentially jumping over a barrier into a region of the solution set that does not reflect physical reality. For example, there may be either true or local minima along the optimization trajectory that involve unrealistically small σ values that can be manually avoided.

A nuanced consideration here is the choice of the dependent versus independent weighting terms in the context of which are being actively optimized. That is to say, for a set of three weightings whose sum is unity, w_1 , w_2 and w_3 , the optimization can proceed, for example, by directly varying w_1 and w_2 , while w_3 is subject to the values for the other two terms. This could potentially lead to a problem if given the optimal value for w_3 , the current parameter set sits on a saddle point of the optimization hypersurface in the w_1 and w_2 dimensions, one optimization trajectory leading to the correct solution set and the other to a spurious one. A change in w_3 could potentially skew the trajectory inappropriately. If only w_1 is an actively optimized parameter, any change in w_1 will change w_3 inappropriately. In order to maintain the correct value for w_3 , changes in w_1 must accompany equal and opposite changes in w_2 , requiring both weighting terms w_1 and w_2 to be simultaneously active, however since in this case w_3 is the dependent variable, it is impossible to have w_1 and w_2 active and w_3 inactive. In practice it is difficult to anticipate the optimization trajectory and if optimization results are poor, particularly with respect the weightings, it may be wise to re-conduct the optimization with inverted dependent and independent weighting term selections prior to a re-assessment of the more fundamental assumptions of the model, in this hypothetical case

by making w_1 active and independent, w_2 active and dependent, and w_3 inactive and independent.

Another valid concern is that coupling profiles with the largest dynamic range may be over emphasized in terms of the χ^2 . Deviations from the experimental coupling in couplings with small dynamic ranges contribute less to the χ^2 despite the fact that they may be more significant in conformational terms than deviations of an equivalent absolute magnitude in coupling profiles with a larger dynamic range. This is subject to the relative regional topologies of the individual coupling profiles explored. For example, DFT calculations show that $^1J_{C1,H1}$ and $^1J_{C2,H2}$ are both exquisitely sensitive to ϕ , ξ , and ζ in **9C** with dynamic coupling ranges of 26.2 Hz and 22.6 Hz respectively, however the distribution of these couplings is asymmetric about the rotational itineraries and very susceptible to correlated conformational effects between the nearby hydroxyls. This leads to difficulty in an accurate parameterization of these couplings. If the parameterization is not very good, and there is large dynamic range that is potentially contributing to the optimization of the overall χ^2 , the contribution from these couplings can swamp out the contributions from couplings with smaller dynamic ranges and lead the optimization trajectory astray. Even if the parameterization is fairly good, contributions from these couplings will still overwhelm couplings with smaller dynamic ranges and dictate the optimization gradients controlling the trajectory. There are a number of methods for addressing this problem. The simplest solution, and the one adopted here, is to completely remove the couplings from the analysis. This can be problematic because it eliminates a potentially useful source of information, however if the ensemble of couplings is large the overall impact of losing the coupling(s) is minimized and

meaningful results can still be obtained. Alternatively, if the parameterization is reasonably good, the problem can be addressed by minimizing a normalized form of χ^2 . Individual contributions to the total χ^2 are normalized against the dynamic range of that coupling profile prior to calculation of the squared difference, then summed and divided by N , the number of couplings. The result is a deviance from the experimental couplings as a percentage of the largest possible deviation across the ensemble of couplings. A value of zero is an exact fit. A further step would be to assign specific weightings to the squared differences for each coupling profile based upon relative confidence in the robustness of the parameterizations. An assessment of error in the parameterization of the DFT data could be applied to this weighting term, such as a scalar factor based on the standard deviation, or a parametric transformation based upon the individual error functions for each coupling profile (*vide infra*). A third solution is to remove the parameterization step altogether and fit the experimental couplings from the DFT data directly using interpolation algorithms.

A minor consideration is the validity of the parameterization of a particular coupling derived from calculations conducted on a reducing sugar to the conformational analysis of a sugar containing a methyl aglycone. In general this seems justified in light of the distal proximity of the methyl group in relation to the nuclei involved in the various couplings employed in the study. Certainly couplings to the methyl carbon could be useful both in terms of the conformational analysis of the dihedrals considered here, but also as a bridge towards understanding glycosidic conformation.

A final assumption intrinsic to this method of revolving the actively optimized parameters is that the correct solution can be reached from the initial conditions by

simultaneously varying 10 or less parameters. This technique breaks down if access to the global minimum lies on the other side of a 11 or higher dimensional barrier, however it is reasonably unlikely that the conformational hypersurface is this complex. This possibility is exquisitely dependent upon the optimization trajectory and therefore upon the choice of initial conditions which can be varied to hopefully avoid the problem if it is suspected to exist. A computationally intensive solution to this problem is to calculate the entire grid hypersurface and then manually select the solution set with a minimal χ^2 value that agrees with chemical intuition (for example there are likely to be solution sets with a small χ^2 but which display unreasonable populational distributions).

Table 7.10 contains experimental spin-spin couplings sensitive to ϕ , ξ , and ζ orientation in $^2\text{H}_2\text{O}$ and $^2\text{H}_6\text{-DMSO}$ and the corresponding coupling values calculated using a variety of different conformational models, whereas Table 7.11 contains the specific μ , σ , and w values for each model set beside a graphical representation of the populational profiles (Figure 7.21). Three of these models came from a series of molecular mechanics and molecular dynamics simulations conducted on *O*-methyl- β -D-glucopyranoside. These were conducted in tandem with the fitting procedure to serve as benchmarks and as a potential source of appropriate initial parameters for the fitting of the experimental data, and will be briefly discussed first. Because these simulations were run as a part of a larger study and were primarily conducted by others, a full description of the computational details is omitted.

The first was a molecular mechanics simulation run at 300K for 10k ps sampled every ps using the MM3 force field⁹⁸ and an atomic solvent parameter (ASP)⁹⁹ set for DMSO ($\epsilon_r = 47.2$). The trajectory was typified by a highly correlated and essentially

TABLE 7.10. SPIN-SPIN COUPLING CONSTANT COMPARISONS BETWEEN
EXPERIMENT AND VARIOUS CONFORMATIONAL MODELS.

Coupling	H ₂ O	DMSO	Initial	$\sigma = 20^\circ$ (I)	Iterative C (II)	Iterative U (III)	MM3	CHARMm	Amber
¹ J _{C1,C2}	46.0	47.1 (1)	46.0	46.8	(45.0)	(46.1)	43.7	43.6	45.3
¹ J _{C2,C3}	38.8	39.2	40.3	41.0	38.9	39.1	37.7	39.2	38.4
¹ J _{C1,H1}	161.3	157.2	160.9	163.4	(158.3)	(157.2)	159.5	157.9	159.5
¹ J _{C2,H2}	145.0	141.7	147.7	143.5	(150.4)	(151.6)	150.0	151.0	150.1
² J _{C1,C3}	4.5	4.5	4.6	4.4	4.7	4.7	5.3	5.7	5.2
² J _{C1,H2}	-6.3	-6.0 (0)	-5.4	-4.9	-5.4	-5.5	-5.7	-5.5	-5.5
² J _{C2,H1}	0.0 (pr)	0.0 (pr)	2.4	2.0	(2.3)	(1.9)	2.6	3.1	2.6
³ J _{H1,H2}	8.0	7.5 (1)	7.3	7.5	7.5	7.6	7.7	7.3	7.3
² J _{C2,OH2}		-2.5 (1)	-2.3	-2.3	-2.2	-2.3	-2.2	-2.3	-2.3
² J _{C3,OH3}		-2.6 (3)	-2.5	-2.5	-2.5	-2.5	-2.5	-2.6	-2.6
³ J _{C1,OH2}		2.5 (0)	4.0	1.9	2.6	2.5	1.2	-0.7	0.5
³ J _{C2,OH3}		2.3	3.3	3.2	2.3	2.3	1.7	3.5	3.0
³ J _{C3,OH2}		3.4 (1)	3.2	5.6	2.8	3.4	1.0	2.1	3.0
³ J _{H2,OH2}		4.9 (3)	5.8	4.8	(8.2)	(7.5)	12.9	13.9	11.0
³ J _{H3,OH3}		4.7 (3)	5.7	5.3	4.8	4.7	1.3	3.7	4.8
15 couplings		χ^2	63.5	56.8	98.7	108.9	180.3	205.7	130.7
		RMS	2.1	1.9	2.6	2.7	3.5	3.7	3.0
		SD	9.4	9.8	19.5	25.0	22.7	28.8	19.8
10 couplings		χ^2	6.1	10.7	0.7	0.3	21.9	15.7	6.1
		RMS	0.8	1.0	0.3	0.2	1.5	1.3	0.8
		SD	0.8	1.7	0.1	0.1	3.6	3.1	1.2

All coupling values and errors are reported in Hz. Descriptions of models **I** through **III** are found in the text. Numbers in parenthesis beside experimental couplings are the standard deviation in the last digit. Experimental couplings to hydroxyl protons were not measurable in H₂O due to rapid exchange. The couplings reported for the various models were back calculated using the coupling profile parameters found in Table 7.9 and the conformational parameters found in Table 7.11. Model **I** was optimized by minimizing the 15 coupling χ^2 (value in bold), whereas models **II** and **III** were optimized through the minimization of the 10 coupling χ^2 (values in bold). The couplings for models **II** and **III** in parenthesis were excluded from the calculation of the 10 coupling χ^2 .

TABLE 7.11. CONFORMATIONAL MODEL PARAMETERS FOR ϕ , ξ AND ζ IN **9C**.

Dihedral	Rotamer	Parameter	Initial	$\sigma=20^\circ$ (I)	Iterative C (II)	Iterative U (III)	MM3	CHARMm	Amber
ϕ	0° to 120°	σ	20	20	17	18		13	22
		μ	60	71	104	149		155	122
		w	1/3	0.17	0.21	0.23	0.00	0.16	0.28
	120° to 240°	σ	20	20	29	28	17	9	14
		μ	180	120	180	146	184	168	165
		w	1/3	0.06	0.63	0.38	0.90	0.83	0.70
	240° to 360°	σ	20	20	25	24	33	41	13
		μ	300	326	275	245	277	281	294
		w	1/3	0.77	0.16	0.39	0.10	0.01	0.01
ξ	0° to 120°	σ	20	20	21		34	21	25
		μ	60	55	93		47	51	54
		w	1/3	0.25	0.35	0.00	0.03	0.53	0.35
	120° to 240°	σ	20	20	21	21	16	22	16
		μ	180	145	175	130	297	180	184
		w	1/3	0.32	0.08	0.50	0.68	0.02	0.20
	240° to 360°	σ	20	20	20	20	25	20	22
		μ	300	306	309	290	306	314	309
		w	1/3	0.42	0.57	0.50	0.29	0.45	0.45
ζ	0° to 120°	σ	20	20	24	27		18	19
		μ	60	63	68	60		60	60
		w	1/3	0.37	0.18	0.18	0.00	0.03	0.17
	120° to 240°	σ	20	20	29	34	50	23	24
		μ	180	163	199	175	286	194	192
		w	1/3	0.27	0.24	0.24	0.11	0.45	0.33
	240° to 360°	σ	20	20	24	24	16	21	23
		μ	300	316	299	294	306	292	296
		w	1/3	0.36	0.58	0.58	0.88	0.52	0.50

Definitions of the ϕ , ξ and ζ dihedrals, the s , m and w conformational parameters along with descriptions of models **I** through **III** are found in the text. The weighting term (w) is reported in fractional units while s and m are reported in degrees. When the weighting term for a particular rotamer was zero, the s and m parameters were omitted.

single state model for ϕ , ξ and ζ , in what is certainly an over representation of the counter-clockwise (*ccw*) hydrogen bonding pattern (greater than 92% in the *ccw* pattern) with the corresponding μ values of an ideally staggered rotamer. This is likely due to an intrinsic overestimation in the force field of either the energetic boundaries to rotamer transition, the description of H-bonding energies, or both. The ϕ dihedral was centered at

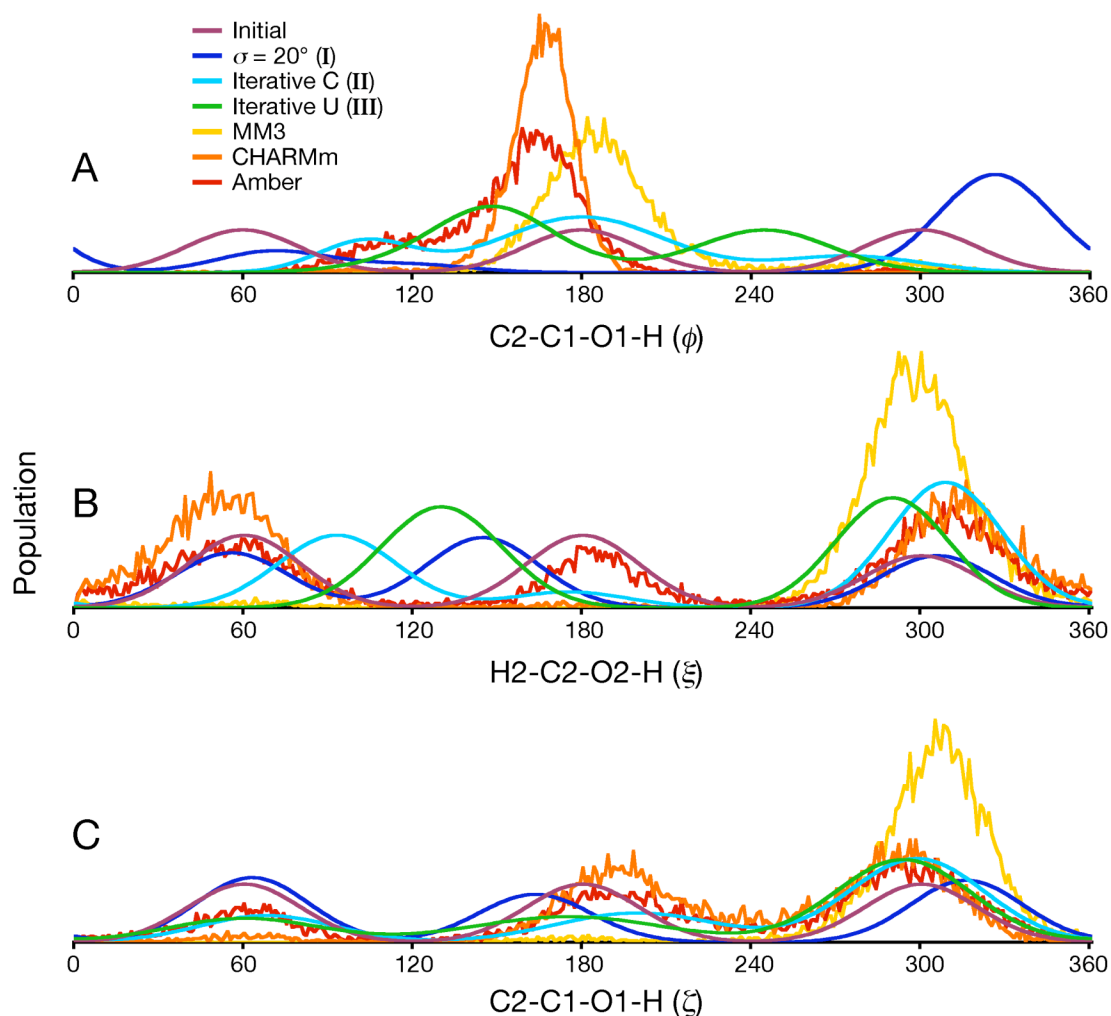


Figure 7.21. Population Histograms for Conformational Models about the Dihedrals ϕ , ξ and ζ in 9^C . The dihedrals for ϕ (A), ξ (B) and ζ (C) are shown in degrees whereas population is shown in arbitrary units. The raw data for the three simulations is shown rather than the best fit gaussian models for the simulations represented in Tables 7.10 and 7.11. The bin size for the three simulation histograms was every 1° . The definitions for models **I**, **II** and **III** are given in the text.

the typical *exo*-anomeric orientation in which $C2-C1-O1-H \approx 180^\circ$. Despite the drawbacks of this method, when the couplings were back calculated using the parameters in Table 7.9 the MM3 trajectory returned a slightly better fit to the experimental

couplings than the CHARMM trajectory as measured by the χ^2 and the standard deviation in the squared differences (Table 7.10).

The second trajectory utilized the CHARMM/CSFF force field¹⁰⁰ (v c28b1) and was also ran at 300K for 10k ps and sampled every ps. The simulation was run with constant energy and constant volume (NVE) with periodic boundary conditions. There was a 300 ps annealing period at the front of the trajectory and a TIP3P water model was used¹⁰¹. The CHARMM forcefield did considerably better in terms of chemical intuition, with a roughly equal representation of the clockwise (*cw*) and *ccw* H-bonding patterns, however there was an extremely small amount of intermediate, non-intramolecular H-bonded conformations. The ϕ dihedral was represented by a single state that was within a narrow range ($\sigma_\phi \approx 9^\circ$) centered at C2-C1-O1-H $\approx 165^\circ$, signifying that the C1-O1 dihedral was essentially static during the course of trajectory. The ξ and ζ dihedrals each had two roughly equally populated regions that correspond to the *cw* and *ccw* idealized H-bond geometries.

The third simulation was conducted with the Amber-8.0/GLYCAM-04 force field¹⁰². The computational details were similar to the CHARMM trajectory except that the simulation was run with constant energy and pressure (NPE). The trajectory exhibited slightly elevated ξ and ζ rotamer populations characteristic of the *cw* and *ccw* H-bonding patterns, however the three state rotamer distribution for these two dihedrals was fully crossed and essentially stochastic, meaning that all 9 conformational combinations of ξ and ζ were represented to some substantial extent. The ϕ distribution had a major peak ($w_{\phi_1} = 72\%$) roughly concurrent with that found in the CHARMM trajectory ($\mu_{\phi_1} = 165^\circ$),

however with a broader range ($\sigma_{\phi_1} = 14^\circ$). There was also a broad shoulder to this main peak extending across the eclipsed $\phi = 120^\circ$ region that accounted for the additional 27% of the population. The Amber trajectory showed the best fit to the experimental couplings of the three simulations by all measures (Table 7.10).

The initial structural parameters used for the first round of optimization will be termed the “idealized model” which included a three state model for all three dihedrals. Each state was equally weighted and had the same peak width ($\sigma = 20^\circ$) centered at the three idealized rotamer values for μ (60° , 180° , 300°). As expected, this model did not return drastically poor χ^2 or rms values when couplings were back calculated using the appropriate coupling profiles (63.5 and 2.1 Hz respectively), however refinement was clearly necessary. It is noteworthy (perhaps ironically so) that this platonic model was a significantly closer fit to the experimental data than any of the simulations described above, a fact that leads one to question the validity of the force constants intrinsic to the various force fields, and the results that they return. This implies a system dependent uniqueness that defies the formulation of generally applicable force-fields from which precise quantitative conclusions may be drawn.

The first attempt to optimize the parameters of the initial model involved keeping $\sigma = 20^\circ$ in all cases while optimizing the other parameters $\{\mu_{\theta_{1-3}}, w_{\theta_{1-3}}\}$ on a rotating basis using all 15 couplings in Table 7.9 as constraints. These results will be referred to as optimization **I**. Integrations were performed from -60° to 420° to minimize boundary effects discussed above, and all active μ parameters were allowed to vary across their respective 120° rotameric ranges centered at 60° , 180° and 300° . The weighting terms for all $\mu = 60^\circ$ and 180° rotamers were allowed to independently vary between 0 and 1 while

the w value for $\mu = 300^\circ$ was a dependent term whose value was arithmetically linked to the weighting terms from the other two rotamers and whose range was constrained to the range 0 to 1 as well. This approach returned slightly better χ^2 and RMS values than the initial conditions (5.6 and 1.9 Hz respectively) and achieved reasonable results for ξ and ζ , however the μ_ϕ values deviated considerably from expectations based upon both the *exo*-anomeric effect, and general considerations regarding standard staggered rotameric orientations. The most practical aspect of this optimization was in the identification of which geometric constraints were and were not useful towards this purpose. The following couplings were identified as being inappropriate for use in the optimization based upon a variety of factors including the quality or applicability of parameterization or the accuracy of the experimental coupling used as a constraint: $^1J_{C1,C2}$, $^1J_{C1,H1}$, $^1J_{C2,H2}$, $^2J_{C2,H1}$ and $^3J_{H2,OH2}$.

Subsequent attempts at optimization used an iteratively revolving subset of the 27 parameters $\{\mu_{\theta_{1-3}}, \sigma_{\theta_{1-3}}, w_{\theta_{1-3}}\}$ against the 10 remaining couplings across the three dihedrals ϕ , ξ and ζ . The integration and parameter ranges followed the conventions described above except as follows: there were two entirely separate series of optimizations; the first used the rotameric μ boundaries described above and will be called the ‘constrained iterative’ optimization (II), while the second which allowed all three μ values to vary across the entire 360° itinerary for a particular dihedral will be called the ‘unconstrained iterative’ optimization (III). Separate optimizations were conducted using initial conditions from the idealized model described above and models derived from the Amber and CHARMM trajectories, either with or without rotameric

boundary constraints on μ . Optimizations conducted using these three different initial conditions all converged on the same parameter values within statistical significance (ANOVA, $p \gg 0.05$) for a given set of μ boundary conditions, and will be considered identical.

Before I enter into a detailed discussion of the population distributions of **II** and **III** on a dihedral-by-dihedral basis, I wish to make a few general observations regarding the optimization results. The solution sets from **I** – **III** are considerably better fits to the ensemble of couplings when compared to any of the simulations, however it must be understood that this statement is based upon a comparison between the measures of error appropriate to the number of couplings utilized as constraints for that particular optimization; 15 in the case of **I** and 10 for both **II** and **III**. Since utilizing a different ensemble of couplings inherently imposes different constraints upon the optimization trajectories, the χ^2 , rms or SD are not comparable strictly speaking between **I** and either **II** or **III**. For example, the 15 coupling χ^2 for **I** was better than the initial conditions because it was the actual quantity that was being minimized, however it became worse in **II** and **III** because it was the 10 coupling χ^2 that was being actively minimized in those optimizations independent from the behavior of the 15 coupling χ^2 (Table 7.10). This being said, both **II** and **III** demonstrated 10 coupling standard deviations of 0.1 Hz, which is within the intrinsic error of the experimental couplings. The 10 coupling χ^2 from **III**, was roughly half that of **II** indicating the slightly better fit to the experimental couplings. Amongst the six models presented (three optimizations and three simulations) these were the only two refined models whose SD came even within an order of magnitude of the experimental error. While there are numerous potential sources of error

that make it difficult to assign an exact error to the numbers reported for **I** – **III** (*vide infra*), it is worth noting here that the primary goal of any model is not to reproduce all data without deviation, rather modeling is a means by which a meaningful understanding of the data can be gainfully applied. Despite the fact that an absolutely precise fit to the experimental data was not achieved ($\chi^2 > 0^\circ$) for any of the models, patterns between the various model distributions suggest some common conclusions that are in accord with chemical intuition.

From panel C in Figure 7.21 it can be seen that the **II** and **III** optimizations are nearly identical in their distributions about ζ and that neither deviated grossly from the Amber and CHARMM models for ζ in that the population distribution about the rotational itinerary is characterized by three states, however the experimental data indicate that the population distribution from the *trans* ζ rotamer is broader than predicted by MD, implying lowered energetic barriers to rotameric transitions involving this region. This will be discussed in more detail in the context of rotameric transition pathways and potential intramolecular H-bonding geometries (*vide infra*), however this exemplifies a theme throughout the conformational analysis of ϕ , ξ and ζ in DMSO, namely that the energetic boundaries that are conventionally thought to separate distinct C-O rotameric states in aqueous solution are attenuated by contravening forces in DMSO. These forces derive from both bulk and direct solvent-solute interactions that arise from the polar aprotic nature of DMSO. Specifically these forces lead to *i*) a heightened potential for intramolecular H-bonding which can in turn be modulated by *ii*) direct solute-solvent interactions unique to DMSO, and electronic considerations such as *iii*) altered dipolar and *iv*) lone pair repulsive energies as well as *v*) steric and

hyperconjugative interactions between O2 lone pairs and the proximal C-C and C-H bonding and anti-bonding orbitals (*vide supra*). The *gauche*⁺ ξ rotamer is less populated than the other two in line with the reduced capacity of this O3 hydroxyl orientation to participate in either *cw* or *ccw* H-bonding networks and the dual repulsive 1,3-diaxial interactions between the hydroxyl proton and both H2 and H4.

In contrast to the tight agreement between the **II** and **III** population distributions for ζ , the population distribution about ξ that best fits the ensemble of experimental couplings is distinct between the two optimization trajectories. The **II** trajectory settled into a three state model about ξ in which the *gauche*⁺ ($\mu_{\xi_1} = 93^\circ$) and *gauche*⁻ ($\mu_{\xi_3} = 309^\circ$) rotamer peaks were roughly in a 3:5 ratio with a third broad and sparsely populated $\mu_{\xi_2} = 175^\circ$ peak ($\sigma_{\xi_2} = 21$ Hz, $w_{\xi_2} = 0.08$), while the **III** trajectory converged to a two state model between the *trans* ($\mu_{\xi_1} = 130^\circ$) and *gauche*⁻ ($\mu_{\xi_2} = 290^\circ$) peaks which were populated in an even 1:1 ratio. It can be seen from panel B of Figure 7.21 that the population distribution of **C** appears to be in a local minimum created the imposition of rotameric boundary constraints, and that the weighting between the *g*⁺ and *t* peaks is an approximation of the $\mu_{\xi_1} = 130^\circ$ state in the **III** distribution. Despite the differences between the **II** and **III** profiles about ξ , there are some strong similarities between the two. Both models show major peaks that are well within a range that ideally situates the O2 hydroxyl to engage either the O1 lone pairs in a *ccw* H-bonding pattern or the O3 lonepairs in a *cw* H-bonding network. The difference in the ξ profiles between **II** and **III** is likely guided by the divergence between the two models in terms of their respective ϕ

population distributions as mediated by coupling profiles that show a dual dependence on the two dihedrals.

The ϕ population distribution model delivered by the analysis of the experimental coupling ensemble is most simply described as a continuum model spanning the *trans* rotamer, although there are clearly discernable nuances between **II** and **III** for this dihedral. For this reason the two models will be considered separately. The **II** model has three peaks, the largest and broadest of which is in the *trans* rotamer centered at the canonical *exo*-anomeric orientation ($\sigma_{\phi_2} = 29^\circ$, $\mu_{\phi_2} = 180^\circ$, $w_{\phi_2} = 0.63$). This broad peak extends into the two adjoining rotameric regimes where it blends with the g^+ and g^- peaks which are themselves shifted towards their boundaries with the *t* rotamer ($\mu_{\phi_1} = 104^\circ$, $w_{\phi_1} = 0.21$ and $\mu_{\phi_3} = 245^\circ$, $w_{\phi_3} = 0.16$ respectively). Interestingly the g^+ peak is roughly coincident with a similar shoulder in the Amber model for ϕ (Figure 7.21A). While the furthest extent of this μ_{ϕ_1} conformational ensemble extends into the ϕ hydroxyl orientation in which the *exo*-anomeric effect is presumably negated ($C2-C1-O1-H \approx 60^\circ$ and $O5-C1-O1-H \approx 180^\circ$), it is tempting to suggest that this conformation is the result of a *cw* O1-H \rightarrow O2 H-bond, however inspection of the Amber trajectory for conformations in which $\phi < 120^\circ$ shows minimal correlation with the O2 hydroxyl orientation (data not shown). The g^- peak is extremely broad for its relative population. This implies a reduced energetic barrier to transitions between the *trans* and *gauche*⁻ regimes. This point will be revisited in a discussion of the various interpretations of the *exo*-anomeric effect following a discussion of the ϕ population distribution in **III**.

The ϕ population distribution in **III** converged to two states with maxima at $\mu_{\phi_1} \approx 150^\circ$ (the μ_{ϕ_1} peak is the superposition of the g^+ and t peaks from the initial conditions) and $\mu_{\phi_2} \approx 240^\circ$, and a ratio between the two of roughly 5:3 (Table 7.11). These two broad peaks are confluent and span the canonical *exo*-anomeric orientation C2-C1-O1-H = 180° , which is the approximate weighted average between the μ_{ϕ_1} and μ_{ϕ_2} geometries. This model, which is the best fit amongst the models explored to the most reliably parameterized of the experimental couplings, seems at first to conflict with the traditional conformational preferences inherent in the description of the *exo*-anomeric effect. Despite this there are a number of considerations that may lend credence to this model.

First of all, while recognizing the dangers of a naïve confidence in the quantitative results of molecular dynamics simulations, it is notable that *both* the Amber and CHARMM simulations show population maxima in the *trans* ϕ regime of approximately 165° , merely 15° from the center of the μ_{ϕ_1} state in **III**. As noted above, the Amber simulation shows significant population density throughout the range of the μ_{ϕ_1} state in **III**, whereas the CHARMM model exhibits a sharp decrease in population past $\phi \approx 150^\circ$.

A second consideration is the effects of the enhanced potential for intra-molecular H-bonding of saccharides dissolved in DMSO⁶⁸. A detailed discussion of the specific patterns, geometries and energetics of H-bonding will follow, however due to the unique lone pair electron distribution about O1 as a result of the *exo*-anomeric effect a geometry of $\phi \approx 150^\circ$ and 240° may be the ideal conformations to accommodate both an intra molecular H-bond donation from O2 to either of the two O1 lone pair lobes while

simultaneously capitalizing on the energetic stabilization offered by the *exo*-anomeric effect.

Alternatively, there may be a local solvation effect on the O1 hydroxyl in DMSO that specifically mediates the ϕ conformational preferences. It has been shown that the anomeric hydroxyl is the strongest H-bond donor in β -D-glucopyranose⁹², which along with the fact that DMSO is a stronger H-bond acceptor than either hydroxyl or water lone pairs suggests enhanced conditions for a persistent intermolecular H-bond. Despite the increased capacity to be an H-bond donor, the anomeric hydroxyl has a reduced potential to participate as an intra-molecular H-bond donor: the only hydroxyl within range of O1 that can act as an acceptor is O2, however the formation of this H-bond is hampered by the contravening geometric constraints of the *exo*-anomeric effect. These considerations taken together indicate a heightened potential of a persistent intermolecular H-bond between O1 and a DMSO molecule, which in turn influences the ϕ conformational preference.

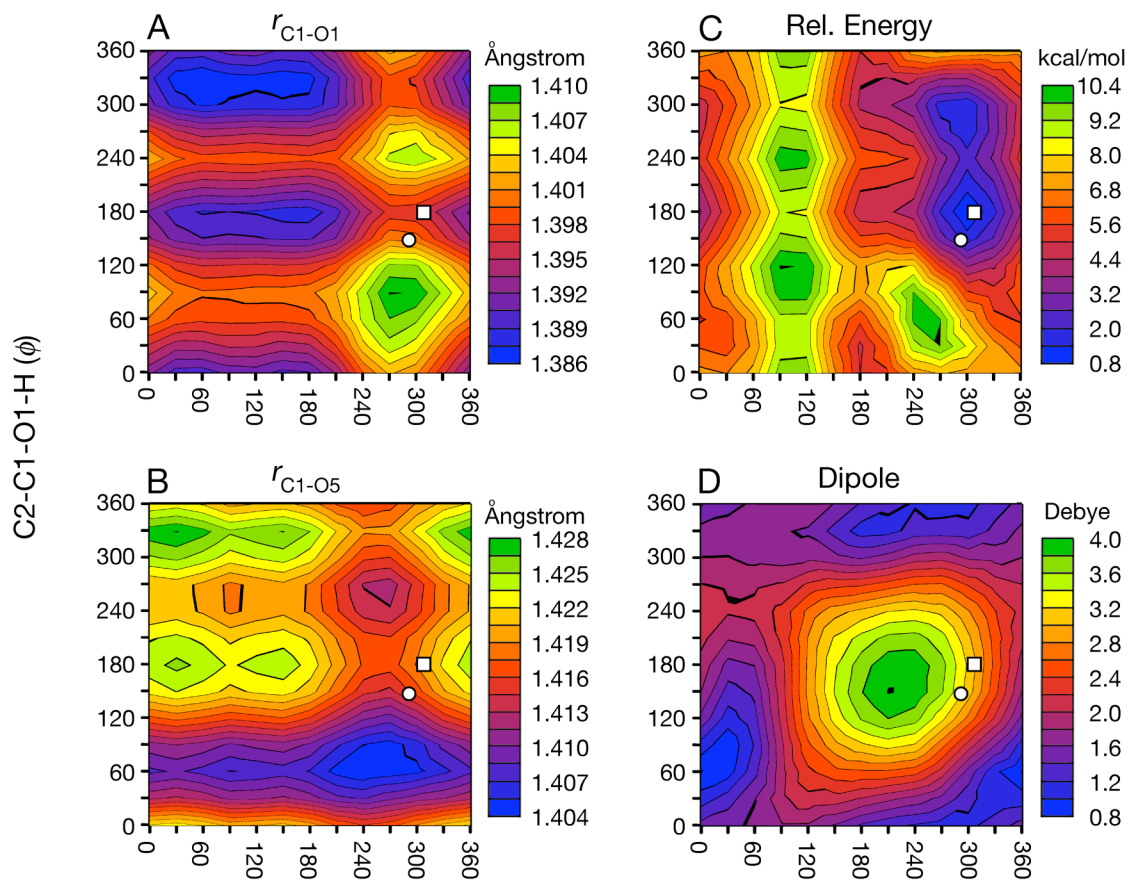
Finally a brief dissection of the various interpretations of the *exo*-anomeric effect and their agreement with model **III** may in fact lead to a more modern understanding of this phenomenon. The anomeric effect was first described in 1959^{94a}, and expanded to include the *exo*-anomeric effect ten years later^{94b}. There have been a variety of conceptual models invoked to account for the presumed orientation effect on the anomeric hydroxyl⁹⁵, which include arguments based on hyperconjugation^{95b-f}, favorable dipole interactions^{95g}, electrostatic stabilization^{95h} and other valence bond or steric descriptions^{95i-j}.

By far the most common description of the anomeric effect invokes some form of molecular orbital interaction leading to a stabilizing electronic delocalization⁹⁶. In terms of the *exo*-anomeric effect this manifests as a C1-O1 orientation in which one of the O1 lone pair orbitals is favorably aligned with $\sigma_{C1,O5}^*$ to allow for a hyperconjugative stabilization^{95a}. A recent combined DFT and NBO study concluded that such orbital interactions are the energetically dominant factor over electrostatic or steric considerations in dictating *exo*-anomeric behaviour^{95a}. It was demonstrated that the *endo*- and *exo*-anomeric effects do not compete energetically due to the mutual orthogonality of the orbitals involved in those two distinct processes, a conclusion contrary to a valence bond interpretation^{95i-j}. One drawback was that this study exclusively explored ideally staggered rotamers about C1-O1, precluding an evaluation of geometric nuances inherent in this phenomenon.

The dipole interpretation of the *exo*-anomeric effect considers fragment molecular dipoles whose directional Cartesian coordinates undergo vector subtraction as being the origin of the conformational stabilization^{95g}. This effect should be reflected in the magnitude of the total molecular dipole with the *exo*-anomeric stabilized geometries exhibiting a decreased dipole moment, however a map of the scalar molecular dipole as a function of ϕ and ξ (Figure 7.22D,H) showed an inverse trend upon comparison to the conformations predicted by **II** and **III**.

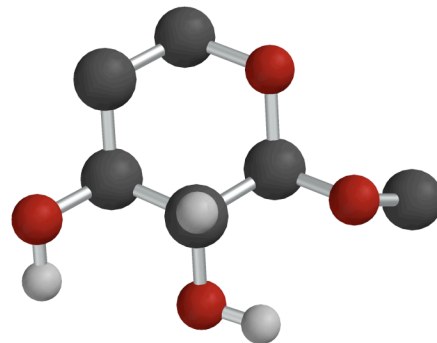
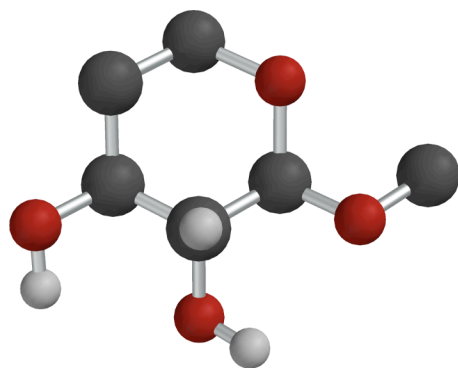
If the outcome of the *exo*-anomeric effect is predominately dictated by a $O1p \rightarrow \sigma_{C1,O5}^*$ hyperconjugation, then the effect should be acutely sensitive to the alignment between these two orbitals. The specific ϕ geometry that leads to the optimal

C2-C3-O3-H (ζ) = 300°



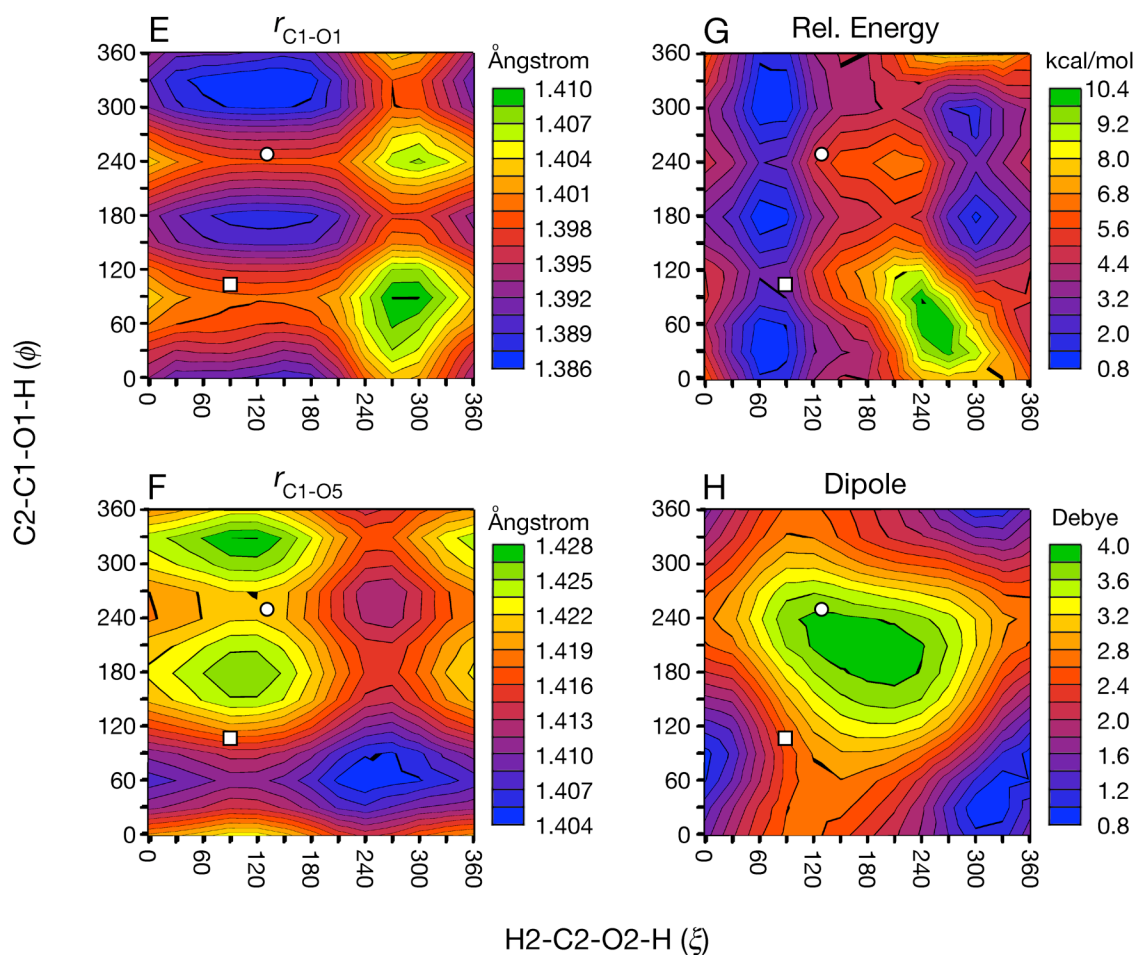
H2-C2-O2-H (ξ)

II (□)	ϕ	ξ	ζ	III (○)	ϕ	ξ	ζ
μ	180°	309°	299°	μ	147°	290°	294°
w	0.63	0.57	0.58	w	0.61	0.50	0.58



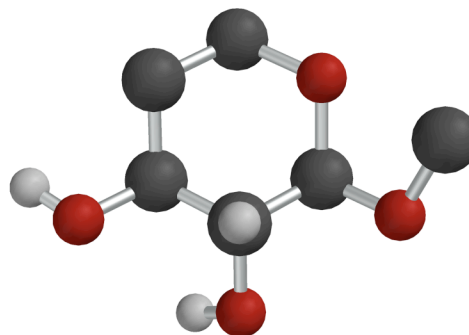
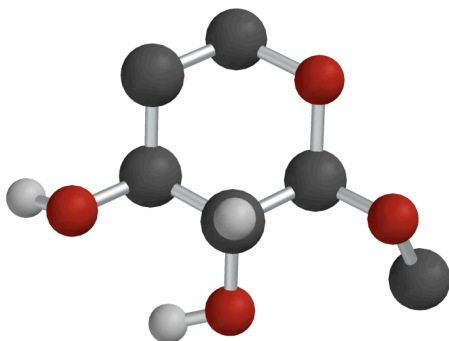
CCW

C2-C3-O3-H (ζ) = 180°



II (□)	ϕ	ξ	ζ
μ	104°	93°	199°
w	0.21	0.35	0.24

III (○)	ϕ	ξ	ζ
μ	245°	130°	175°
w	0.39	0.50	0.24



CW

Figure 7.22. Bond Length, Relative Energy, and Dipole Hypersurfaces as a Function of ϕ , ξ and ζ for the *ccw* and *cw* H-bonding Conformers in **9C**. The dihedrals ϕ and ξ are measured in degrees. The global dipole is represented as the non-directional scalar quantity. Putative *ccw* and *cw* H-bonding geometries are depicted for the **II** and **III** models and are indicated on the various hypersurfaces by open squares and circles respectively. The 0 kcal/mol global relative energy minimum used as a reference for panels C and G occurred in a slice through the ζ dimension not shown. The hydroxymethyl group, the O4 hydroxyl and all non-hydroxyl protons are omitted from the molecular models for clarity, except for H2 which was retained as a reference.

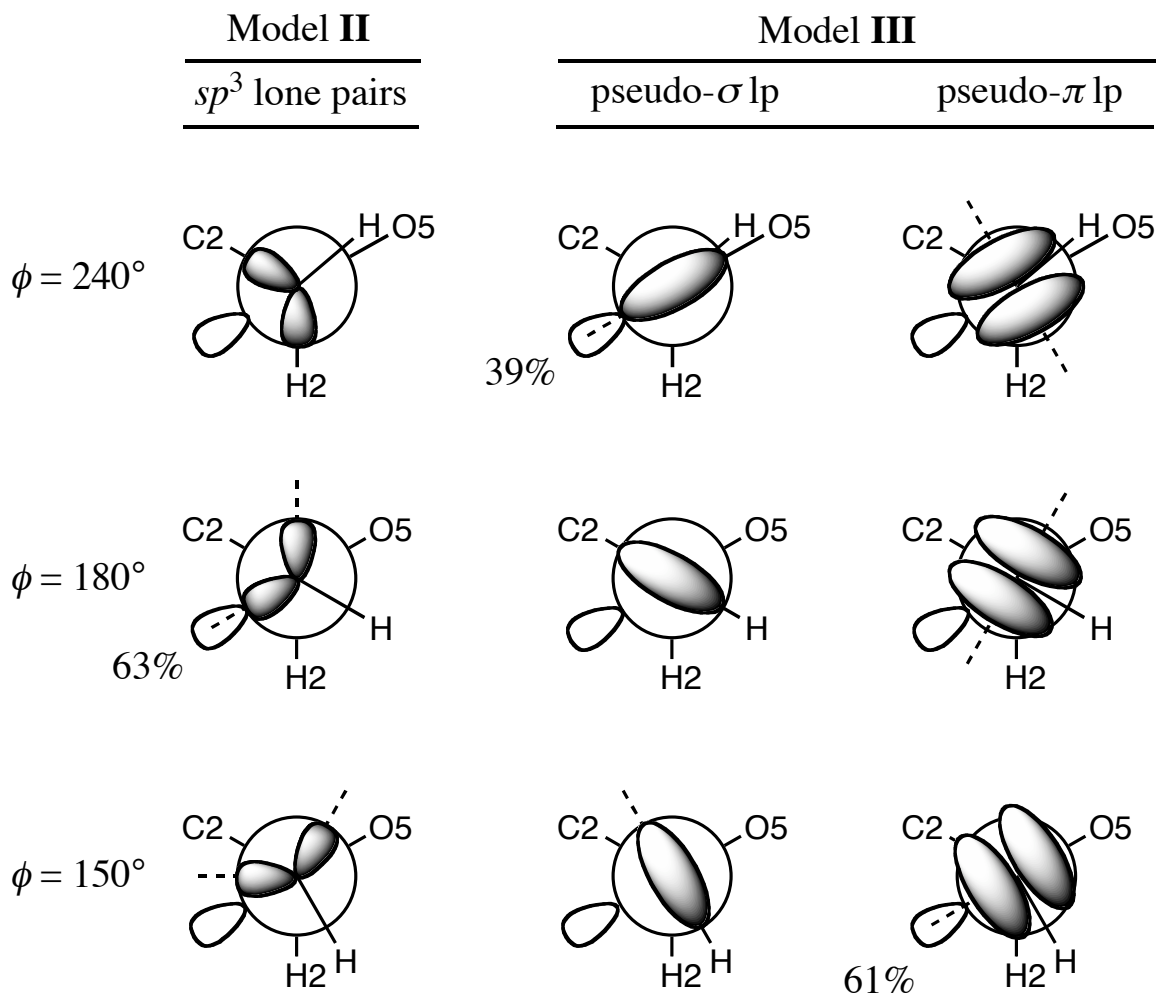
orbital interactions is dependent upon the specific electronic distribution of the pertinent O1 lone pair. The distribution of the O1 lone pair electron density is in turn sensitive to its concomitant participation as an acceptor orbital for an intra-molecular H-bond from the O2 hydroxyl proton. As mentioned above, model **III** predicts a $\mu_{\xi_2} = 290^\circ$ peak populated 50% of the time. The O2 hydroxyl is undoubtedly engaging O1 as an H-bond donor for the majority of the time that it is in this conformational range. While it is possible that in the absence of the other effect impinging upon the O1 lone pair, the optimal ϕ geometry for either the *exo*-anomeric effect or an O2-H \rightarrow O1 H-bond may be $\sim 180^\circ$, it is not a forgone conclusion that this is the optimal geometry for ϕ in the presence of both. The formation of an O2-H \rightarrow O1 H-bond may distort the O1 lone pair such that the optimal alignment between the O1 lone pair and $\sigma_{C1,O5}^*$ is now not $\phi = 180^\circ$, and conversely the presence of the *exo*-anomeric effect may distort the O1 lone pair such that the optimal alignment between this orbital and $\sigma_{O2,H}^*$ is also not $\phi = 180^\circ$.

There is evidence that suggests this may be the case. Based on a detailed analysis of hydroxyl lone pair and $\sigma_{O,H}$ localized orbitals from a sophisticated series of QMMD

simulations of β -D-glucose with 60 explicit water molecules, Suzuki reported that hydroxyls are weaker H-bonding acceptors than the oxygen of water⁹². This trait extends by inference to DMSO. This conclusion was based in part on the observed correlation between the contraction of lone pair Wannier functions and the degree to which the hydroxyl oxygen is participating in an H-bond. The hydroxyl lone pairs have more *p* character when participating as an acceptor in an H-bond and more *s* character when they are not. If O1 acts as an H-bonding acceptor to O2, leading to more *p* character in the O1 lone pair engaged, this lone pair will be distorted from the canonical tetrahedral geometry about O1. This same study also found that indeed the hydrogen bonding geometry about the hydroxyls in pyranose rings deviates from the tetrahedral arrangement⁹²: a typical pyranosyl hydroxyl oxygen participates in an average of two H-bonds at any given time, not the three that is characteristic of a tetrahedral array.

Inspection of the series 1 NBO partitioning for the O1 lone pairs in **2C** indicates that one of the lone pairs is described by >99% *p*-character while the other lone pair is roughly 50% *s* and 50% *p* for all geometries, with the average between the two approximating the conventional *sp*³ description. Since this computational series was conducted solely on staggered rotamers about the C1-C3 hydroxyls, it is difficult to say how this may vary by any better than a 120° resolution, however it is suggestive that differences in the electron distribution between the two lone pairs may affect the geometries of ideal alignment with $\sigma_{C1,O5^*}$. Scheme 7.4 outlines the interpretations of **II** and **III** for ϕ in terms of either the *sp*³ or the pseudo- σ and pseudo- π descriptions of the O1 lone pair electron densities respectively. The *exo*-anomeric effect for the $\mu_{\phi_2} = 180^\circ$,

$w_{\phi_2} = 0.63$ state in **II** is well described by the alignment of an sp^3 hybridized lone pair with the back lobe of $\sigma_{C1,O5}^*$. In contrast the $\mu_{\phi_1} \approx 150^\circ$, $w_{\phi_1} = 0.61$ and $\mu_{\phi_2} \approx 240^\circ$, $w_{\phi_2} = 0.39$ states of **III** are best described by the alignment of either a pseudo- π or pseudo- σ lone pair with the back lobe of $\sigma_{C1,O5}^*$ respectively.



Scheme 7.4. Orbital Interpretations of Models **II** and **III**.

A recent *ab initio* conformational study of the *exo*-anomeric effect in 2-methoxytetrahydropyranoside, a methyl- β -glucopyranoside analogue, correlated %*s*-character with C-O bond lengths about the acetal and the energetic stabilization of the

exo-anomeric effect^{95j}. Figure 7.22 depicts the variation in $r_{C1,O1}$, $r_{C1,O5}$, relative energy and the magnitude of the molecular dipole as a function of ϕ and ξ for $\zeta = 180^\circ$ and 300° . These ζ values were chosen because they roughly represent the two major peaks about this dihedral for both **II** and **III**. Additionally, from the specific pattern of conformations about the three dihedrals ϕ , ξ and ζ , and the understanding that DMSO increases the potential for intramolecular H-bonding, a picture of the hydrogen bonding patterns present in solution can be formulated.

When peaks with similar populations across the three dihedrals are correlated with each other for both **II** and **III**, H-bonding patterns typical of the *cw* and *ccw* H-bonding networks emerge. The major *cw* and *ccw* conformations for the two models are superimposed upon the bond length, relative energy and dipole maps found in Figure 7.22. If the *exo*-anomeric effect is due to an $O2lp \rightarrow \sigma_{C1,O5}^*$ mechanism, then when present there should be a shortening of $r_{C1,O1}$ and a concomitant lengthening of $r_{C1,O5}$. From the bond length maps it can be seen that the formation of an $O2-H \rightarrow O1$ H-bond ($\xi \approx 300^\circ$) dramatically attenuates the C-O bond length correlates of the *exo*-anomeric effect, while the effect of an $O1-H \rightarrow O2$ H-bond ($\xi \approx 60^\circ$) is considerably more modest. The major population density peaks for **II** and **III** are both located in the energetic well on the ϕ vs. ξ energetic hypersurface ($\zeta = 300^\circ$; Figure 7.22C) with the central peak for **II** being approximately 1 kcal/mol lower in energy than **III**. The minor population density peaks on the ϕ vs. ξ energetic hypersurface ($\zeta = 180^\circ$; Figure 7.22G) for **II** and **III** both border the minimum energy regime. This area in the *in vacuo* DFT calculations, which extends from $\xi \approx 60^\circ$ to 90° with three roughly staggered minima in the ϕ dimension is

clearly defined by the O2-H→O3 H-bond characteristic of the *cw* H-bonding network. The peak for **II** is roughly 2 kcal/mol from the global energetic minimum (for $\zeta = 180^\circ$) on the *in vacuo* DFT energetic hypersurface, whereas the peak from **III** is ~4 kcal/mol from the lowest energy DFT structure. The slight disparity between the geometries predicted by these models and the range of the energetic minima from the hypersurface calculations is likely due to solvent mediated H-bonding effects on the energetic hypersurface in DMSO. Additionally the multiple sources of error intrinsic to the analysis certainly contribute to this minor discrepancy.

All the while maintaining the caveats inherent in the comparison between the various hypersurfaces generated on an *in vacuo* hemi-acetal model system with experimental results collected on a methyl glycoside in DMSO, model **II** seems to fit the chemically intuitive geometries offered by the theoretical bond length and energetic hypersurfaces slightly better than **III**, particularly with respect to the $\mu_{\phi_2} \approx 240^\circ$ value of the minor peak in **III**. How the *exo*-anomeric effect might work in such a geometry in which the aglycone is eclipsed with O5 may be rationalized based upon donation into $\sigma_{C1,O5}^*$ from the pseudo- σ lone pair orbital which is aligned with the C1-O5 antibonding orbital in that conformation. Alternatively, a combination of mitigating solvent effects or steric issues related to the aglycone may be stabilizing this geometry in spite of pressures from the *exo*-anomeric effect.

Indeed the energetic stabilization offered by the *exo*-anomeric effect as defined by a comparison between the relative energies of the three canonical staggered rotamers about ϕ seems to be almost negligible in contrast to the energy of an intramolecular H-bond between O2 and O1. For example, from a comparison of panels C and G in Figure

7.22 across the range $\xi = 180^\circ$ to 360° , the relative energy profiles are essentially superimposable with two minima at $\xi \approx 300^\circ$ and $\phi \approx 180^\circ$ and 300° . These two minima, which correspond to the conventionally described *exo*-anomeric orientations, certainly benefit from the energetic stabilization of an O2-H \rightarrow O1 H-bond. Even as it is tempting to attribute the energy difference about the three ϕ rotamers at $\xi \approx 300^\circ$ to the presence of the *exo*-anomeric effect, the difference in relative conformational energy is also dictated by the energetic differences between the putative H-bond and conformations which incur steric repulsions between the conflicting O1 and O2 hydroxyl protons ($\phi \approx 60^\circ$, $\xi \approx 300^\circ$). This is confirmed by a detailed inspection of the other half of the rotational itinerary about ξ for the two ζ values presented. From the $\zeta \approx 180^\circ$ energetic hypersurface (Figure 7.22G) there are three minima centered at $\xi \approx 60^\circ$ that roughly correspond to the three staggered rotamers about ϕ . These three minima have essentially the same relative conformational energies, despite their qualitatively different conformational characteristics; $\phi \approx 180^\circ$ and 300° are the presumptive ideal *exo*-anomeric orientations whereas $\phi \approx 60^\circ$ is an ideal orientation for an O1-H \rightarrow O2 H-bond when $\xi \approx 75^\circ$ (when $\zeta \approx 180^\circ$ and $\xi \approx 60^\circ$ the O2 hydroxyl is additionally able to form an O2-H \rightarrow O3 H-bond). The fact that the three ϕ rotamers have the same relative conformational energy indicates that any energetic stabilization offered by the *exo*-anomeric effect is comparable to that afforded by the O1-H \rightarrow O2 H-bond. When the similar ϕ and ξ conformational region is inspected in the $\zeta \approx 300^\circ$ energetic hypersurface (Figure 7.22C), the profile is dominated by the steric clashes between the O2 and O3 hydroxyl protons.

While it is impossible to estimate the strength of the O1-H \rightarrow O2 H-bond in isolation from either the *exo*-anomeric effect or the O1-H \rightarrow \leftarrow H-O2 steric repulsions, the

value of the *in vacuo* stabilization provided by either the O2-H→O3 or O2-H→O1 H-bonds can be gained by a comparison between the relative energies for $\xi \approx 60^\circ$ and 180° or $\xi \approx 300^\circ$ and 180° respectively when both ϕ and $\xi = 180^\circ$. This value is approximately 3.5 kcal/mol for either H-bond, however this number is likely to be attenuated in a solvent dependent manner. It has been shown that the either intra- or intermolecular donor/acceptor characteristics vary as a function of the hydroxyl identity about the pyranose ring with the O1 lone pairs exhibiting the greatest *s* character which translates into the anomeric hydroxyl being the weakest donor and strongest acceptor⁹². By logical inference any energetic stabilization provided by the *exo*-anomeric effect must be less than 3.5 kcal/mol, perhaps considerably so.

Another observation in support of the notion that the *exo*-anomeric effect is of limited energetic benefit with respect to other forces such as intramolecular H-bonding and various steric interactions including rotameric transition boundaries comes from a comparison of the bond length and relative energy hypersurfaces in Figure 7.22. If ϕ -rotational variation in C1-O and C1-O5 bond lengths particularly in the absence of an O2-H→O1 H-bond is diagnostic of the presence or absence of the *exo*-anomeric effect and the presence of the effect is a significant source of conformational stabilization, then the C1-O1 and the C1-O5 hypersurfaces should directly and inversely map onto the relative conformational energy hypersurface respectively. Although there is definitely some overlap, the correlation is far from perfect. A simple observation indicative of this is that in the set of $\xi = 180^\circ$ hypersurfaces (Figure 7.22E-G) for the $\xi \approx 60^\circ$ slices (a representative ξ conformation in which there is no O2-H→O1 H-bond) the bond length hypersurfaces are both bi-modal in the ϕ dimension reflecting the geometric aspects of

the $O2lp \rightarrow \sigma_{C1,O5}^*$ orbital interactions, whereas the energetic hypersurface is tri-modal in the ϕ dimension as a result of the rotameric transition boundaries about this dihedral.

Although it is difficult to absolutely assess the correlation effects in this system or unequivocally assign percentages to the correlated conformations about the three dihedrals, the relative percentages found in Table 7.11 for the various μ_ϕ , μ_ξ and μ_ζ values is suggestive of a high occurrence of the *cw* and *ccw* H-bonding network patterns for methyl- β -D-glucopyranoside in a DMSO solution. There have been a number of theoretical studies that have shown that the hydroxyls in saccharides participate in either bifurcated or 2-coordinate H-bonds which deviate from a tetrahedral arrangement about the oxygen^{53,92}. The hydrogen bond is a complex phenomenon that can be described either in terms of electrostatic interactions, charge transfer between the lone pair on the acceptor and a covalent anti-bonding orbital of the donor group, exchange repulsion or dispersion^{31,92-93}. There is evidence that most appropriate description for a particular H-bond is dependent upon the species involved^{93d}. The reduced polarity of the C-O bond versus the H-O bond in water or the highly polar S=O bond in DMSO is partially responsible for the reduced acceptor capacity of hydroxyls⁹². This leads to a reduced electrostatic attraction in the H-bond, a decreased local polarization between the oxygen nucleus and its lone pair acceptors, and thereby a reduced potential for orbital overlap between the lone pair and the anti-bonding orbital of the donor. This view of intra- and intermolecular H-bonds in saccharides that deviate from the idealized tetrahedral arrangement found in ice is in accord with some of the H-bonds formed from non-ideally staggered geometries suggested by the *cw* and *ccw* models in **II** and **III**.

The proposition of distinct *ccw* and *cw* H-bonding networks raises the question of rotameric transition paths for the conversion between the two H-bonding systems. Figure 7.23 depicts the population density profiles about the ϕ , ξ and ζ dihedrals as radial histograms. This is a convenient way to identify confluent populations about the three rotational itineraries in addition to regions that present significant boundaries. Because effects due to conformational correlation between ϕ , ξ and ζ may complicate the transitions between what appears to be overlapping populated regions about a particular dihedral in Figure 7.23, the conclusions about various transition paths were cross validated by the energetic hypersurfaces found in Figure 7.22.

It appears that there is a direct transition between the *ccw* and *cw* H-bonding patterns for both **II** and **III**. Other minor populations about ϕ , ξ and ζ appear to be deviations from the *ccw* or *cw* conformers and not transitions between them. The $\text{II}_{ccw} \rightarrow \text{II}_{cw}$ transition (the subscripts denote the *ccw* and *cw* conformers for the models presented in Figure 7.23) passes through the global dipole minimum, whereas $\text{III}_{ccw} \rightarrow \text{III}_{cw}$ skirts the region of maximal global dipole (Figure 7.22). This conclusion is primarily a result of the differing transition pathways indicated between **II** and **III** for ξ . The significant populations that characterize the boundaries between peaks about ϕ for both models indicate facile reorientation about this dihedral. The lack of any significant population in the upper hemisphere of the itinerary (the $\pm 90^\circ$ about $\phi = 0^\circ$) is in line with steric considerations regarding the aglycone.

There is a difference in the direction of the *ccw* to *cw* interconversion about ξ suggested by either **II** or **III**. The most direct pathway between II_{ccw} and II_{cw} passes through $\xi = 0^\circ$, whereas the III_{ccw} to III_{cw} inter-conversion most likely transits $\xi = 180^\circ$.

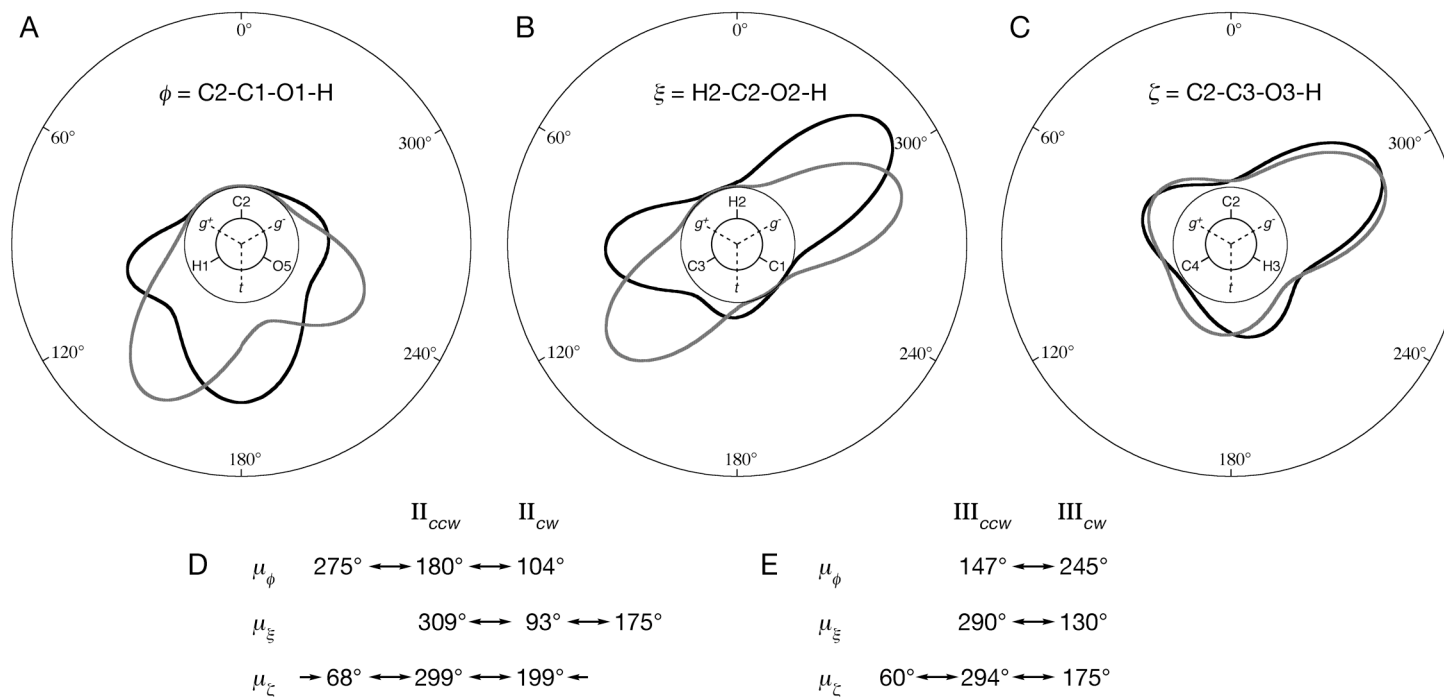


Figure 7.23. Radial Population Density Histograms about ϕ , ξ and ζ for **II** and **III** with Correlated Rotameric Transition Paths for the *ccw* and *cw* H-bonding Networks. Polar coordinates defined with arbitrary radial population units and a polar angle corresponding to the ϕ (A), ξ (B) or ζ (C) dihedrals measured in degrees. The annulus containing the nested Newman projections represents zero population. Black solid line represents model **II** and the grey line represents model **III**. Rotameric transition paths for the interconversion between *ccw* and *cw* conformers of **II** (D) and **III** (E) assume transitions proceed through adjacent rotameric regimes. Half arrows indicate transitional continuity about the entire rotational itinerary. The value for μ_ϕ used in the III_{ccw} model (E) is the weighted average of the μ_{ϕ_1} and μ_{ϕ_2} peaks from Table 7.11.

From an energetic map standpoint, the two pathways are of roughly equal energetic topology, however it is clear that due to the O2-H \rightarrow \leftarrow H-O3 steric clash the ζ *ccw* to *cw* shift must be concerted with any transition in ξ , if not precedent to it. The minimal population indicated in ξ boundary regions for both models implies higher transition energies than those for the other two dihedrals.

The transition between the *ccw* and *cw* conformers about ζ are able to pass directly from the *gauche*⁻ to the *trans* rotamers for both models, however the ζ profile for **II** suggests that there may be a partial contribution from a $g^- \rightarrow g^+ \rightarrow t$ mechanism. For the **II**_{*ccw*} \rightarrow **II**_{*cw*} transition, it is easy to envision a concerted mechanism in which the O3 hydroxyl is initially engaged in an O3-H \rightarrow O2 H-bond and makes the $g^- \rightarrow t$ transition about ζ , initially pulling the O2 lone pair with it, thereby placing a torque on the C2-O2 dihedral and facilitating the $g^- \rightarrow g^+$ transition about ξ . The transition is complete with the formation of an O2-H \rightarrow O3 H-bond characteristic of the *cw* H-bonding network. Transitions about ζ are likely to be highly dependent upon the orientation of the O4 hydroxyl. Recent experimental UV and IR evidence suggests that specific solvation and desolvation events involving O4 and O6 acts as a trigger controlling transitions between *ccw* and *cw* conformers⁶⁷. Additionally a high level QMMD of explicitly solvated β -D-glucopyranose concluded that vibrational coupling between a solvent molecule and the O6 hydroxyl led to rotameric reorientation in the hydroxymethyl molecular fragment⁹⁷. There have been other studies which have shown that solvent mediation of H-bonding involving O6 can affect the C-O rotameric distribution throughout the entire pyranose ring⁶³. The conformational models presented here do not conflict with these conclusions,

however a further analysis of the remaining conformational degrees of freedom in β -D-glucopyranose is required for a more complete picture of the solution behavior as derived from NMR spin-spin couplings.

Generalized Discussion of Error: Sources and Analysis — When reporting the results of a conformational analysis based upon the theoretical interpretation of experimental couplings, conducting a comprehensive error analysis and assignment of appropriate radial confidence intervals in terms of degrees is of practical importance, insofar as the direction of studies subsequent to a conformational analysis are in part subject to the accuracy of the original work. This applies to the experimental and theoretical methods employed, and has implications for both the accuracy and precision of the ultimate analysis. Contributions to this radial confidence interval arise from three sources of error and uncertainty; *a*) experimental, *b*) theoretical, *c*) parametrical *d*) and the inherent robustness of the model used in the analysis.

The first source of error is that intrinsic to the accuracy and precision of the experimental measurement. The reproducibility of spectral data is highly accurate given rigorous methodology and modern high field instrumentation. The precision is readily measured in Hz and is relatively uniform within any given spectra. Translating this error into a confidence interval measured in degrees is more difficult due to the topology of the coupling hypersurface. Since the second derivative of a trigonometric function returns another trigonometric function, the confidence interval in degrees for a given experimental value will vary non-linearly across the dimension of any given geometric parameter.

The second source of error arises from how well the DFT calculations accurately reproduce the coupling that arises from a specific geometry in solution. This source of error is harder to estimate than the previous, however there are a number of controllable factors that influence the magnitude of the error. These factors include functional⁷⁴ and basis set identity^{74b,75}, use of solvent continuum models¹⁶, and inclusion of the non-Fermi contact Ramsey^{75c} and ρ -vibrational¹² coupling contributions. Benchmarks for the accuracy of these calculations are based upon comparisons to conformationally restricted experimental systems. The B3LYP hybrid functional has been shown to perform well in this manner as a balance of calculational accuracy and computational efficiency. The 6-31G* basis used for geometric optimizations was designed to capture an accurate representation of valence energetics with a maximum of computational efficiency by using an inflexible representation of the 1s cores on carbon and oxygen. In contrast, the relatively flexible [5s2p1d|3s1p] extended double-zeta basis set used in these studies for spin-spin coupling calculation is specifically designed to reliably recover a balance of core and valence contributions to the Fermi contact term. This basis has returned *in vacuo* results estimated to deviate by less than 5% of the total coupling when compared to experimental values^{6b,76}

A charge present on an ionizable group can significantly affect the magnitude of proximal couplings. In uncharged systems, the addition of solvent continuum has a minimal impact upon the magnitude of the spin-spin coupling returned by DFT (*vide supra*, §B), however the presence of a solvent continuum helps to locally ‘contain’ the effects of the charge on proximal couplings⁷⁷. For example, a comparison of the corresponding *in vacuo* and *in solvo* data for **2C** reveals a rms <0.2 Hz or <3.5% of the

total average coupling across the ensemble. In contrast the Series 7 calculations on O2 ionized forms of **2**^C demonstrate a rms of ~3.5 Hz or 35% of the ensemble averaged coupling when comparing *in vacuo* and *in solvo* data, with the *in solvo* values more closely representing experimental values (*vide supra*, §B).

The spin-dipole (SD), paramagnetic spin orbital (PSO), and diamagnetic spin orbital (DSO) portions of the total coupling tend to be very small in *sp*³ hybridized organic compounds compared to the Fermi contact term. In the case of the Series 1 calculations on **2**^C, the three terms were additive in their attenuation of the Fermi contact contribution. There was a rms of <0.02 Hz in the ensembled sum of the non-Fermi contact terms which represented <0.5% of the ensemble averaged coupling.

The underlying theme of this work is that NMR properties are dependant upon variations in molecular geometry. This is most often considered for gross structural changes such as dihedral rotameric regime shift, however smaller temperature and isotope dependant fluctuations in the molecular framework affect nuclear shielding and indirect spin-spin coupling as well, in some instances by more than 10%.¹² This implies that the converged energetically optimized structure returned by the *ab initio* calculation that is used for the spin-spin coupling calculation at the particular conformational coordinate of interest represents only the most likely local structure and does not necessarily reflect other slightly higher energy geometries that the molecule samples in solution for the same conformational coordinate of interest. There are numerous methods for calculating these so called p-vibrational contributions to NMR couplings such as perturbation methods,^{12a, 13} numerical solutions to the vibrational differential equation,¹⁴ or quantum Monte-Carlo (QMC) simulations on the molecular framework¹⁵ to name a

few. All methods have advantages and disadvantages, however a drawback common to all is the prohibitive computational cost of such calculations on systems with more than a half dozen heavy atoms such as carbon or oxygen. Furthermore, the ρ -vibrational coupling contribution is likely to vary synchronously with gross conformational changes, making estimation of this factor unfeasible in a conformationally quantitative system such as this with 1,728 grossly distinct geometries. Bock and Duus estimated the ρ -vibrational contributions about the staggered rotamers in a conformational analysis of the hydroxymethyl group in a spectrum of hexopyranoses and found the effect to be essentially negligible^{87a}, a result subsequently verified by others^{87b}. Furthermore, the ρ -vibrational coupling contributions represent <5% of the unscaled $^1J_{CC}$ and $^2J_{CH}$ and <2% of the unscaled $^1J_{CH}$ at 300 K in ethane. These effects are not always so nominal as evinced by the ~30% zero point vibrational contribution to the total $^2J_{CC}$ in benzene.^{12a}

The third potential source of error in the structural interpretation of experimental spin-spin couplings arises from how well the DFT data is fit. The two factors that control this are *a*) the form of the equation used to fit the data, and *b*) the optimization of the coefficients for each term in the equation. Due to specifics of the fitting procedure it is much easier to estimate the error associated with the goodness of fit in terms of Hz. Practically however the use of such a parameterization seeks to interpret the coupling in terms of a set of associated geometric constraints measured in degrees. There are overlapping nonlinearities to consider when transforming the error function from Hz to radial confidence intervals in the parameterized conformational space. The first nonlinearity to consider is that of the error measured in Hz as a function of conformation. That is to say that the difference between the DFT coupling value for a given geometry

and the value determined by back calculation at the same conformational coordinates from the parameterization varies non uniformly throughout the conformational space as determined by the quality of the data fit. In Figure 7.18 for example, this is represented by deviations between the DFT data and the parameterized hypersurface along the z -axis. The second nonlinearity is the response of the ‘goodness of fit’ portion of the radial confidence interval for a constant error measured in Hz across the conformational space. This response is also dependent upon the coupling hypersurface topology and is measured in Figure 7.18 by differences between the DFT data and the parameterized hypersurface in the xy -plane.

An additional dimension of complexity arises from the interdependence of the radial confidence intervals for each of the geometric parameters, in this case ϕ , ξ and ζ . For example, solving the parameterized ${}^2J_{C1,C3}$ equation in Table 7.9 for ζ and using the DFT coupling value produced at a particular ϕ , ξ and ζ conformational coordinate yields two values predicted by the parameterizations designated $\zeta_{\pm\text{error}}$. There are two values due to the symmetric redundancy of the unimodal ζ coupling profile (*i.e.* for a given set of ϕ and ξ values, $\zeta = 30$ and 330 return similar coupling results from the parameterized equation). This quantity $\zeta_{\pm\text{error}}$ measured in degrees is a function of the original DFT ϕ , ξ and ζ values and represents the maximally erroneous ζ value possibly returned from the parameterized coupling equation in the interpretation of an experimental coupling. Because the exact DFT values for ϕ and ξ were used to determine $\zeta_{\pm\text{error}}$, this term implies an absolute accuracy in the ϕ and ξ values also returned by the parameterized equation in the interpretation of the experimental coupling. This is equivalent to a

linearization of the error into the ζ dimension. If the coupling hypersurface happens to pass directly through the DFT point with the same conformational coordinate (*i.e.* the DFT and the parameterization have the same coupling value at a particular conformational coordinate corresponding to an exact fit at that point) then $\zeta_{\pm\text{error}}$ has the same value as the corresponding ζ value for a particular set of ϕ and ξ values from the DFT data. A plot of the difference between the DFT derived coupling and the parameterized coupling value(s) back calculated at the DFT conformational coordinate ($\Delta_{\text{DFT-Fit}}^2J_{\text{C1,C3}}$) versus $\zeta_{\pm\text{error}}$ (Figure 7.24B) gives a picture of the correlation between the goodness of fit error in ζ measured in Hz versus degrees. Mapped onto Figure 7.18, this is ostensibly a measure of error in the z -dimension versus the xy -plane as a function of the dihedral ζ . The $\Delta_{\text{DFT-Fit}}^2J_{\text{C1,C3}}$ versus $\zeta_{\pm\text{error}}$ profile for a specific ζ slice shows a simple sinusoidal relationship with variation in the ϕ and ξ values. This is due to the simplicity of the term describing the behavior of ζ in the parameterized $^2J_{\text{C1,C3}}$ equation in Table 7.9, $E \cos(\zeta)$. This is in contrast to the analogous unimodal plots of $\Delta_{\text{DFT-Fit}}^2J_{\text{C1,C3}}$ versus $\phi_{\pm\text{error}}$ and bimodal plots of $\Delta_{\text{DFT-Fit}}^2J_{\text{C1,C3}}$ versus $\xi_{\pm\text{error}}$ in panels C and D of Figure 7.24. The scatter in these latter two profiles is due to the multiple appearances of both ϕ and ξ in the parameterization, particularly in the mixed angle term $F \cos(\xi)\sin(\phi)$, and is systematic in the corresponding additional conformational dimension. In theory this graphical method could be used to assign individual confidence intervals measured in degrees to the values found in Table 7.11.

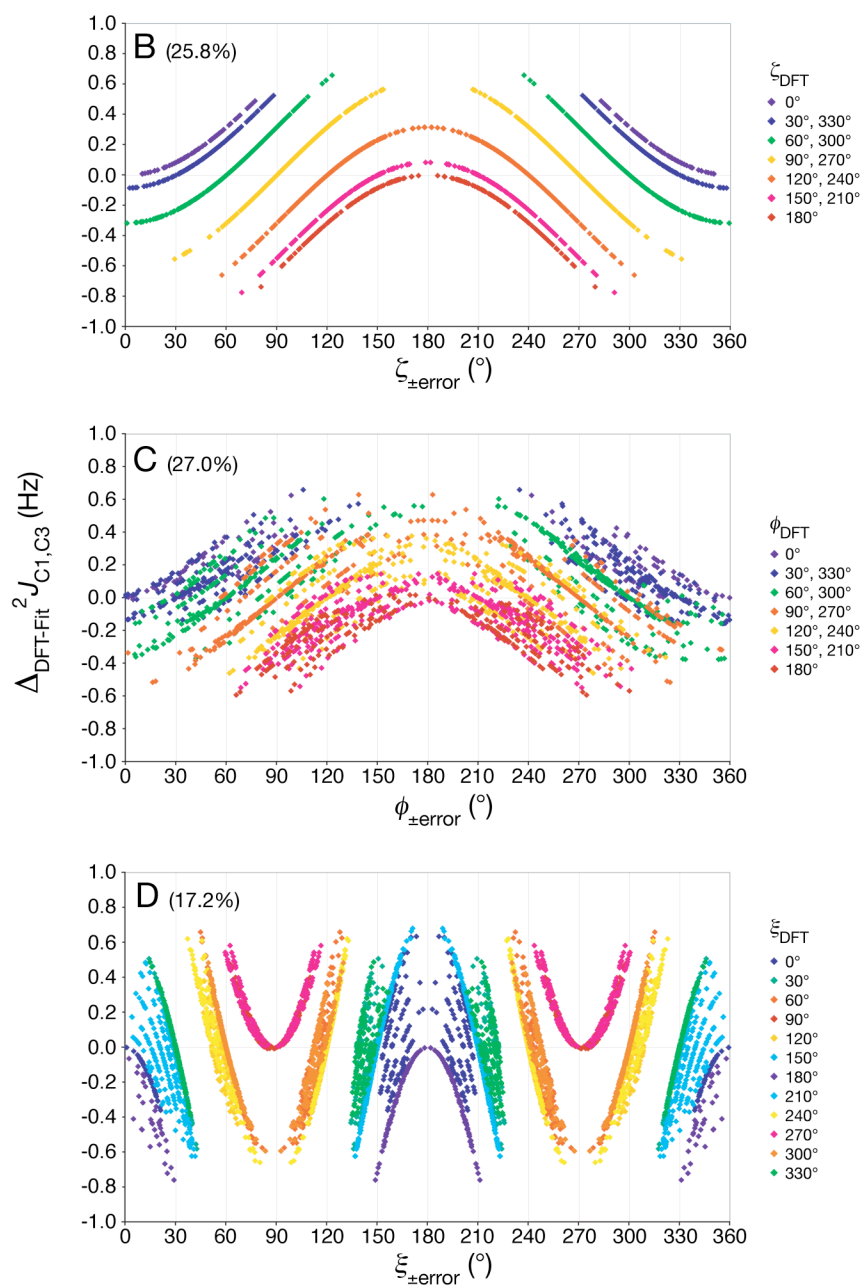
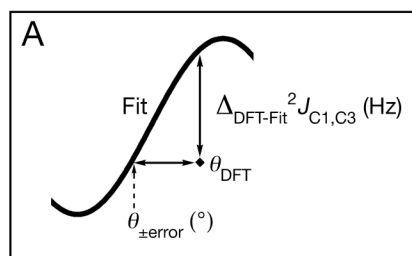


Figure 7.24. The Goodness of Fit Error for ${}^2J_{C1,C3}$ in 9C as a Function of Dihedral. The definitions for the x and y values in panels B-D are found in panel A, where θ represents one of the dihedrals ζ , ϕ or ξ respectively. Data in panels B-D are color coded by virtue of the respective ζ_{DFT} , ϕ_{DFT} or ξ_{DFT} angles used to calculate $\Delta_{DFT-Fit}{}^2J_{C1,C3}$. The percentage of the 1728 DFT data points whose coupling magnitude was outside the range of the parameterization thereby precluding the calculation of $\theta_{\pm error}$ is indicated in parenthesis beside the panel label. Definitions for ζ , ϕ and ξ are found in the text.

There are however a number of drawbacks to the use of this graphical error transformation. One drawback discussed below is that the usefulness is limited to the resolution of the DFT data sampling in the conformational space. Secondly the error functions found in Figure 7.24 are only useful for a conformational model that has equal weighting throughout. The practical application would require calculation of similar relationships using the weighting functions from the experimental coupling analysis. Additionally there is the problem of DFT (or experimental) coupling values that fall entirely outside the range of coupling in the parameterizations. For example when considering the $\zeta_{\pm error}$ profile, 25.8% of the couplings from the 1,728 DFT structures would return no ζ from the use of Equation 7.1 for the given ϕ and ξ values of that DFT structure. Analogous percentages for $\phi_{\pm error}$ and $\xi_{\pm error}$ profiles are 27.0% and 17.2% respectively. This indicates that the parameterization for the dihedral ξ upon which ${}^2J_{C1,C3}$ is primarily conformationally dependent is better, or at least more inclusive, than that for the other two dihedrals of secondary dependence ϕ and ζ . These large percentages can be misleading however, because the percentage of possible couplings outside the interpretive power of the parameterized coupling equation would be significantly reduced if there was an integration of the $\zeta_{\pm error}$, $\phi_{\pm error}$ and $\xi_{\pm error}$

functions into a single correlated error function described by the interdependent error ranges for all three dihedrals centered around a conformational coordinate. A single correlated error function could be derived based upon the difference between a spline multivariate trigonometric interpolation of the DFT data and the original parameterization, however calculation of such an interpolation should logically supplant the continuous parameter function as the interpretive tool for experimental couplings. Unless a variant of the Whittaker–Shannon interpolation formula is adapted to the purpose, the spline interpolation method has the disadvantage that there is the potential for n number of discrete interpolant intervals, n being the number of data points fit.⁴¹ Alternatively a numeric approach to the ‘goodness of fit’ error problem could be used, however the precision of this method would be limited by the coarseness of DFT sampling in the conformational space.

This ‘goodness of fit’ source of error can be avoided entirely by using the DFT data directly to interpret the experimental coupling, however the limit of precision in the interpretation is dictated by the resolution of the DFT structures sampled for any of the pertinent structural parameters, in this instance using the simplified case of $\sigma = 0$, there would be $\pm 15^\circ$ range on the precision of $\{\mu_\phi, \mu_\xi, \mu_\zeta\}$. This problem could be potentially avoided by application of one of the previously mentioned interpolation protocols.

The final source of error arises from the assumptions intrinsic to the model to which the experimental data is being fit. Obvious sources of error in this regard arise from choices about which parameters regulating the shape of the Gaussian distribution to optimize and which to hold constant as discussed above.

The applicability of a Gaussian distribution to the modeling of any quantitative phenomena is based upon the central limit theorem, which states that the sum of an independent and identically-distributed (*i.i.d.*) random sampling with a finite variance approaches normality as the sampling size reaches infinity⁸⁵. The fundamental assumption of *i.i.d.*, while providing a good intuitive approximation of a single state torsional reorientation in terms of a single bell curve, does not account for changes to the probability distribution within the range of the variance (2π) due to stereo-electronic or stereo-chemical effects. These forces can complicate the true conformational profile with multiple overlapping bellcurves. For example, when the hypothetical dihedral I is in the rotameric regime *a* the average value is dependent upon the conformation of the neighboring dihedral II that alternates between two of its own rotameric regimes, *b* and *c*, the first of which enables a hydrogen-bond between the atomic constituents of I and II and second which does not. The average value of *Ia* in the situation *Ia•IIb* is likely to be slightly different than *Ia•IIc*. A complicated superposition of multiple correlation effects is not well modeled by a single normal distribution. The number and features of such overlapping bell curves arising from correlation effects is difficult to predict *a priori*.

Since the the optimization protocol integrates the product of a periodic function and an aperiodic scalar weighting function over a discrete range, error may also arise from potential discontinuities at the boundaries of the integral range. For boundary discontinuity in a model to be a problem there must be a significant population from the tail of the scalar weighting function that gets cut off. Formally the integral should be conducted from ∞ to $-\infty$, however the problem is minimized by integrating over a range suitably larger than 2π such that there is a negligible remaining contribution from the

weighting function. In the current model the range of the integration was from $-\pi/3$ to $7\pi/3$, with overlap about the rotational axes from $-\pi/3$ to $\pi/3$. This region was selected for the overlap because it straddles an eclipsed geometry that was expected to be minimally populated.

A final consideration is the intrinsic weaknesses and strengths of the trajectory optimization method used in this study compared to a hypersurface grid search approach^{37a}. The strength of a trajectory optimization is its relatively small computational cost; this study was conducted on a single Dual-Core Intel processor. The problem with this method is that it is inexhaustive. It is difficult to absolutely conclude that the model that is the global best fit to the experimental data has been found. This global best fit may lie on the other side of a boundary from the initial conditions that precludes arriving at this state using even fairly liberal aleatoric optimization algorithms. The only way to curtail this potential difficulty is by using a variety of reasonable initial conditions. In contrast, a grid search across the $\{\mu_{\theta_n}, \sigma_{\theta_n}, w_{\theta_n}\}$ conformational hypersurface will provide all points on the hypersurface within a user defined resolution, which can then be sorted by χ^2 to obtain the best global conformational model. The primary drawback of this technique is the computational cost involved. This study, which reported results with a 1° and 1% precision for 27 parameters across a maximum of 15 experimental constraints, would require $\sim 6.6 \times 10^8$ separate integrations from $-\pi/3$ to $7\pi/3$ given the defined ranges in the parameters. Given that a single integration took ~ 45 seconds for a single 2.16 GHz Dual-Core processor, it can be appreciated how long this might take. One factor that can alleviate this bottleneck is the high degree of parallelizability possible for a grid search algorithm. A practical compromise would be to initially use a

parallelized grid search method with a fairly coarse resolution in the parameter set, followed by a trajectory optimization refinement of several of the highest ranked models.

7.6. Conclusions

There is a wealth of theoretical studies attempting to interpret NMR spin-spin coupling mechanisms by invoking any number of methods such as the bond vector formalism⁸¹, the Sum Over States technique⁸², or finite perturbation methods⁸³ such as Natural *J*-coupling Analysis^{24,29}. Typical objectives of these studies include identifying the physical determinants of coupling sign and magnitude and understanding to what extent the coupling is mediated through steric versus hyperconjugative mechanisms³⁴. However, such studies are often conducted on model systems that are conveniently simplified, have minimal biological relevance or are otherwise amenable to the theoretical principles employed^{29c,43,84}. On the other hand, the development of sophisticated protocols for the strategic isotopic enrichment of biologically relevant molecules enables the continued systematic compilation of a battery of scalar coupling data³⁻⁴. As a result, there is a burgeoning literature of empirical correlations between measurable NMR parameters and molecular geometry in proteins³⁵, nucleic acids³⁶ and carbohydrates³⁷. These efforts have been facilitated in some measure by the advent of accurate high-level DFT calculation of *J*-coupling values in moderately sized molecules. In instances where the interpretation of experimental data in simplistic structural terms such as individual bond length is not transparent, an appeal to a unifying and underlying quantum mechanical description can become necessary; that is to say that a small number of geometric constraints and a scalar coupling sensitive in part to those constraints may

have different susceptibilities to the fundamental MO description. This report attempts in part to bridge the gap between a first principles description of localized electron density and solid phenomenologic correlations between empirical NMR spin-spin couplings and theoretically determined molecular geometry in saccharides.

While it is true that NBO analysis is in some sense an arbitrary, if clearly delineated recapitulation of the MO description, it is useful in terms of tracking and localizing subtle shifts in electron density that traditionally represent the diffuse tail of a standard molecular orbital. These discreet localizations can in turn be cross-correlated with a panoply of inter-related geometric constraints such as bond length, angle or dihedral, as well as spectroscopic data to gain an intuitive understanding of the connection between molecular structure and J -coupling mechanisms. An example of a study that interprets empirical structure-coupling relationships in a biologically relevant system through a first principles approach was described by Munzarová and Sklenář in 2002⁴⁴. The work resolves a trans-glycoside $^3J_{C6/8,H1'}$ Karplus profile discrepancy between purine and pyrimidine nucleotides in terms of a differential through space contribution to the coupling from the highest occupied molecular orbital.

The first post-Hartree–Fock calculation of geminal spin-spin coupling constants was conducted in 1967 by Armour and Stone⁴³. In this study they found that electron correlation effects are crucial to an accurate calculation of $^2J_{HCH}$ in ethylene, the most important of which to consider involved the bonding and anti-bonding C-H orbitals. They showed that within the molecular plane the nodal lines for the bonding and anti-bonding C-H orbitals are virtually coincident in the proximity of the second proton, implying that even small changes in the location of the nodes can significantly alter the $\sigma \rightarrow \sigma^*$

delocalizations and thereby the magnitude and sign of the contribution from these orbitals to the total 2J .

We have shown that subtle configuration and conformation dependent shifts in $\sigma_{CC} \rightarrow \sigma_{CO}^*$ and O2 $lp\sigma \rightarrow \sigma_{CC}^*$ interactions respectively lead to modulation of $^2J_{CC}$, thus confirming the importance of the relative disposition of the σ and σ^* orbitals for the bonds along the nuclear coupling pathway. In addition, the geometric modulation of steric interactions involving coupling path σ_{CC} plays an equal part in determination of the sign and magnitude of $^2J_{CC}$.

The predictive ability of the projection resultant method¹¹ to calculate geminal coupling sign and magnitude is a result of the underlying modal characteristics of the specific orbital interactions that modulate the spin-spin coupling mechanisms. These orbital interactions vary as a function of C-C rotation, however within the context of the conformationally rigid pyranose ring, variation in the orbital interaction as a function of C-C rotamer manifests in terms of hydroxyl configuration, specifically configuration at the coupled nuclei. The relative disposition of the axial versus equatorial hydroxyls at either of the coupled sites are non-equivalent, and therefore represent different C-C rotamers; i.e. an equatorial hydroxyl at a coupled carbon is in plane with respect to the coupling pathway leading a specific $\sigma_{CC} \rightarrow \sigma_{CO}^*$ orbital interaction, whereas an axial hydroxyl is at $\pm 120^\circ$ with respect to the coupling pathway (sign depending on which coupled carbon is considered) leading to a distinctly different σ_{CO}^* orbital disposition and consequently different $\sigma_{CC} \rightarrow \sigma_{CO}^*$ orbital interaction. The lack of O2 configurational

dependence in ${}^2J_{C1,C3}$ can be understood by recognizing the dispositional redundancy of the O2 axial and equatorial epimers with respect to the coupling path.

The heightened conformational influence of hydroxyls on the central atom of a three-bond path is qualitatively validated by the *PR* method. In the calculation of a projection resultant, any hydroxyl on the central atom is counted in two projections: in the case of ${}^2J_{C1,C3}$, the first down the C1-C2 bond and the second down the C2-C3 bond. Hydroxyls on the terminal (coupled) nuclei are counted only once in the projection. It intuitively follows that the conformation of the O2 hydroxyl will have a greater effect on the transfer of spin-density from C1 to C3 than the terminal hydroxyls. This ability of O2 to simultaneously affect σ_{C1C2}^* and σ_{C2C3}^* in equal measure due to proximity versus the primary interaction of O1 with σ_{C1C2}^* and O3 with σ_{C2C3}^* is at the root of the bimodal versus unimodal dependence of ${}^2J_{C1,C3}$ on O2 versus O1 and O3 respectively. An analogous proximity effect is likely at the root of the strong bimodal dependence of ${}^2J_{CCH}$ on hydroxyls appended to the central carbon and weaker unimodal dependence on hydroxyls attached to the coupled carbon reported previously³⁸.

The understanding of a variety of spin-spin coupling phenomena in saccharides could benefit from the NJC deconvolution of spin-spin coupling into distinct coupling mechanisms. For example, the subtleties of the electronegative substituent effect on the shifting of Karplus profiles observed in glycosidic torsions, empirically described by the Haasnoot formulation⁷³ of the generalized Karplus equation^{2a-b}, could be understood in terms of orbital interactions by this approach. Other examples include the vicinal lone pair effect on ${}^1J_{CH}$ as a potential $\mathcal{J}^{(deloc)}$ spin-transfer mechanism, or the 1,3-diaxial effect

on $^1J_{\text{CH}}$ as mediated by a steric $J(\text{L})$ type process. Furthermore, based upon a theoretical interpretation of a large battery of experimental scalar couplings, there is evidence to suggest that perhaps the NBO description of pseudo- σ and pseudo- π lone pair orbitals may be useful in the understanding of *exo*-anomeric C1-O1 torsional preferences.

Befitting its rank as the most common bio-monomer on the planet, there has been a wealth of theoretical data collected for glucopyranose such as a variety of force-field applications⁴⁵⁻⁴⁸, quantum mechanical calculations⁴⁹⁻⁵⁵ and calculation of crystalline structures^{49a,56-58}. Given the volume of theoretical data generated, there are a significant number of conflicting results, and even recent studies conducted at a relatively sophisticated level of theory have produced chemically counter-intuitive results⁶⁶. For example, a 3 picosecond QMMD study⁵⁹ on either α - or β -glucopyranose using the Car-Parrinello method⁶⁰⁻⁶¹ with 58 explicit water molecules showed that, within their criteria for identifying a hydrogen bond, α -glucopyranose had an average of 0.2 intramolecular H-bonds during the course of the trajectory and β -glucopyranose had ostensibly none. The average solvent H-bond number ranking amongst the hydroxyls in β -glucopyranose from greatest to least was O4, O3, O2, O6 and O1 with O4 – O2 each being at or slightly above two H-bonds, indicating that hydroxyl C-O rotamers are heavily dependent upon the aqueous solvation sphere. These results are in contrast to a recently published 200 ns simulation on β -glucopyranose using the carbohydrate optimized GROMOS 45A4 force field⁶⁶. In this work the authors use an extensive literature survey to compare against their report of a 15%, 39%, 59% and 35% occurrence of an intra-molecular hydrogen bond (be it donor or acceptor) in the solvated molecule for O1 through O4 respectively. This work is notable for its thorough treatment of correlated hydroxyl conformation throughout the

molecule, although it suffers from a number of anomalous force-field dependent results related to the preferred *exo*-anomeric hydroxyl orientation and to ring form inter-conversions.

It has been shown that variations in experimental spin-spin couplings sensitive to C-O rotamer orientation can indicate the presence of intra-molecular H-bonding in saccharides^{32a}. Woods *et al.*⁶³ recently showed through a combined experimental ¹H-NMR and MD approach that a solvent dependent balance between O6 intra- and inter-molecular H-bonding can affect exocyclic torsions around the ring in 2,3-di-*O*-methyl- α -D-glucopyranoside, a molecule whose capacity for intramolecular H-bonds involving O2 and O3 is hampered by the presence of the methyl-ethers. The paper concluded that in CH₂Cl₂ the lack of solvent hydrogen bond partners led to a strengthening of intramolecular H-bonds and a minimization of repulsive oxygen–oxygen interactions, whereas water disrupts stabilizing intramolecular H-bonding, allowing internal electrostatic repulsions to dictate exocyclic torsional properties.

Dix *et al.*⁶⁸ established a linear relationship between the empirical expression $E_T(30) + 100\beta_{KT}$, and a ΔG_i for the formation of an intramolecular H-bond in the hydroxyethers 10 α -(hydroxymethyl)-2-oxabicyclo[4.4.0.]decane and 4-*tert*-butyl-2-(hydroxymethyl)-1-methoxycyclohexane across a wide spectrum of solvents including CD₂Cl₂, D₂O and DMSO based on experimental ³J_{HH} and IR measurements. The term β_{KT} refers to the (Kamlet-Taft) empirical solvent hydrogen bond acceptor scale and is a measure of H-bond acceptor ability based on the basicity of the acceptor (the 100 pre-factor is simply an order of magnitude scaling factor), whereas $E_T(30)$ is an empirical

measure of solvent polarity. The study assigned approximate ΔG_f values of 0.5, -0.5, and -2 in DMSO, D₂O and CD₂Cl₂ respectively for the formation of a hydrogen bond between the hydroxymethyl hydroxyl and the methoxy oxygen in the second compound, implying an increasing potential for intramolecular H-bonding acceptance along the same continuum. The same relationships did not hold for α_{KT} , the (Kamlet-Taft) empirical solvent hydrogen bond donor ability, potentially due to the fact that the putative H-bond that they were observing was uni-directional, meaning that the methyl ether can act only as an acceptor.

By analogy to these results, a mental image begins to appear in which the highly polar DMSO acts as a slightly better H-bond acceptor than either water or the hydroxyl groups of the saccharide while partially mitigating the repulsive oxygen–oxygen interactions in the sugar. However, due to the aprotic nature of DMSO all donor protons derive from solute hydroxyl groups and there is essentially no competition for lone pair density on the solute. Thus, while the DMSO is a slightly more attractive H-bond acceptor than any particular oxygen in the saccharide, a solute-solvent H-bond in this context is invariably uni-directional, whereas the majority of possible intramolecular H-bonds have a bi-directional potential, meaning that either hydroxyl can act as a donor or acceptor^{49d}. Thought of another way, there are no solvent H-bond donors that are in competition with a hydroxyl donor for electron density on a proximal acceptor in the same molecule.

A recent gas phase mass-selected infrared laser spectroscopy study conducted by Simons and coworkers⁶⁷ on selectively hydrated phenyl-*O*- β -D-glucopyranoside elegantly identified three distinct un-hydrated conformers characterized by the “counter clockwise”

cooperative H-bond network (O4→O3→O2→O1) with O6 donating to either O4 or O5. Upon partial hydration, a single water molecule entered the H-bonding network by either *a*) mediating the O6→O5 interaction such that O6→O_{water}→O5 meanwhile maintaining the counter clockwise O4→O3→O2→O1 chain or *b*) instigating a clockwise H-bond network (O2→O3→O4→O_{water}→O6→O5).

We have shown that methyl-*O*-β-D-glucopyranoside exhibits C-O rotamers for C1 through C3 consistent with the *ccw* and *cw* H-bonding networks in DMSO at room temperature. These results are based on the analysis of a large ensemble of NMR spin-spin couplings that include both ¹³C and ¹H couplings to the hydroxyl protons of C2 and C3. This study sets the groundwork for the practical extension to other parts of the pyranose ring, both as a monosaccharide and in larger molecular contexts which may effect the hydroxyl conformations.

7.7. References

1. (a) Brooks, A.; Dwek M.V.; Schumacher, U. *Functional and molecular Glycobiology* **2002**, BIOS Scientific Publishers, Ltd. (b) Maddox, J. *Nature* **1993**, *364*, 669. (c) Helenius, A.; Aebi, M. *Annu. Rev. Biochem.* **2004**, *73*, 1019
2. (a) Karplus, M. *J. Am. Chem. Soc.* **1963**, *85*, 2870 (b) Karplus, M. *J. Am. Chem. Soc.* **1969**, *91*, 1-10 (c) Kraszni, M.; Szakács, Z.; Noszál, B. *Anal. Bioanal. Chem.* **2004**, *378*, 1449-1463
3. (a) Wu, J.; Bondo, P.B.; Vuorinen, T.; Serianni, A.S. *J. Am. Chem. Soc.* **1992**, *114*, 3499 (b) King-Morris, M.J.; Serianni, A.S. *J. Am. Chem. Soc.* **1987**, *109*, 3501
4. (a) Danilova, V.A.; Krivdin, L.B. *Russ. J. Org. Chem.*, **2003**, *39*, 663 (b) Danilova, V.A.; Isotomina, N.V.; Krivdin, L.B. *Russ. J. Org. Chem.*, **2004**, *40*, 1194
5. (a) Malkin, V.G.; Malkina, O.L.; Casida, M.E.; Salahub, D.R. *J. Am. Chem. Soc.* **1994**, *116*, 5898 (b) Malkin, V.G.; Malkina, O.L.; Eriksson, L.A.; Salahub, D.R. Modern Density Functional Theory: A Tool for Chemistry. In *Theoretical and*

- Computational Chemistry*, vol. 2 (eds. Seminario, J.M.; Politzer, P.) **1995**, Elsevier, Amsterdam (c) Malkina, O.L.; Salahub, D.R.; Malkin, V.G. *J. Chem. Phys.* **1996**, *105*, 8793 (d) Malkin, V.G.; Malkina, O.L.; Steinbrunner, G.; Huber, H. *Chem. Eur. J.* **1996**, *2*, 452
6. (a) Bose, B.; Zhao, S.; Stenutz, R.; Cloran, F.; Bondo, P.B.; Bondo, G.; Hertz, B.; Carmichael, I.; Serianni, A.S. *J. Am. Chem. Soc.* **1998**, *120*, 11158 (b) Cloran, F.; Carmichael, I.; Serianni, A.S. *J. Phys. Chem. A* **1999**, *103*, 3783 (c) Church, T.J.; Carmichael, I.; Serianni, A.S. *J. Am. Chem. Soc.* **1997**, *119*, 8946
 7. (a) Grant, D.M.; Facelli, J.C.; Alderman, D.W.; Sherwood, M.H. *Nuclear Magnetic Shieldings and Molecular Structure* (ed. Tossell, J.A.) **1993**, p.367, Kulwer Academic, Dordrecht (b) Barfield, M. *Nuclear Magnetic Shieldings and Molecular Structure* (ed. Tossell, J.A.) **1993**, p.523, Kulwer Academic, Dordrecht (c) Liu, F.; Phung, C.; Alderman, D.W.; Grant, D.M. *J. Mag. Reson. Ser. A* **1996**, *120*, 231
 9. Bracewell, R. Pentagram Notation for Cross Correlation. In *The Fourier Transform and Its Applications*, **1999**, pp. 46 and 243, 3rd ed., McGraw-Hill, New York
 10. (a) new JCC ref in press (b) King-Morris, M.J.; Serianni, A.S. *J. Am. Chem. Soc.* **1987**, *109*, 3501-3508.
 11. Church, T.; Carmichael, I.; Serianni, A.S. *Carbohydrate Research* **1996**, *280*, 177 – 186
 12. (a) Ruden, T.A.; Ruud, K. Ro-Vibrational Corrections to NMR Parameters. In *Calculation of NMR and EPR Parameters*, (eds. Kaupp, M; Bühl, M.; Malkin, V.G.) **2004**, 153 – 173. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim (b) Ruden, T.A.; Lutnæs, O.B.; Helgaker, T.; et al. *J. Chem. Phys.* **2003**, *118*, 9572
 13. (a) Ruud, K.; Åstrand, P.-O.; Taylor, P.R. *Int. J. Comput. Methods. Sci. Eng.* **2003**, *3*, 7 (b) Kern, C.W.; Matcha, R.L. *J. Chem. Phys.* **1968**, *49*, 2081 (c) Åstrand, P.-O.; Ruud, K.; Taylor, P.R. *J. Chem. Phys.* **2000**, *112*, 2655
 14. (a) Clifford, C.E.; Malik, D.J. *J. Chem. Phys.* **1987**, *87*, 2806 (b) Sundholm, D.; Gauss, J.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *243*, 264
 15. Böhm, M.C.; Schulte, J. Ramírez, R. *Int. J. Quantum Chem.* **2002**, *86*, 280
 16. M. T. Cancès, B. Mennucci, and J. Tomasi, *J. Chem. Phys.* **1997**, *107*, 3032.
 17. (a) B. Mennucci and J. Tomasi, *J. Chem. Phys.* **1997**, *106*, 5151. (b) B. Mennucci, E. Cancès, and J. Tomasi, *J. Phys. Chem. B* **1997**, *101*, 10506. (c) J. Tomasi, B. Mennucci, and E. Cancès, *J. Mol. Struct. (Theochem)* **1999**, *464*, 211. (d) R. Cammi, B. Mennucci, and J. Tomasi, *J. Phys. Chem. A* **1999**, *103*, 9100. (e) M.

- Cossi, V. Barone, and M. A. Robb, *J. Chem. Phys.* **1999**, *111*, 5295. (f) R. Cammi, B. Mennucci, and J. Tomasi, *J. Phys. Chem. A* **2000**, *104*, 5631.
18. Stenutz, R.; Carmichael, I.; Widmalm, G.; Serianni, A.S. *J. Org. Chem.* **2002**, *67*, 949 – 958.
 19. (a) Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, **2004**. (b) GaussView, Version 4.12, Dennington II, R.; Keith, T.; and Millam, J.; Semichem, Inc., Shawnee Mission, KS, **2007**.
 20. Becke, A.D. *J. Chem. Phys.* **1993**, *98*, 5648 – 5652.
 21. Hehre, W.J.; Ditchfield, R.; Pople, J.A. *J. Chem. Phys.* **1972**, *56*, 2257 – 2261.
 22. Cloran, F.; Zhu, Y.; Osborn, J.; Carmichael, I.; Serianni, A.S. *J. Am. Chem. Soc.* **2000**, *122*, 6435 – 6448.
 23. Cloran, F.; Carmichael, I.; Serianni, A.S. *J. Am. Chem. Soc.* **2001**, *123*, 4781 – 4791.
 24. F. Weinhold and J. E. Carpenter, *Plenum* **1988**, 227
 25. Appl Environ Microbiol. **1976** Jul ;32 (1):166-71 Adsorption of bacteria to roots as related to host specificity in the Rhizobium-clover symbiosis.
 26. Krivdin, L.B.; Della, E.W. *Progress in NMR Spectroscopy*, **1991**, *23*, 301 – 610
 27. Contreras, R.H.; Peralta, J.E. *Progress in NMR Spectroscopy*, **2000**, *37*, 321 – 425
 28. Bernstein, F.C.; Koetzle, T.F.; Williams, G.J.B.; Meyer, E.F. Jr.; Brice, M.D.; Rogers, J.R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.*, **1977**, *112*, 535 – 542

29. (a) Foster, J.P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211 – 7218 (b) Weinhold, F.; Landis, C.R. *Chem. Ed.: Res. & Pract.* **2001**, *2*, 91–104 (c) Wilkins, S.J.; Westler, W.M.; Markley, J.M.; Weinhold, F. *J. Am. Chem. Soc.* **2001**, *123*, 12026 – 12036
30. (a) Badenhop, J.K.; Weinhold, F. *J. Chem. Phys.* **1997**, *107*, 5406 (b) Badenhop, J.K.; Weinhold, F. *J. Chem. Phys.* **1997**, *107*, 5422
31. Reed, A.E.; Curtiss, L.A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899-926
32. (a) Zhao HQ, Pan QF, Zhang WH, Carmichael, I.; Serianni, A.S. *J. Org. Chem.* **2007**, *72* (19), 7071 – 7082 (b) Fraser, R.R.; Kaufman, M.; Morand, P.; Govil, G. *Can. J. Chem.* **1969**, *47*, 403-409; (c) Fukui, H.; Baba, T.; Inomata, H.; Miura, K.; Matsuda, H. *Mol. Phys.* **1997**, *92*, 161-165 (d) Alkorta, I.; Elguero, J. *Theor. Chem. Acc.* **2004**, *111*, 31-35.
33. Manuscript in Preparation
34. Malkina, O.L.; Interpretation of Indirect Nuclear Spin-Spin Coupling Constants. In *Calculation of NMR and EPR Parameters*, (eds. Kaupp, M; Bühl, M.; Malkin, V.G.) **2004**, 307 – 324. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
35. Case, D.A.; Scheurer, C.; Brüschweiler, R. *J. Am. Chem. Soc.* **2000**, *122*, 10390 – 10397
36. Sklenár, V.; Munzarová, M.L. *J. Am. Chem. Soc.* **2003**, *125*, 3649 – 3658
37. (a) Thibaudeau, T.; Stenutz, R.; Hertz, B.; Klepach, T.E.; Zhao, S.; Wu, Q.; Carmichael, I.; Serianni, A.S. *J. Am. Chem. Soc.* **2004**, *126*, 15668 – 15685 (b) Klepach, T.E., Carmichael, I.; Serianni, A.S. *J. Am. Chem. Soc.* **2005**, *127*, 9781 – 9793 (c) Bose-Basu, B.; Klepach, T.E.; Bondo, G.; Bondo, P.B.; Zhang, W.; Carmichael, I.; Serianni, A.S. *J. Org. Chem.* **2007**, *72*, 7511-7522
38. Podlasek, C.A.; Wu, J.; Stripe, W.A.; Bondo, P.B.; Serianni, A.S. *J. Am. Chem. Soc.* **1995**, *117*, 8635-8644.
39. Weisskopf, V.F. *Science* **1975**, *187*, 605 – 612
40. Cremer, D.; Gräfenstein, J. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2791 – 2816
41. Whittaker, J.M. *Interpolatory Function Theory*, **1935** Cambridge Univ. Press, Cambridge, England
42. Pan, Q.; Hu, X.; Klepach, T.E.; Reed, M.; Brady, J.W.; Carmichael, I.; Serianni, A.S. *Manuscript in preparation*

43. Armour, E.A.; Stone, A.J. *Proc. R. Soc. London Ser. A* **1967**, *302*, 25
44. Munzarová, M.L.; Sklenář, V. *J. Am. Chem. Soc.* **2002**, *124*, 10666
45. (a) Brady, J.W. *J. Am. Chem. Soc.* **1986**, *108*, 8153-8160 (b) Brady, J.W. *Carbohydr. Res.* **1987**, *165*, 306-312 (c) Ha, S.N.; Giammona, A.; Field, M.; Brady, J.W. *Carbohydr. Res.* **1988**, *180*, 207-221
46. Dashnau, J.L.; Sharp, K.A.; Vanderkooi, J.M. *J. Phys. Chem. B* **2005**, *109*, 24152-21459
47. Dowd, M.K.; French, A.D.; Reilly, P.J. *Carbohydr. Res.* **1994**, *264*, 1-19
48. Astley, T.; Birch, G.G.; Drew, M.G.B.; Rodger, P.M. *J. Phys. Chem. A* **1999**, *103*, 5080-5090
49. (a) Barrows, S.E.; Dulles, F.J.; Cramer, C.J.; French, A.D.; Truhlar, D.G. *Carbohydr. Res.* **1995**, *276*, 219-251 (b) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1993**, *115*, 5745-5753. (c) Cramer, C. J.; Truhlar, D. G.; French, A. D. *Carbohydrate Res.* **1997**, *298*, 1-14 (d) French, A.D.; Dowd, M.K.; Reilly, P.J. *J. Mol. Struct.: Theochem.* **1997**, *395*, 271-287
50. Csonka, G. I.; Elias, K.; Csizmadia, I. G. *Chem. Phys. Lett.* **1996**, *257*, 49-60
51. Brown, J. W.; Wladkowski, B. D. *J. Am. Chem. Soc.* **1996**, *118*, 1190-1993
52. Polavarapu, P. L.; Edwig, C. S. *J. Comput. Chem.* **1992**, *13*, 1255-1261
53. Momany, F.A.; Appell, M.; Strati, G.; Willett, J.L. *Carbohydrate Res.* **2004**, *339*, 553-567
54. Jebber, K.A.; Zhang, K.; Cassady, C.J.; Chung-Phillips, A. *J. Am. Chem. Soc.* **1996**, *118*, 10515-10524
55. (a) Zuccarello, F.; Buemi, G. *Carbohydrate Res.* **1995**, *273*, 129-145 (b) Bagno, A.; Rastrelli, F.; Saielli, G. *J. Org. Chem.* **2007**, *72*, 7373-7381
56. Molteni, C.; Parrinello, M. *Chem. Phys. Lett.* **1997**, *275*, 409-413
57. Kouwijzer, M.L.C.E.; van Eijck, B.P.; Kroes, S.J.; Kroon, J. *J. Comput. Chem.* **1993**, *14*, 1281-1289
58. Kouwijzer, M.L.C.E.; van Eijck, B.P.; Kooijman, H.; Kroon, J. *Acta Crystallogr. B* **1995**, *51*, 209-220
59. Molteni, C.; Parrinello, M., *J. Am. Chem. Soc.* **1998**, *120*, 2168-2171

60. Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471-2474
61. Parrinello, M. *Solid State Commun.* **1997**, *102*, 107-120
62. (a) Malkina, O.L.; Hricovini, M.; Bizik, F.; Malkin, V.G. *J. Phys. Chem. A* **2001**, *105*, 9188 (b) Hricovini, M.; Malkina, O.L.; Bizik, F.; Turi Nagy, L.; Malkin, V.G. *J. Phys. Chem. A* **1997**, *101*, 9756
63. Gonzalez-Outeirino, J.; Kirschner, K.N.; Thobhani, S.; Woods, R.J. *Can. J. Chem.* **2006**, *84*, 569-579
64. (a) Krivdin, L.B.; Shcherbakov, V.V.; Kalabin, G.A. *Zh. Org. Khim.*, **1986**, *22*, 342 (b) Krivdin, L.B.; Zinchenko, S.V.; Shcherbakov, V.V.; Kalabin, G.A.; Contreras, R.H.; Tufro, M.F.; Ruiz De Azuna, M.C.; Giribet, C.G. *J. Magn. Reson.* **1989**, *84*, 1 (c) Cramer, C.J.; Truhlar, D.G.; French, A.D. *Carbohydr. Res.* **1997**, *298*, 1
65. (a) Gil, V.M.S.; Teixeira-Dias, J.J.C. *Mol. Phys.* **1968**, *15*, 47 (b) Gil, V.M.S.; Alves, A.C.P. *Mol. Phys.* **1969**, *16*, 525
66. Kräutler, V.; Müller, M.; Hünenberger, P.H., *Carbohydr. Res.* **2007**, *342*, 2097-2124
67. Çarçabal, P.; Jockusch, R.A.; Hünig, I.; Snoek, L.C.; Kroemer, R.T.; Davis, B.G.; Gamblin, D.P.; Compagnon, I.; Oomens, J.; Simons, J.P. *J. Am. Chem. Soc.* **2005**, *127*, 11414-11425
68. Beeson, C.; Pham, N.; Shipps, G.; Dix, T.A. *J. Am. Chem. Soc.* **1993**, *115*, 6803-6812
69. Lobsanov, Y.D.; Gitt, M.A.; Leffler, H.; Barondes, S.H.; Rini, J.M. *J. Biol. Chem.* **1993**, *268*, 27034-27038
70. (a) Lemieux, R.U. *Acc. Chem. Res.* **1996**, *29*, 373-380 (b) Clarke, C.; Woods, R.J.; Gluska, J.; Cooper, A.; Nutley, M.A.; Boons, G. *J. Am. Chem. Soc.* **2001**, *123*, 12238-12247
71. (a) Claridge, T.D.W. *High-Resolution NMR Techniques in Organic Chemistry* **1999**, Elsevier, Netherlands (b) Lipsitz, R.S.; Tjandra, N. New Applications for Residual Dipolar Couplings: Extending the Range of NMR in Structural Biology. In *Modern Magnetic Resonance Part I: Applications in Chemistry, Biological and Marine Sciences* (ed. Webb, G.A.) **2006**, 653-660. Springer, Netherlands
72. (a) Schneider, H.J.; Hoppen, V. *J. Org. Chem.* **1978**, *43*, 3866-3873 (b) Matsui, M.; Okada, M. *Chem. Pharm. Bull.* **1971**, *19*, 395 (c) Matsui, M.; Okada, M. *Chem.*

- Pharm. Bull.* **1970**, *18*, 2129 (d) Huitric, A.C.; Carr, J.B.; Trager, W.F. *J. Pharm. Sci.* **1966**, *55*, 211
73. Haasnoot, C.A.G.; DeLeeuw, F.A.A.M.; Altona, C. *Tetrahedron* **1980**, *36*, 2783-2792
 74. (a) Parr, R.G.; Yang, W.T *Annu. Rev. Phys. Chem.* **1995**, *46*, 701-728 (b) Helgaker, T.; Watson, M.; Handy, N.C. *J. Chem. Phys.* **2000**, *113*, 9402-9409 (c) VandeVondele, J.; Mohamed, F.; Krack, M.; Hutter, J.; Sprik, M.; Parrinello, M. *J. Chem. Phys.* **2005**, *122*, 014515
 75. (a) Helgaker, T.; Jaszunski, M.; Ruud, K.; Gorska, A. *Theor. Chem. Acc.* **1998**, *99*, 175-182 (b) Guilleme, J.; San Fabian, J. *J. Chem. Phys.* **1998**, *109*, 8168-8181 (c) Barone, V.; Peralta, J.E.; Contreras, R.H.; Snyder, J.P. *J. Phys. Chem.* **2002**, *106*, 5607-5612
 76. Serianni, A.S.; Wu, J.; Carmichael, I. *J. Am. Chem. Soc.* **1995**, *117*, 8645-8650
 77. (a) Klepach, T.E.; Carmichael, I.; Serianni, A.S *J. Am. Chem. Soc.* In press (b) Ciofini, I. Use of Solvent Continuum Models in Magnetic Resonance Parameter Calculations. In *Calculation of NMR and EPR Parameters*, (eds. Kaupp, M; Bühl, M.; Malkin, V.G.) **2004**, 191-208. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
 78. (a) Barban, S.; Schultze, H.O. *J. Biol. Chem.* **1961**, *236*, 1887 (b) Garriga-Canut, M.; Schoenike, B.; Qazi, R.; Bergendahl, K.; Daley, T.J.; Pfender, R.M.; Morrison, J.F.; Ockuly, J.; Stafstrom, C.; Sutula, T.; Roopra, A. *Nature Neurosci.* **2006**, *9*, 1382 (c) Kilbourne, E.D. *Nature* **1959**, *183*, 271 (d) Huang, R.; Dietsch, E.; Lockhoff, O. *FEBS Lett.* **1991**, *291*, 199 (e) Purohit, S.C.; Pohlitz, W. *Int. J. Radiat. Oncol. Biol. Phys.* **1982**, *8*, 495
 79. Hu, X.; Klepach, T.E.; Carmichael, I.; Serianni, A.S *Manuscript in Preparation*
 80. Houdier, S.; Perez, S. *J. Carbohydr. Chem.* **1995**, *14*, 1117
 81. Duval, E.; Koide, S. *Phys. Lett.* **1964**, *8*, 314
 82. (a) Ramsey, N.F.; *Phys. Rev.* **1953**, *91*, 303 (b) Pople, J.A.; Santry, D.P. *Mol. Phys.* **1964**, *8*, 1
 83. Pople, J.A.; McIver Jr., J.W.; Ostlund, N.S. *Chem. Phys. Lett.* **1967**, *1*, 465
 84. Case, D.A. NMR Parameters in Nucleic Acids. In *Calculation of NMR and EPR Parameters*, (eds. Kaupp, M; Bühl, M.; Malkin, V.G.) **2004**, 341 – 351. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

85. Fischer, H. The Central Limit Theorem from Laplace to Cauchy: Changes in Stochastic Objectives and in Analytical Methods In *The Different Forms and Applications of the Central Limit Theorem in its Development from Classical to Modern Probability Theory*, **2000**, 1-36. Shaker Verlag, Aachen
86. (a) Muddaseni, P.R.; Bernet, B.; Vasella, A. *Helv. Chim. Acta* **1994**, *77*, 334 (b) Tang, T.H.; Whitfield, D.M.; Douglas, S.P.; Csizmadia, L.G.; Krepinsky, J.J. *Can. J. Chem.* **1992**, *70*, 2234 (c) Martin, X.; Moreno, M.; Lluch, J.M. *Tetrahedron* **1994**, *50*, 6689
87. (a) Bock, K.; Duus, J.Ø. *J. Carbohydr. Chem.* **1994**, *13*, 513-543 (b) Rockwell, G.D.; Grindley, T.B. *J. Am. Chem. Soc.* **1998**, *120*, 10953-10963
88. (a) Lin, W.C. *J. Chem. Phys.* **1969**, *50*, 1890-1891 (b) Lin, W.C. *J. Chem. Phys.* **1970**, *52*, 2805-2814 (c) Lin, W.C. *J. Chem. Phys.* **1971**, *55*, 4971-4981 (d) Hägele, G.; Batz, M.; Peters, R.; Niemeyer, U.; Engel, J. *Arzneim-Forsch / Drug Res.* **1990**, *40*, 599-602
89. Gwin, W.D.; Pitzer, K.S. *J. Chem. Phys.* **1948**, *16*, 303
90. (a) Džakula, Ž.; Westler, W.M.; Edison, A.S.; Markley, J.L. *J. Am. Chem. Soc.* **1992**, *114*, 6195-6199 (b) Džakula, Ž.; Edison, A.S.; Westler, W.M.; Markley, J.L. *J. Am. Chem. Soc.* **1992**, *114*, 63200-63207 (c) Džakula, Ž.; Westler, W.M.; Markley, J.L. *J. Mag. Res. Ser. B* **1996**, *111*, 109-126 (d) Džakula, Ž.; DeRider, M.L.; Markley, J.L. *J. Am. Chem. Soc.* **1996**, *118*, 12796-12803
91. (a) Ko, H.; Shim, G.; Kim, Y. *Bull. Korean Chem. Soc.* **2005**, *26*, 2001-2006 (b) Lee, K.; Lee, S.; Jhon, G.; Kim, Y. *Bull. Korean Chem. Soc.* **1998**, *19*, 569-575 (c) Acquotti, D.; Sonnino, S. *Methods Enzymol.* **2000**, *312*, 247
92. Suzuki, T. *Phys. Chem. Chem. Phys.* **2008**, *10*, 96-105
93. (a) Steiner, T. *Angew. Chem., Int. Ed.* **2002**, *41*, 48 (b) Umeyama, H.; Morokuma, K. *J. Am. Chem. Soc.* **1977**, *99*, 1316 (c) Ludwig, R. *Angew. Chem., Int. Ed.* **2001**, *40*, 1808 (d) Dannenberg, J.J.; Haskamp, L.; Masunov, A. *J. Phys. Chem. A* **1999**, *103*, 7083
94. (a) Lemieux, R.U. *Abstr. Papers, Am. Chem. Soc.* **1959**, *135*, 5E (b) Lemieux, R.U.; Pavia, A.A.; Martin, J.C.; Watanabe, K.A. *Can. J. Chem.* **1969**, *47*, 4427
95. (a) Trapp, M.L.; Watts, J.K.; Weinberg, N.; Pinto, B.M. *Can. J. Chem.* **2006**, *84*, 692 (b) Lücken, E.A.C. *J. Chem. Soc.* **1959**, *111*, 2954 (c) Romers, C.; Altona, C.; Buys, H.R.; Havinga, E. *top. Stereochem.* **1969**, *4*, 39 (d) Hoffmann, R.; Radom, L.; Pople, J.A.; Schleyer, P.v.R.; Hehre, W.J.; Salem, L. *J. Am. Chem. Soc.* **1972**, *94*, 6221 (e) Eisenstein, O.; Anh, N.T.; Jean, Y.; Devaquet, A.; Cantacuzene, J.; Salem,

- L. *Tetrahedron*, **1974**, 30, 1717 (f) Wolfe, S.; Whangbo, M.-H.; Mitchell, D.J. *Carbohydr. Res.* **1979**, 69, 1 (g) Edward, J.T. *Chem. Ind. (London)* **1955**, 1102 (h) Lemieux, R.U.; Chü, N.J. *Abstr. Papers, Am. Chem. Soc.* **1958**, 133, 31N (i) Praly J.-P.; Lemieux, R.U. *Can. J. Chem.* **1987**, 65, 213 (j) Bitzer, R.S.; Barbosa, A.G.H.; Silva, C.O.d.; Nascimento, M.A.C. *Carbohydr. Res.* **2005**, 340, 2171–2184
96. Juaristi, E.; Cuevas, G.; *The Anomeric Effect* **1995**, CRC Press, Boca Raton
 97. Suzuki, T.; Kawashima, H.; Sota, T. *J. Phys. Chem. B* **2006**, 110, 2405-2418
 98. (a) Allinger, N.L.; Yuh, Y.H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, 111, 8551-8565 (b) Lii, J.-H.; Allinger, N.L. *J. Am. Chem. Soc.* **1989**, 111, 8566-8575 (c) Lii, J.-H.; Allinger, N.L. *J. Am. Chem. Soc.* **1989**, 111, 8576-8582
 99. (a) Eisenberg, D.; McLachlan, A.D. *Nature* **1986**, 319, 199-203 (b) Wesson, L.; Eisenberg, D. *Protein Science* **1992**, 1, 227-235
 100. (a) Brooks, B.R.; Bruccoleri, R.E.; Olafson, B.D.; States, D.J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, 4, 187 (b) Kuttel, M.; Brady, J.W.; Naidoo, K.J. *J. Comput. Chem.* **2002**, 23, 1236-1243
 101. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. *J. Chem. Phys.* **1983**, 79, 926-935
 102. (a) Case, D.A.; Pearlman, D.A.; Caldwell, J.W.; Cheatham III, T.E.; Ross, W.S.; Simmerling, C.S.; Darden, T.A.; Merz, K.M.; Stanton, R.V.; Cheng, A.L.; Vincent, J.J.; Crowley, M.; Ferguson, D.M.; Radmer, R.J.; Seibel, G.L.; Singh, U.C.; Weiner, P.K.; Kollman, P.A. AMBER 8.0 [computer program] **1997**, University of California San Francisco, San Francisco, California (b) Woods, R.J.; Dwek, R.A.; Edge, C.J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, 99, 3832

CHAPTER 8:
CRYSTAL STRUCTURE METHYL β -ALLOLACTOSIDE (METHYL 6-O-BETA-D-
[1- ^{13}C]-GALACTOPYRANOSYL-BETA-D-GLUCOPYRANOSIDE)
MONOHYDRATE⁷

“Art, Glory, Freedom fail, but Nature still is fair!”

– Lord George Gordon Byron

*“Subjects which disclose their full power, meaning and beauty as soon as they are
presented to the mind have very little of those qualities to disclose.”*

– Charles Dutton

8.1. Introduction

The disaccharide β -D-galactopyranosyl-(1 \rightarrow 6)-D-glucopyranose, commonly known as allolactose, is most widely known as an inducer (in its reducing form) of the *lac* operon. Allolactose acts as a negative allosteric effector of the *lac* repressor (LacI)¹⁻⁴. In *E. coli* lactose is converted to allolactose through trans-galactosylation by β -galactosidase¹⁻³. Notwithstanding the biological significance of this basic disaccharide in prokaryotic metabolism, allolactose is not currently available commercially⁵. Synthetic

⁷ I gratefully acknowledge my co-authors Bruce Noll and Anthony S. Serianni for their assistance in preparing this chapter, which is in preparation for submission to the Acta Crystallographica. I also wish to acknowledge Meredith Reed from Omicron Biochemicals for her synthetic efforts in this report.

routes to allolactose have either been via *a*) traditional Koenigs-Knorr glycosylation involving the condensation of a galactosyl donor and a glucosyl acceptor², or *b*) chemo-enzymatic trans-galactosylation^{6,7}. A novel synthetic route for the production of allolactose from the naturally occurring and commercially available cyanogenic glycoside amygdalin via an epimerization at C4 of the nascent galactose residue⁵ as well as the solid phase synthesis of Gal β 1 \rightarrow 6Glc containing glyco-peptides have also been reported⁸.

While the crystal structure of the isolated disaccharide has not been solved previous to this report, allolactose appears as a 2.0 Å and 2.3 Å and resolution co-crystal with the C-lobe of lactoferrin (PDB ID: 2dyx and 2dwj respectively)^{9,13}, and twice in the 1.5 Å resolution co-crystal with β -galactosidase (PDB ID: 1jz8)¹⁰.

In addition to its presence as a free disaccharide, the β -D-galactopyranosyl-(1 \rightarrow 6)-D-glucopyranose linkage is found in a variety of glyco-conjugated contexts. An example of this includes a ceramide trihexoside (Gal β 1 \rightarrow 6Gal β 1 \rightarrow 6Glc β 1 \rightarrow 1Cer) isolated from eggs of the sea urchin, *H. pulcherrimus*¹².

8.2. Comment

As the body of structural information involving the Gal β 1 \rightarrow 6Glc linkage in differing contexts continues to grow, it is crucial to have a reference geometry with which comparisons can be drawn. Additionally, having an X-ray structure of a crystal isotopically enriched with ¹³C at the anomeric site of the Gal residue also sets the stage for structural correlations with carbon based solid state NMR measurements of indirect spin-spin coupling constants.

In an attempt to separate the influence of crystal packing forces upon structure from intrinsic energetic minima, the coordinates from the crystal structure of methyl 6-O- β -D-[1- ^{13}C]-galactopyranosyl- β -D-glucopyranoside•mono-hydrate (I_E) (Figure 8.1) were used as the starting geometry for an unconstrained *in vacuo* DFT geometry optimization at the B3LYP⁴²/6-31G* level⁴³ (I_C) with *Gaussian03*⁴⁴. The ring numbering system conforms to the convention in which C1 is the anomeric carbon and C6 is the exocyclic hydroxymethyl group in both residues. The only (1 \rightarrow 6) glycosidically linked disaccharide structure found as an isolated crystal in the literature prior to this report are that of the C4 epimeric cognate of allolactose, gentiobiose (II) (β -D-glucopyranosyl-(1 \rightarrow 6)- β -D-glucopyranose)^{14,15}. This is likely due to the difficulty in crystallizing (1 \rightarrow 6) linkages, which have an additional degree of conformational mobility about the C5-C6 torsion, ω , when compared to other glycosides.

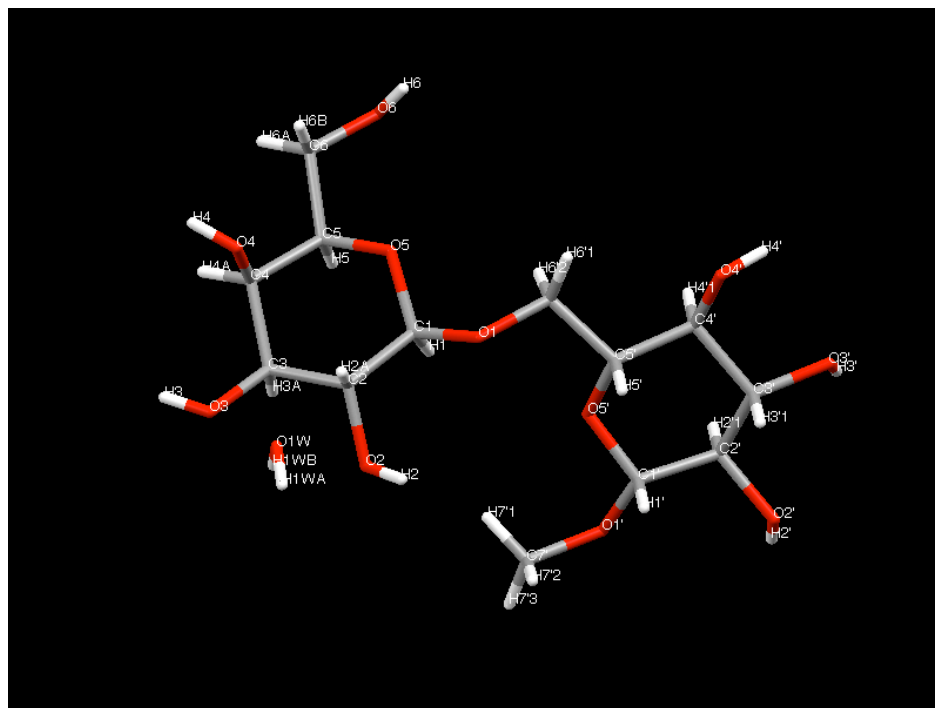


Figure 8.1. Atomic Numbering for I_E

A comparison between the structural parameters in I_E , I_C , and II^{15} , as well as select comparisons with the Gal β 1 \rightarrow 6Glc linkage as it is found in 1jz8¹⁰ (III₁, III₂ and III₃), 2dyx⁹ (IV), and 2dwj¹³ (V) can be found in Table 8.1. The average C-C and C-O bond lengths are 1.525 and 1.424 Å respectively. As noted previously in Gal β 1 \rightarrow 4Glc linked disaccharides¹¹, the endocyclic C-C bonds are uniformly longer than the exocyclic hydroxymethyl C-C bond lengths in I_E , however this is less pronounced in I_C and II . The shortening of β anomeric C1-O1 bonds with respect to other C-O bond has long been recognized^{15,16}, and is reproduced in the crystal of I_E where both anomeric C1-O1 bonds have a length of 1.398 Å. This difference is due to the influence of the *exo*-anomeric effect as a result of the orientation of the glycosidic torsion angles ϕ_1 (O5_{gal}-C1_{gal}-O1_{gal}-C6_{glc}) = -67.83°, and ϕ_2 (O5_{glc}-C1_{glc}-O1_{glc}-C7) = -63.30°. Despite the occurrence of the *exo*-anomeric effect, the fact that the two C1-O5 bond lengths in I_E are not longer than average is consonant with previous reports for II^{15} .

It is interesting that the C4-O4 bondlength in I_E and II are identical despite the inverted configuration about C4 between them. This is not the first report of configuration independence in the relative C-O bond lengths in saccharides¹¹. Conformation dependent variation in this bond length cannot be at the root of this anomaly, considering that the O4 hydroxyls in I_E and II are in equivalent conformations that orient the hydroxyl proton *syn*-periplanar to H4. The geometric optimization of I_C in principle should provide a picture of the relaxed disaccharide free of the constraints of the crystalline lattice, and indeed there is a ~0.02 Å difference in the C4-O4 bond length between I_E and I_C . Despite the fact that the bond length counter-intuitively shortened

TABLE 8.1. GEOMETRIC PARAMETERS FOR I_E AND I_C (Å, °) WITH SELECT
COMPARISONS TO II-V.

	I _E	I _C	II
C1-C2	1.527	1.533	1.515
C1'-C2'	1.526	1.529	1.521
C2-C3	1.526	1.524	1.526
C2'-C3'	1.532	1.523	1.516
C3-C4	1.535	1.546	1.521
C3'-C4'	1.540	1.528	1.514
C4-C5	1.522	1.533	1.533
C4'-C5'	1.522	1.537	1.528
C5-C6	1.510	1.527	1.514
C5'-C6'	1.514	1.523	1.510
C1-O1	1.398	1.396	1.390
C1'-O1'	1.398	1.387	1.393
C2-O2	1.423	1.420	1.427
C2'-O2'	1.438	1.421	1.420
C3-O3	1.414	1.433	1.422
C3'-O3'	1.431	1.422	1.433
C4-O4	1.433	1.415	1.433
C4'-O4'	1.418	1.418	1.438
C5-O5	1.442	1.433	1.424
C5'-O5'	1.429	1.430	1.430
C1-O5	1.426	1.418	1.415
C1'-O5'	1.408	1.425	1.428
C6-O6	1.436	1.418	1.426
C6'-O1	1.441	1.430	1.425
C1-O1-C6'	114.30	115.43	113.30
C1'-O1'-CH ₃	112.22	114.28	
C5'-C6'-O1	109.62	110.95	109.00
C5-C6-O6	111.17	110.95	112.24
C5-O5-C1	111.76	112.28	114.80
C5'-O5'-C1'	111.98	113.33	111.84
O5-C1-O1	106.85	108.97	107.19
O5'-C1'-O1'	106.89	109.49	107.88

TABLE 8.1. (CONTINUED)

	I _E	I _C	II	III ₁	III ₂	III ₃	IV	V
C1-O5-C5-C4	66.49	60.26	60.60	60.91	57.62	61.29	55.83	-59.15
C2-C1-O5-C5	-61.33	-63.92	-63.29	-60.19	-63.34	-64.63	-71.80	-59.15
C3-C2-C1-O5	51.39	58.44	56.25	56.40	59.35	59.15	75.53	40.38
C4-C3-C2-C1	-48.17	-50.90	-51.30	-53.43	-51.48	-53.62	-61.62	46.20
C5-C4-C3-C2	52.71	48.23	49.18	51.07	49.79	52.33	45.19	46.20
O5-C5-C4-C3	-60.57	-51.59	-51.68	-53.00	-54.71	-56.22	-42.04	-63.54
C1'-O5'-C5'-O4'	64.79	61.66	64.60	65.37	43.66	68.29	37.98	14.72
C2'-C1'-O5'-C5'	-65.54	-63.42	-66.44	-48.22	-74.32	-70.55	-56.92	14.72
C3'-C2'-C1'-O5'	58.89	57.94	58.75	58.47	57.99	43.30	70.62	30.48
C4'-C3'-C2'-C1'	-54.87	-54.38	-52.30	-50.46	-65.70	-37.45	-65.98	53.77
C5'-C4'-C3'-C2'	55.07	53.12	49.69	63.16	31.06	29.16	47.70	53.77
O5'-C5'-C4'-C3'	-57.83	-54.97	-54.30	-50.48	-44.25	-53.15	-32.74	-10.58
C1-C2-O2-H	74.16	144.77						
C1'-C2'-O2'-H	92.28	-159.66						
C2-C3-O3-H	167.46	-107.82						
C2'-C3'-O3'-H	18.29	-119.39						
C3-C4-O4-H	-127.53	124.98						
C3'-C4'-O4'-H	67.02	124.23						
O6-C6-C5-O5	59.48 (<i>gt</i>)	59.58 (<i>gt</i>)	-53.60 (<i>gg</i>)	-179.77 (<i>tg</i>)	-179.71 (<i>tg</i>)	-171.58 (<i>tg</i>)	114.95	49.50 (<i>g</i>)
O6-C6-C5-C4	179.92 (<i>gt</i>)	-177.88 (<i>gt</i>)	66.10 (<i>gg</i>)	-57.74 (<i>tg</i>)	-56.26 (<i>tg</i>)	-50.80 (<i>tg</i>)	-115.49	171.19 (<i>g</i>)
C5-C6-O6-H	-104.78	62.58						
O5'-C1'-O1'-CH ₃	-63.30	-73.40						
C2'-C1'-O1'-CH ₃	177.37	167.43						
O5-C1-O1-C6'	-67.83	-79.53	-58.30	-65.49	-67.67	-63.52	-19.09	-0.34
C2-C1-O1-C6'	172.93	160.75	-176.40	171.05	174.01	170.03	-141.38	-124.76
H1-C1-O1-C6'	52.37	40.93	63.25					
C1-O1-C6'-C5'	-117.84	-104.61	-156.30	-162.36	-158.41	-168.31	-77.85	81.00
C1-O1-C6'-H6'R	2.76	17.80	83.92					
C1-O1-C6'-H6'S	121.56	136.05	-41.05					
O1-C6'-C5'-O5'	63.81 (<i>gt</i>)	61.79 (<i>gt</i>)	-61.50 (<i>gg</i>)	-86.85 (<i>gg</i>)	69.41 (<i>gt</i>)	66.08 (<i>gt</i>)	114.07	4.50
O1-C6'-C5'-C4'	-176.23 (<i>gt</i>)	-177.58 (<i>gt</i>)	59.50 (<i>gg</i>)	32.71 (<i>gg</i>)	-169.73 (<i>gt</i>)	-176.03 (<i>gt</i>)	-121.36	138.39
O1-C6'-C5'-H5'	-53.77 (<i>gt</i>)	-57.28 (<i>gt</i>)	-177.88 (<i>gg</i>)					

Symbols *gg*, *gt* and *tg* denote *gauche-gauche* and *gauche-trans* and *trans-gauche*.

after energetic minimization, the result still indicates that the origin of the finding is likely due to crystal packing forces involving the O4 hydroxyl.

The valence bond angles in I_E are all within with published ranges^{15,17-18}. Consistent with previous observation¹¹, the glycosidic C-O-C bond angle for the linkage ($\sim 114^\circ$) is larger than that of the less sterically encumbered methyl aglycone ($\sim 112^\circ$).

The galacto- and the glucopyranose rings of I_E are both in the 4C_1 chair conformation, as expected. The structure of the gluco-pyranose ring in I_E as determined by the endocyclic torsions is comparable to that in I_I and as such conforms to the Arnott-Scott pyranose model¹⁹. In contrast, the galacto-pyranose ring in I_E deviates slightly from the average values in this idealized 4C_1 chair model. This is reflected in the Cremer-Pople²⁵ ring puckering parameter θ found in Table 8.2, and can be visualized using the stereographical projection convention of Jeffrey and Yates³³ (Figure 8.2). The *endo*-cyclic dihedrals $C1_{gal}-C2_{gal}-C3_{gal}-C4_{gal} = -48.17^\circ$ and $O5_{gal}-C1_{gal}-C2_{gal}-C3_{gal} = 51.39^\circ$ are the two smallest torsions found in the ring reflecting the deviation of the ring towards a puckering conformation intermediate to the canonical 0H_5 and E_5 ring forms. Upon DFT optimization, θ remained relatively constant, but φ changed by approximately -90° . This is equivalent to the ring maintaining the same relative amount of distortion from the 4C_1 conformation, but traversing the pseudorotational itinerary counter-clockwise towards the 0H_1 ring form. This is evident in the I_C torsion $C2_{gal}-C3_{gal}-C4_{gal}-C5_{gal} = 48.23^\circ$. The change is likely due to the shift in the H-bonding pattern involving the $C3_{gal}$ and $C4_{gal}$ hydroxyls from an independent intermolecular arrangement in I_E (*vide infra*) to the direct intramolecular H-bond ($O4_{gal}-H \rightarrow O3_{gal}$) found in I_C . Upon DFT optimization, the ring puckering amplitudes, Q , reduced by 3.2% and 3.5% for the galacto- and the gluco- rings respectively.

TABLE 8.2. HYDROGEN-BOND GEOMETRY (Å, °) IN CRYSTALLINE I_E.

$D-H\cdots A$	$D-H$	$H\cdots A$	$D\cdots A$	$D-H\cdots A$
O2—H2 \cdots O1 ^{vii}	0.840	1.895	2.693(4)	158
O3—H3 \cdots O3 ^{ti}	0.840	1.985	2.796(4)	162
O4—H4 \cdots O1 ^{iv}	0.840	1.868	2.699(4)	170
O6—H6 \cdots O2 ⁱⁱ	0.840	1.950	2.782(4)	170
O3'—H3 \cdots O4 ^{vi}	0.840	1.903	2.696(3)	157
O4'—H4 \cdots O1 ^v	0.840	2.037	2.829(4)	157
O1 ^W —H1 ^{WA} \cdots O6 ⁱⁱⁱ	0.88(4)	1.85(5)	2.723(4)	173
O1 ^W —H1 ^{WB} \cdots O2 ^{viii}	0.78(6)	1.99(6)	2.750(4)	165

Symmetry codes: (i) x, y, z-1; (ii) x-1, y, z; (iii) x+1, y, z; (iv) -x+1, y-1/2, -z+1; (v) -x+1, y-1/2, -z+2; (vi) -x+1, y+1/2, -z+2; (vii) 2-x, y-1/2, 2-z; (viii) -x+2, y+1/2, -z+2

In the galacto-pyranose residue of I_E, O5 is *gauche* to O6 (59.48°) and *trans* to C4 (179.92°). This ‘*gt*’ orientation of the C5-C6 dihedral is consistent with the finding that this is the most highly populated and energetically stable rotamer for the monomer in solution as determined by a scalar coupling analysis²⁰. However, in accord with the solution structure studies, other orientations for this rotamer are possible particularly within a larger context as illustrated by structures III-V. Eclipsed structures are even possible as demonstrated by IV. The O6 hydroxyl orientation in the gal- residue of I_E is defined by the dihedral C5-C6-O6-H = -104.78°. This is within the *g*-rotameric regime for this dihedral, which was identified as the most highly populated C6-O6 rotamer for the monomer in solution²⁰.

The global conformation of I_E is described by the glycosidic torsion angles ϕ (O5_{gal}-C1_{gal}-O1_{gal}-C6_{glc}) = 67.83°, ψ (C1_{gal}-O1_{gal}-C6_{glc}-C5_{glc}) = 117.84°, and ω (O1_{gal}-C6_{glc}-C5_{glc}-O5_{glc}) = -63.81°. In terms of the ϕ / ψ conformational map, I_E is

located within a cluster composed of I_C, II and III₁₋₃ spanning the relatively narrow range in ϕ from roughly -80° to just past -60° (the preferred *exo*-anomeric conformation^{23,24}), and the range in ψ from roughly -160° to -100° . This relatively tight conformational range can be misleading however, unless correlations with ω are taken into consideration. The ϕ / ψ clustering breaks up into two distinct groups across ω , one composed of I_E, I_C, III₂ and III₃ characterized by $\omega \approx 60^\circ$ (the *gt* rotamer), and the second composed of II and III₁ in which $\omega \approx -60^\circ$ (the *gg* rotamer). Early attempts to quantify torsional preferences about ω in the crystalline state identified *gg* as the most stable conformation^{21,22}. More recent reports identify the *gt* rotamer as the most highly populated conformation about ω in glucopyranose, with ω occupying *gg* roughly a third of the time in solution²⁰. While structures IV and V both show a ϕ value bordering the range of the *exo*-anomeric orientation (-19.09° and -0.34° respectively), both structures represent intermediate values for ω and ψ , demonstrating that a context dependent flattening of the energetic hypersurface can allow for divergent values across these torsions.

Ring pucker can play an important role in modulating the glycosidic conformational space, that is to say that deformation of one or both pyranose rings in a $\beta(1\rightarrow6)$ linkage from the ideal 4C_1 conformation may allow access to regions of the $\phi / \psi / \omega$ conformational map not normally explored. For example, across the spectrum of structures II – V there is a qualitative correlation between deviance from 4C_1 and glycosidic conformation (Figure 8.2). This implies that in solution there may be a correlation between fluctuations in pyranose ring pucker and glycosidic reorientation. This sort of behavior has been previously observed in nucleosides³⁴. While in isolation

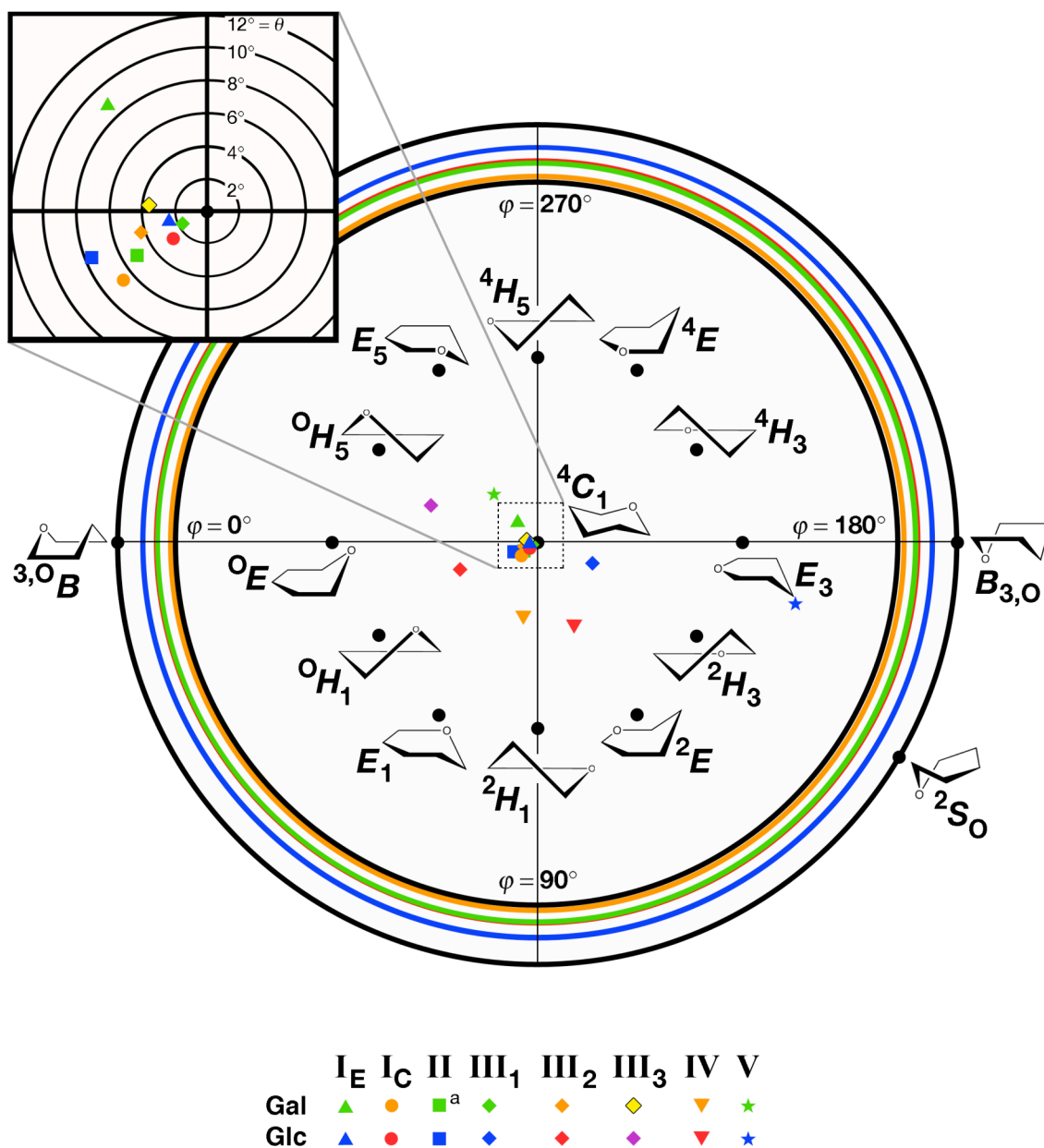


Figure 8.2. Northern Hemispheric Stereographical Projection of Pyranose Pseudorotational Itinerary with Ring Puckering Coordinates for I-V. ^aRing at the non-reducing end of the disaccharide II is glucopyranose instead of galactopyranose. The ring puckering parameters φ and θ are represented by angular and radial displacements about the ⁴C₁ origin (center point) respectively. The displacements for all crystal data in the main figure are based upon individual total puckering amplitudes, Q , while displacements in the inset are based upon an average amplitude for the depicted data ($Q_{\text{avg}} = 0.5734$). The inner and outer black circumferential rings represent the minimal and maximal Q values for the total data set, while the colored rings correspond to the Q values for the galacto and gluco rings of I_E and

I_C as per the color code in the legend. The southern hemisphere (not shown) relates via symmetry to the northern hemisphere with 1C_4 as the origin.

furanosyl rings have been recognized to be more conformationally flexible than their pyranosyl counterparts³⁵⁻³⁷, the pseudorotational itinerary of pyranose rings may be rendered more energetically accessible due to impinging macro-molecular forces. This is borne out for the Gal β 1 \rightarrow 6Glc linkage in the IV and V structures.

There are no intramolecular hydrogen bonds in the crystal of I_E . The single water molecule within the crystalline lattice acts as the nexus for four finite hydrogen bonded chains that account for all H-bonds in the crystal (Table 8.3). The water acts as donor to *a*) O6_{gal} in a chain that proceeds O6_{gal}-H \rightarrow O2_{gal}-H \rightarrow O1_{glc}, and *b*) O2_{glc}. The water molecule is the terminal H-bond acceptor for the chains O3_{gal}-H \rightarrow O3_{glc}-H \rightarrow O4_{gal}-H \rightarrow O_{wat}, and *b*) O4_{glc}-H \rightarrow O_{wat}. The fact that the O2_{glc} hydroxyl is the only one that does not act a H-bond donor is reflected in the lengthened C2_{glc}-O2_{glc} bond length (1.438 Å) with respect to the other exocyclic C-O bond lengths. Conversely, the two shortest exocyclic C-O bond lengths are C3_{gal}-O3_{gal} (1.414 Å) and C4_{glc}-O4_{glc} (1.418 Å), the only two hydroxyls that do not act as H-bond acceptors in the lattice. In addition, neither O5_{gal}, O1_{gal}, nor O5_{glc} act as H-bond acceptors.

There is an intermolecular hydrophobic contact between the methyl aglycone and the H6R_{glc} and H6S_{glc} protons adjacent to the glycoside from a neighboring unit cell, with methyl proton to hydroxy-methyl proton inter-nuclear distances of 2.602 and 2.522 Å respectively. The third methyl proton from the aglycone has a weaker intramolecular contact with H1_{gal} (3.014 Å).

TABLE 8.3. CREMER-POPLE PYRANOSE RING PUCKERING PARAMETERS²⁵
FOR I-V.

Ring: Galactopyr	q_2	q_3	Q	φ_2 (°)	θ (°)
I _E	0.0877	0.5700	0.5767	312.8	8.8
I _C	0.0620	0.5545	0.5580	39.8	6.4
II ^a	0.0481	0.5491	0.5512	33.3	5.0
III ₁	0.0159	0.5532	0.5534	28.0	1.7
III ₂	0.0414	0.5730	0.5745	18.2	4.1
III ₃	0.0356	0.5850	0.5861	353.6	3.5
IV	0.2231	0.5974	0.6378	79.2	20.5
V	0.1905	0.5156	0.5496	311.7	20.3
Ring: Glucopyr	q_2	q_3	Q	φ_2 (°)	θ (°)
I _E	0.0237	0.5999	0.6004	6.4	2.3
I _C	0.0279	0.5785	0.5791	37.2	2.8
II	0.0756	0.5763	0.5812	22.3	7.5
III ₁	0.1737	0.5589	0.5852	159.1	17.3
III ₂	0.2363	0.5088	0.5610	19.5	24.9
III ₃	0.3039	0.4793	0.5675	347.5	32.4
IV	0.2626	0.5229	0.5852	113.5	26.7
V	0.5495	0.2200	0.5918	166.7	68.2

^aRing at the non-reducing end of the disaccharide II is glucopyranose.

8.3.Experimental

8.3.1. Reagents

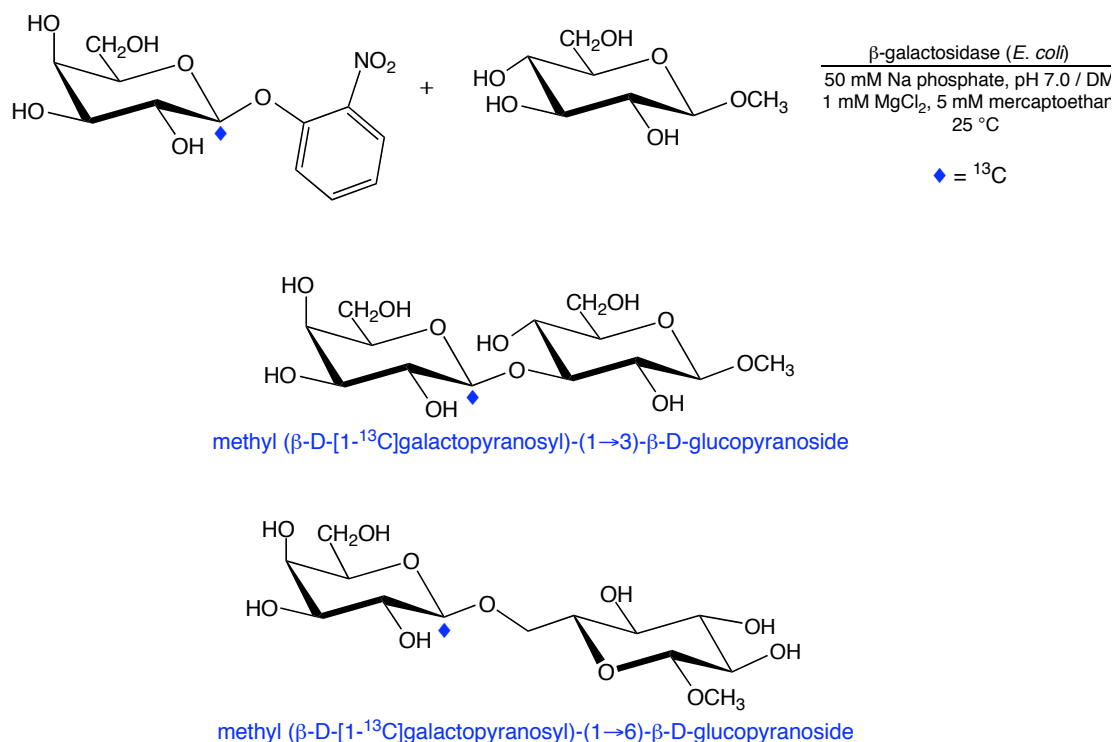
O-Nitrophenol, Sepharose G10, Dowex 1 × 2 (200-400 mesh) (Cl⁻) ion-exchange resin, β-galactosidase (E.C. 3.2.1.23) (*E. coli*), and methyl β-D-glucopyranoside (**1**) were purchased from Sigma Chemical Co.

8.3.2. Synthesis of *o*-Nitrophenyl β -D-[1- ^{13}C]-galactopyranoside (**2**)

D-[1- ^{13}C]Galactose was prepared from D-lyxose and K^{13}CN (Cambridge Isotope Laboratories; 99 atom-% ^{13}C) by cyanohydrin reduction, yielding D-[1- ^{13}C]galactose and D-[1- ^{13}C]talose;⁴⁵⁻⁴⁷ The galacto and talo epimers were purified by chromatography on Dowex 50 \times 8 (200-400 mesh) (Ca^{2+}),²⁹ with the galacto isomer eluting first. *o*-Nitrophenyl β -D-[1- ^{13}C]galactopyranoside **2** was prepared as described previously from D-[1- ^{13}C]galactose in an overall yield of 30%⁴⁸.

8.3.3. Synthesis of Disaccharide I_E via Enzyme-catalyzed Transglycosylation

Conditions for the transglycosylation reaction between **1** and **2** (Scheme 8.1) were identical to those reported by Nilsson;^{26,27} the reaction was conducted with 1.35 g (4.48 mmol) of **2** and 2.50 g (12.9 mmol) of **1**. After the reaction was quenched, the mixture was concentrated in vacuo at 40 °C to ~45 mL and applied to a column (2.5 \times 55 cm) of Sepharose G10. Elution with distilled water gave a phenol-sulfuric acid²⁸ positive peak near the column void. This disaccharide-containing fraction was collected, concentrated to ~35 mL, and the solution was applied to a column (2.5 \times 55 cm) of Dowex 1 \times 2 (200-400 mesh) (OH^-) ion-exchange resin.²⁹ Elution with distilled water (3.8 mL/fraction, 0.8 mL/min) gave three phenol-sulfuric acid positive peaks. NMR analysis of each peak indicated the following: Peak 1, residual **1**; Peak 2, disaccharide I_E ; Peak 3, methyl 3-O- β -D-[1- ^{13}C]-galactopyranosyl- β -D-glucopyranoside. Disaccharide I_E was subsequently reduced to a clear colorless viscous syrup and characterized by ^1H and ^{13}C NMR.



Scheme 8.1. Coupling Reaction for the Synthesis of I_E

8.3.4. Crystal growth of I_E

Individual crystals suitable for X-ray diffraction were grown by an adaptation³⁰ of published protocols³¹ for the use of the lambda tube of Gravatt and Gross³². Approximately 20 mg of I_E was dissolved into 300 μL of distilled water and placed in the sample reservoir of a lambda tube. This apparatus had a Nichrome wire wrapped around the arm nearer to the sample reservoir for heating purposes, and a length of copper wire closely wound about the cross arm acting as cooling ribs. The apparatus was gently filled with a 1:3:6 mixture of butanol:methanol:ethanol to above the level of the cross arm, the open ends were stoppered, and a low current was applied across the Nichrome wire with a Variac such that the temperature of the heated ‘ascending’ arm was ~50° C. After ten days a tiny needle shaped seed crystal had formed on the cooler ‘descending’ arm of the

lambda tube. The seed was removed, placed in a capped vial with the mother liquor, which had been significantly reduced in volume, and set aside on the benchtop. The vial was not tightly capped in the hopes that the remaining solvent would slowly evaporate away, which eventually happened after ~18 months, leaving behind colorless needle shaped crystals.

8.3.5. Data Collection

Crystallographic data were collected through the SCrALS (Service Crystallography at Advanced Light Source) program at the Small-Crystal Crystallography Beamline 11.3.1 (developed by the Experimental Systems Group) at the Advanced Light Source (ALS). The ALS is supported by the U.S. Department of Energy, Office of Energy Sciences Materials Sciences Division, under contract DE-AC02-05CH11231 at Lawrence Berkeley National Laboratory. Data collection: *APEX2*³⁹; cell refinement: *APEX2* and *SAINT*³⁹; data reduction: *SAINT* and *XPREP*^{39,40}; program(s) used to solve structure: *XS*⁴⁰; program(s) used to refine structure: *XL*⁴⁰; molecular graphics: *XP*⁴⁰; software used to prepare material for publication: *XCIF*⁴⁰ and *enCIFer*⁴¹.

8.3.6. Instrumentation

Bruker d8-Apex II CCD diffractometer
 ω scans
10449 measured reflections
2192 independent reflections
1920 reflections with $I > 2\sigma(I)$

$R_{\text{int}} = 0.0867$
 $\theta_{\text{max}} = 29.5344^\circ$
 $h = -10 \rightarrow 10$
 $k = -11 \rightarrow 11$
 $l = -16 \rightarrow 16$

8.3.7. Refinement

Refinement on F^2
 $R[F^2 > 2\sigma(F^2)] = 0.047$
 $wR(F^2) = 0.1124$
 $S = 1.047$
2192 reflections
243 parameters

H-atom parameter treatment mixed
 $w = 1/[\sigma^2(F_o^2) + (0.0534P)^2 + 0.0000P]$
where $P = (F_o^2 + 2F_c^2)/3$
 $(\Delta/\sigma)_{\text{max}} = 0.000$
Absolute structure: Flack³⁸
Flack parameter: 0.0 (10)

8.3.8. Crystal Data

$\text{C}_{13}\text{H}_{24}\text{O}_{11}$, H_2O
 $M_r = 374.34$
Monoclinic, $P2_1$
 $a = 7.528$ (2) Å
 $b = 8.744$ (4) Å
 $c = 12.695$ (11) Å
 $\beta = 100.15$ (4)°
 $V = 822.6$ (8) Å³
 $Z = 2$

$D_x = 1.511$ Mg m⁻³
X-ray synchrotron radiation
Cell parameters from 2490 reflections
 $\theta = 2.9958 - 29.5344^\circ$
 $\mu = 0.135$ mm⁻¹
 $T = 150$ (2) K
Blade, clear colorless
 $0.12 \times 0.06 \times 0.01$ mm

8.4. References

1. Jacob, F.; Monod, J. *J. Mol. Biol.* **1961**, *3*, 318
2. Burstein, C.; Cohn, M.; Kepes, A.; Monod, J. *Biochim. Biophys. Acta* **1965**, *95*, 634
3. Jobe, A.; Bourgeois, S. *J. Mol. Biol.* **1972**, *69*, 397
4. Yildirim, N.; Mackey, M.C. *Biophys. J.* **2003**, *84* (5), 2841
5. Prosperi, D.; Panza, L.; Haltrich, D.; Nonini, M.; Riva, S. *J. Carbo. Chem.* **2003**, *22* (5), 267
6. Huber, R.E.; Wallenfels, K.; Kurz, G. *Can. J. Biochem.* **1975**, *53*, 1035
7. Pazur, J.H.; Tipton, C.L.; Budovich, T.; Marsh, J.M. *J. Am. Chem. Soc.* **1958**, *80*, 119
8. Roberge, J.Y.; Beebe, X.; Danishefsky, S.J. *Science* **1995**, *269*, 202
9. Mir, R.; Premkumar, R.; Sinha, M.; Singh, N.; Sharma, S.; Kaur, P.; Bhushan, A.; Singh, T.P. *Data to be published*
10. Juers, D.H.; Heightman, T.D.; Vasella, A.; McCarter, J.D.; Mackenzie, L.; Withers, S.G.; Matthews, B.W. *Biochemistry* **2001**, *40*, 14781
11. Pan, Q.; Noll, B.; Serianni, A.S. *Acta Cryst. C* **2005**, *61*, 674
12. Kubo, H.; Jiang, G.J.; Irie, A.; Morita, M.; Matsubara, T.; Hoshi, M. *J. Biochem.* **1992**, *111*, 726
13. Mir, R.; Singh, N.; Sinha, M.; Sharma, S.; Bhushan, A.; Singh, T.P. *Data to be published*
14. Arène, F.; Neuman, A.; Longchambon, F. *C.R. Acad. Sc. Paris S.C* **1979**, *288*, 331
15. Rohrer, D.C.; Sarko, A.; Bluhm, T.L.; Lee, Y.N. *Acta Cryst. B* **1980**, *36*, 650
16. Berman, H.M.; Chu, S.C.; Jeffery, G.A. *Science* **1967**, *157*, 1576
17. Jeffrey, G.A. *Am. Chem. Soc. Symp. Ser.* **1979**, *87*, 50
18. Jeffrey, G.A. *Am. Chem. Soc. Symp. Ser.* **1990**, *430*, 20

19. Arnott, S.; Scott, W.E. *J. Chem. Soc. Perkin Trans.* **1972**, 2, 324
20. Thibaudeau, T.; Stenutz, R.; Hertz, B.; Klepach, T.E.; Zhao, S.; Wu, Q.; Carmichael, I.; Serianni, A.S. *J. Am. Chem. Soc.* **2004**, 126, 15668
21. Marchessault, R.H.; Pérez, S. *Biopolymers* **1979**, 18, 2369
22. Tvaroska, I.; Pérez, S.; Marchessault, R.H. *Carbohydr. Res.* **1978**, 61, 97
23. Lemieux, R.U. *Pure Appl. Chem.* **1971**, 25, 527
24. Pérez, S.; Marchessault, R.H. *Carbohydr. Res.* **1978**, 65, 114
25. Cremer, D.; Pople, J.A. *J. Am. Chem. Soc.* **1975**, 97, 1354
26. Nilsson, K.G.I. *Carbohydr. Res.* **1987**, 167, 95
27. Nilsson, K.G.I. *Carbohydr. Res.* **1988**, 180, 53
28. Hodge, J.E.; Hofreiter, B.T. *Methods Carbohydr. Chem.* **1962**, 1, 380
29. Austin, P.W.; Hardy, F.E.; Buchanan, J.C.; Baddiley, J. *J. Chem. Soc.* **1963**, 5350
30. Klepach, T.E. *Pais. Pars. Parc. P.* **2008**, 35, 1 – 7
31. Hope, H. *J. Appl. Cryst.* **1971**, 4, 333
32. Gravatt, C.C.; Gross, P.M. *J. Chem. Educ.* **1969**, 46, 693
33. Jeffrey, G.A.; Yates J.H. *Carbohydr. Res.* **1979**, 74, 319
34. Nilsson, L.; Foloppe, N. *J. Phys. Chem. B* **2005**, 109, 9119
35. Altona, C.; Sundaral, M. *J. Am. Chem. Soc.* **1972**, 94, 8205
36. Altona, C.; Sundaral, M. *J. Am. Chem. Soc.* **1973**, 95, 2333
37. French, A.D.; Finch, P. Monosaccharides: Geometry and Dynamics. In *Carbohydrates: Structures, Syntheses and Dynamics*, (Finch, P. ed.) **1999**, 1 – 46. Kulwer Academic Publishers, Boston
38. Flack, H.D. *Acta Cryst. A* **1983**, 39, 876
39. Bruker-Nonius **2007**, *APEX2* and *SAINT*, Bruker-Nonius AXS, Madison, Wisconsin, USA.

40. Sheldrick, G.M. *Acta Cryst. A* **2008**, 64, 112
41. CCDC **2005**, *enCIFer*, The Cambridge Crystallographic Data Center, Cambridge, UK.
42. Becke, A. D. *J. Chem. Phys.* **1993**, 98, 5648
43. Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J Chem. Phys.* **1972**, 56, 2257
44. Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, 2004.
45. Serianni, A.S.; Barker, R. Synthetic Approaches to Carbohydrates Enriched with Stable Isotopes of Carbon, Hydrogen and Oxygen. In *Isotopes in the Physical and Biomedical Sciences*; (eds. Buncel, E.; Jones, J.R.) **1987**, 211. Elsevier: New York
46. Serianni, A.S.; Vuorinen, T.; Bondo, P.B. *J. Carbohydr. Chem.* **1990**, 9, 513
47. Serianni, A.S.; Nunez, H.A.; Barker, R. *Carbohydr. Res.* **1979**, 72, 71
48. Conchie, J.; Levvy, G.A. *Methods Carbohydr. Chem.* **1963**, 2, 335

CHAPTER 9:

SYNTHESIS

*“It is the pervading law of all things organic and inorganic,
Of all things physical and metaphysical,
Of all things human and all things super-human,
Of all true manifestations of the head,
Of the heart, of the soul,
That the life is recognizable in its expression,
That form ever follows function. This is the law.”*

– *Louis Sullivan*

“Define your terms, you will permit me again to say, or we shall never understand one another”

– *Voltaire*

9.1. Definition of Terms

The above quote from the early 20th century architect Louis Sullivan became the modernist mantra for an era of design that spurned aesthetic consideration in favor of functionality: *form ever follows function*. Sullivan’s ironically expansive and poetic

words allude to his unlikely influences; the biological sciences. Nearly a century earlier Charles Darwin realized that the anatomy of an organism evolves to serve the function that satisfies the selective pressures exerted by the environs. However, the trouble of aesthetics that plagued the modernist architects and designers was moot to Darwin; aesthetics is an essential element of behaviour as a vehicle for determining fitness and as such a transitive aspect of biological function. The structure/function paradigm still reigns supreme in the contemporary biosciences, however with more reconcilable definitions of the terms structure and function.

Traditionally, while function implies activity, form has been thought of in terms of a static three-dimensional structure. More current definitions within research bioscience seek a more holistic reconciliation between the kairological aspects of function and the temporal variation in the structure of the system. For example, in the ecological sciences it has recently been shown that a complete picture of spatial and temporal variation of nutrient uptake in headwater streams requires a detailed consideration of the structural variation in the stream on time scales ranging from daily to seasonal and annual.¹ In the last decade protein biochemists have expanded the definition of protein structure to include a temporal component that describes the functionally significant conformational averaging of a molecule.² Due to the comparatively high degree of conformational mobility in saccharide systems, this concept of time averaged structure becomes of even greater functional importance.

An ordinal hierarchy of structural classification analogous to that used in proteomics is useful in the discussion of structural glycobiology. The *primary* structure of a saccharide refers to the sequential identity of all glycosyl residues in the molecule

including absolute and anomeric configuration, and all glycosidic linkage identities. I refer to the *secondary* structure of a saccharide as meaning the total static three-dimensional conformation at a single point in time including all intra-molecular interactions. *Tertiary* structure refers to the spatially distributed inter-molecular contacts within the larger macromolecular context. This includes both solvent interactions as well as those with other solutes such as proteins or other biomolecules, small organic solutes or inorganic ionic species. The *quaternary* structure of a saccharide is similar to the tertiary structure except that it is considered in the time domain as an ensemble of conformations which exist in relative percentages and whose geometric reorientations are characterized by a potentially broad spectrum of time scales. Thus, quaternary structure has two components: the ensemble of conformational states and the time scale of their samplings. A subtle aspect of this quaternary structure is the potential for both correlated and auto-correlated motions in the molecular framework across these time scales.

9.2. Primary Structure

In addition to being the most evolutionary ancient class of biopolymer, carbohydrates have the highest potential for structural diversity. For example there are 256 possible tetramers of RNA or DNA, ~160,000 possible tetra-peptides, and $>8 \times 10^7$ possible linear tetra-saccharides, only considering biologically common subunits. Before the relatively sophisticated consideration of this quaternary structure can begin, appropriate methods for addressing the lower elements of saccharide structure must be developed and refined. Originating with the 19th century work of Emil Fischer, the emphasis during most of structural glycobiology's long history has been on developing

techniques for primary structure determination. This was no mean feat. A consideration of the wide array of primary structural determinants such as the biologically relevant *a*) members of the aldose and ketose trees, *b*) variety of chemical derivatizations such as *N*-acetylation or oxidation, *c*) anomeric configurations, *d*) pyranosyl and furanosyl ring forms or any of the various open chain species, *e*) potential glycosidic linkage identity and *f*) array of possible multi-antennary branched oligosaccharide structures reveals that from amongst the major classes of biomolecules, saccharides have the highest potential for primary structural diversity.

In the face of such a rich molecular repertoire, many gaps in basic structural knowledge persist. For example, knowledge of the tautomeric distribution that a particular monosaccharide exhibits is fundamental primary structure information. Despite the biological importance of sialic acid, **chapter 2** is the first systematic quantification of all open chain species in this nine carbon α -keto-acid. This report goes on to provide valuable information on the effects of pH on the relative amounts of the tautomeric forms and an account of the variation in the pKa of the carboxyl group across the tautomeric species.

9.3. Secondary Structure

If the saccharide in question can be obtained in a single homogeneous crystal of a sufficient size, X-ray or neutron diffraction crystal structures are the most readily available source of unambiguous secondary structure in saccharides. The major caveat intrinsic to this method arises from the static nature of the structure. The specific three-dimensional conformation that the molecule adopts is highly subject to crystal packing

forces and is not necessarily representative of the major solution conformers or the geometry of a bound state. Crystal structures are an important source of structural information nonetheless. They act as a reference structure for other forms of structural data, as well as offering a reasonable set of starting coordinates for simulations. Additionally, if the crystalline material was synthesized with ^{13}C isotopic enrichment, it is in theory possible to measure the ^{13}C based spin-spin couplings in the solid state. This latter technique offers the possibility of correlating accurate coupling constant measurements with an unequivocal molecular geometry. This can serve as a sterling benchmark for validating theoretically calculated spin-couplings. For example, in **chapter 8** I reported the high resolution crystal structure of methyl 6-*O*- β -D-[1- ^{13}C]-galactopyranosyl- β -D-glucopyranoside•mono-hydrate, commonly known as methyl-allolactoside. This structure provided valuable structural information about the rarely crystallized (1 \rightarrow 6) glycosidic linkage. This sample could be used to measure the J_{CH} and J_{CC} involving C1 in the *galacto* residue in the solid state.

9.4. Higher Order Structure

If there is to be a commensurate definition of structure and function in carbohydrates, quantitative techniques need to be developed to address conformational mobility on multiple time scales. This conformational mobility takes many forms. The time scale of motions is roughly proportional to the total nuclear mass involved in the motion. Ignoring the IR band, the first to consider is along the carbon backbone. In pyranosyl and furanosyl structures this manifests as an endocyclic sampling about the pseudorotational itinerary and sampling about the exocyclic C-C bond(s). Pseudorotation

is typically more facile in furanoid structures than in six-membered rings, although substitution patterns can greatly affect this. The typical ring form for most D-hexopyranoses is the 4C_1 chair although a survey of the crystallographic databases reveals that this can be significantly influenced by contextual forces. For obvious reasons pseudorotation is a highly auto-correlated phenomenon in terms of the individual C-C dihedrals that constitute the ring.

Exocyclic groups such as hydroxyls and *N*-acetyl substituents reorient around their respective C-O or C-N bonds. While the rotation of a single hydroxyl is in an explicit sense independent of the rotation of other hydroxyls, they can be highly correlated with one another. This can be through direct interactions between proximal groups such as electrostatic or steric clashing or intra-molecular H-bonding, or via indirect interactions that are mediated by solvent or intervening substituents. A good example of this is demonstrated in **chapter 7** in which the results of a conformational analysis on methyl- β -D-glucopyranoside in DMSO indicate the presence of two distinct intra-molecular H-bonding networks in roughly equal proportions.

A special case of C-O reorientation is about the ϕ and ψ dihedrals that constitute the glycosidic linkage. The time scales and range of these motions within the allowable conformational space are likely to be highly context dependent. For example, in a typical disaccharide one can imagine relatively fast conformational fluctuations about central ϕ and ψ values that more or less conform to a particular rotameric regime within the allowable conformational space. A second conformational state may exist that is characterized by different central ϕ and ψ values. The interconversion between these two conformational states will take place on a longer time scale than the principal

fluctuations. If this linkage finds itself in the milieu of a larger oligosaccharide, the range of motions in terms of the glycosidic conformational preferences can be affected relative to the isolated disaccharide. Reorientations that represent conformational shifts involving the entire oligosaccharide can have timescales even longer than those associated with the shifts between conformational states about a single linkage.

The above-mentioned motions are typically too fast to be able to cleanly resolve into distinct conformational populations with NMR. Thus, the measurement of a NMR parameter such as a spin-spin coupling constant represents a single time averaged value across the ensemble of sampled conformations. However through the use of multiple spin-spin couplings reporting redundant structural information, the single time averaged geometry can be teased apart into the contributions from the constituent conformational states. This is the first part of quaternary structure as defined above, the second being the relative time scales involved in the conformational reorientations. Luckily these two components can be considered in isolation, the first step being a determination of the conformational ensemble. Throughout chapters the general approach to structural glycobiology determination has been the application of a suite of theoretical methods to the interpretation of various NMR parameters, indirect spin-spin coupling in particular.

The eagerness of many glycoscientists to produce results in large biologically interesting systems has pre-empted the development of robust methods as applied to smaller systems and ultimately produced many conflicting reports with dubious conclusions. There must be a firm understanding of mono-saccharide behaviour before extending into higher order structures. The first step in establishing ^{13}C based NMR spin-spin coupling as a viable tool for conformational glycobiology is the ability to measure

the appropriate J_{CH} and J_{CC} values in the relevant monosaccharide systems. **Chapters 3** and **6** report the ensemble of experimentally measured ^{13}C -based spin-couplings in the anomers of sialic acid and a suite of aldohexopyranoses respectively. The study on sialic acid included the effects of pH on all J_{CH} and J_{CC} couplings to carbons 1 through 3. This represents the first published account of the effect of the ionization state of a proximal carboxyl moiety on ^{13}C -based spin-couplings in a pyranose system. The study of J_{CC} in hexopyranoses demonstrated that the complete ensemble of ^{13}C - ^{13}C couplings can be measured and interpreted throughout the ring.

The magnitude of indirect spin-spin coupling is traditionally thought of as being modulated the spatial disposition of the coupled nuclei with respect to one another³. However the quantitative application of spin-spin coupling measurements to the conformational analysis of saccharides requires a more nuanced view of coupling. The carbon backbone of a saccharide functions as the scaffold for the lone pair electrons decorating the hydroxyl substituents. As such, couplings involving the carbons in the glycosyl backbone are exquisitely sensitive to the spatial disposition of these electrons with respect to the coupling pathways. Consequently the ^{13}C -based spin-couplings report on the orientation of the hydroxyl appendages along the coupling path. I showed that geminal couplings involving ^{13}C are particularly useful in this regard. **Chapters 4** and **5** showed the usefulness of $^2J_{CCH}$ in the conformational analysis of the hydroxy-methyl group and the anomeric hydroxyl respectively, while **chapter 7** demonstrated that $^2J_{CCC}$ is a valuable probe of hydroxyl conformation along the coupling pathway. This latter structural correlation is significantly more pronounced for hydroxyls appended to the central carbon of the coupling path. **Chapter 7** goes on to show that these relationships

between geminal coupling magnitude and proximal hydroxyl conformation can be interpreted in terms of discrete interactions between hydroxyl lone pair electrons and either bonding or anti-bonding orbitals along the coupling path. These interactions with either bonding or anti-bonding orbitals represent distinct and concerted steric (through space) or hyperconjugative (through bond) coupling mechanisms.

9.5. Modular Conformational Analysis

While all spin-spin couplings in a saccharide will show a magnitude sensitivity to at least one geometric parameter, most will have multiple sensitivities. Notable exceptions are vicinal couplings involving hydroxyl protons, which are almost exclusively dependent upon the subtended dihedral of the coupling path. Couplings with multiple conformational sensitivities typically have them break down into primary (larger) and secondary (smaller) dependencies. Due to the presence of multiple geometric coupling determinants, it is difficult to isolate a single geometric parameter with enough redundant couplings to unambiguously deconvolute the time averaging of the coupling(s) into a reliable conformational population distribution. For this reason it is often necessary to consider multiple geometric parameters simultaneously by using the ensemble of coupling constants sensitive to the geometric parameters in question. Thus the pyranose ring can be broken up into distinct structural regions each with its own collection of couplings that is sensitive to the geometric parameters in that region. This is a modular approach to conformational analysis.

Chapter 4 describes the interpretation of a suite of experimental couplings sensitive to CH₂OH conformation in the anomers of methyl-*galacto* and *gluco*pyranoside

in aqueous solution. Due to the dependency of a number of the couplings on both ω (O5-C5-C6-O6) and θ (C5-C6-O6-H) conformation, a picture of the correlated conformational behaviour of the two dihedrals emerged. The hydroxy-methyl group is well suited to modular conformational analysis by virtue of its exocyclic nature and its proximity to the relatively conformationally immobile ring oxygen (*i.e.* there is no hydroxyl substituent on that ring position).

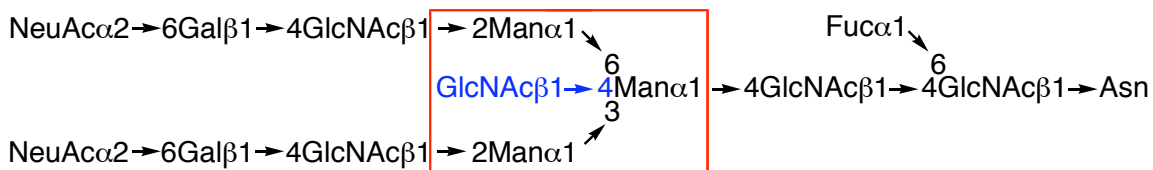
On the other side of the ring oxygen is the anomeric center. **Chapter 7** describes the simultaneous conformational analysis of the C1-O1, C2-O2 and C3-O3 dihedrals of β -methyl-glucopyranoside in DMSO based upon an extensive set of experimental spin-couplings. This molecular fragment was well suited to modular analysis again due to the relative isolation afforded by the proximity to the ring oxygen. Additionally the fact that the substituents on the extreme ends of the molecular fragment, namely the O1 methoxy and O3 hydroxyl were in plane with the C1 through C3 carbon backbone of the pyranose ring conferred a heightened degree of conformational sensitivity to a number of couplings used in the analysis. The analysis identified two distinct intramolecular H-bonding networks as the primary conformational states present in a DMSO solution.

Given the above two modular parameterizations, the sole remaining molecular region that requires geometric parameterization to complete the tools necessary for a fully integrated picture of monosaccharide molecular geometry is the C2 through C4 molecular fragment. Many of the couplings parameterized in the anomeric region analysis of **chapter 7** can be applied to the conformational analysis of the glycosidic torsion ϕ . This alludes to a further potential ‘molecular fragment’, namely the glycosidic region. In fact there has been significant preliminary results on the conformational analysis of the

methyl α -man-(1 \rightarrow 2)- α -man disaccharide indicating that the ensemble of trans-glycoside spin-spin couplings can be interpreted in terms of ϕ and ψ populational profiles. This opens the door to a more holistic conformational analysis via multi-modular approach across various orders of molecular structure that would include *intra*-residue and *inter*-residue modulation throughout a larger oligosaccharide.

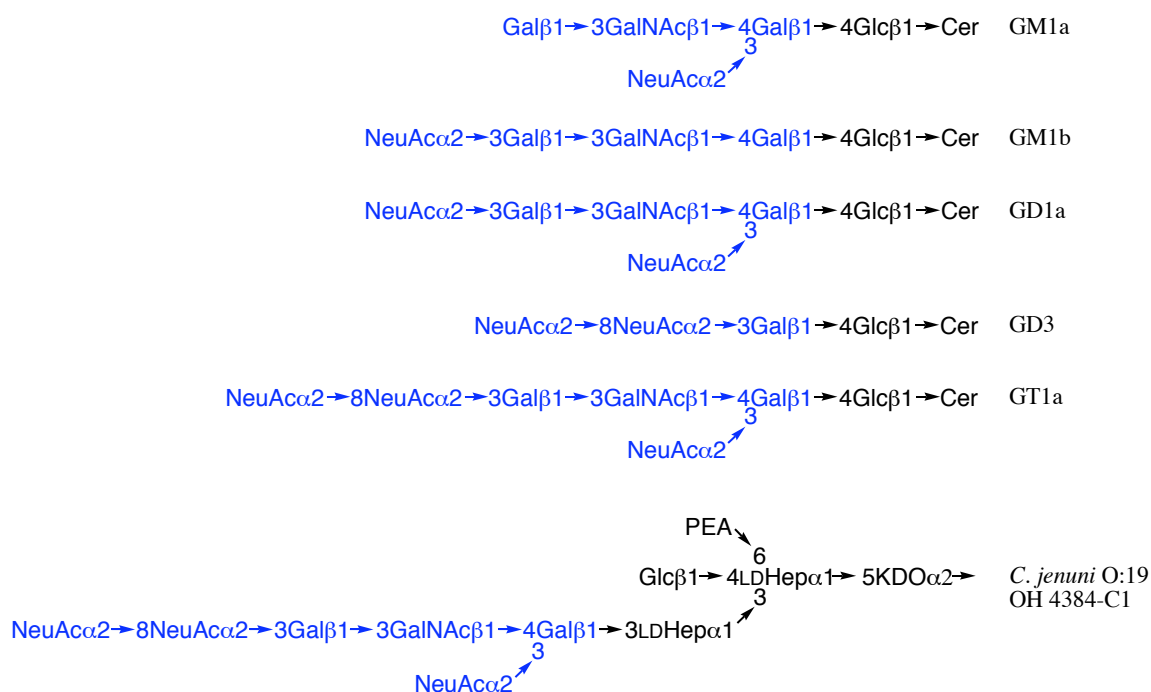
9.6. Biological Applications

There is an inexhaustible number of biological applications for this sort of conformational analysis. For example, the common complex N-glycan in depicted in Scheme 9.1 can be found in a diverse array of biological contexts such as in crops like *Phaseolus vulgaris* or *Triticum vulgaris* to decorating human IgG or sperm cells. MD simulations conducted in our lab on the tetra-saccharide in the red box indicate that the presence of the bifurcating GlcNAc in a β (1 \rightarrow 4) linkage (blue) to the bridgehead Mannose residue alters the conformational profiles of the hydroxyls in the vicinity of the linkage. The serianni lab is in the process of synthesizing an isotopically enriched tetra-saccharide (red box) with strategic labeling at the carbons involved in the glycosidic linkages. This compound will be ideally suited for conformational studies using a variety of NMR methods including spin-spin coupling.



Scheme 9.1. Complex N-Glycan.

An example of the potential application of saccharide conformational analysis to further the understanding of pathogenesis is in the case of intestinal *Campylobacter jejuni* infection as the secondary etiological agent for the development of Guillian Barré Disease. *C. jejuni* causes acute gastroenteritis. *C. jejuni* serotype O:19 evades the host immune system by producing lipopolysaccharides with terminal structures that resemble the human neuronal GD1a, GD3, GM1 and GT1a gangliosides⁴ (Scheme 9.2).



Scheme 9.2. Structure of *Campylobacter jejuni* Serotype O:19 Antigenic Lipopolysaccharide OH 4384-C1 and Related Human Neolacto Series Gangliosides. Residues in blue are structurally related portions of the antigenic terminus of the OH 4384-C1 LPS. Select abbreviations: LDHep, L-glycero-D-manno-heptose; PEA, O-phosphoethanolamine; KDO, 3-deoxy-D-manno-octulosonic acid.

Guillain Barré Disease (GBD) is a form of neuropathy that causes widespread muscular wasting and is the most common cause of generalized paralysis⁵. The link between *C. jejuni* infection and subsequent development of GBD has been recognized for

over 15 years. The method of pathogenesis appears to be mediated by an autoimmune response elicited by *C. jejuni* against the neuronal cerebroside of the neolacto series gangliosides⁶. In principle the autoreactive T-cells that recognize these gangliosides are cleared postnatally during a process called thymic deletion. In light of this, the precise mechanism for antigenic stimulation is not well understood, however a method of antigenic cross-reactivity with a T-cell subpopulation not removed during thymic deletion is likely. This could potentially involve the antigenic determinant from *C. jejuni* adopting a slightly different conformational profile from that present in the native gangliosides. This conformational difference could be contextually mediated by the divergent core structures present in the O:19 serotype versus the native gangliosides. A careful comparative conformational analysis of these two glycosyl fragment pools including the core structures could shed light on the molecular disease mechanism and ultimately lead to improved therapies such as a tunable antigen specific leucopheresis.

9.7. References

1. (a) Hoellein, T.J.; Tank, J.L.; Rosi-Marshall, E.J.; Entrekin, S.A.; Lamberti, G.A. *Limnol.Oceanogr.* **2007**, 52, 1964 – 1977 (b) Roberts, B.J.; Mulholland, P.J. *J. Geophys. Res.* **2007**, 112, G4.
2. Bhalla, J.; Storch, G.B.; MacCarthy, C.M.; Uversky, V.N.; Tcherkasskaya, O. *Mol. Cell. Proteomics* **2006**, 5, 1212 – 1223.
3. (a) Karplus, M. *J. Am. Chem. Soc.* **1963**, 85, 2870 (b) Karplus, M. *J. Am. Chem. Soc.* **1969**, 91, 1-10
4. (a) Aspinall, G.O.; Fujimoto, S.; McDonald, A.G.; Pang, H.; Kurjanczyk, L.A.; Penner, J.L. *Infect. Immun.* **1994**, 62, 2122-2125 (b) Aspinall, G.O.; McDonald, A.G.; Pang, H.; Kurjanczyk, L.A.; Penner, J.L. *Biochemistry* **1994**, 33, 241-249 (c) Aspinall, G.O.; McDonald, A.G.; Pang, H. *Biochemistry* **1994**, 33, 250-255

5. Ropper, A.H. *N. Engl. J. Med.* **1992**, 326, 1130-1136
6. Kuroki, S.; Saida, T.; Nukina, M.; Haruta, T.; Yoshioka, M.; Kobayashi, Y.; Nakanishi, H. *Ann. Neurol.* **1993**, 33, 243-247