Supporting Information

Accurately Modeling the Conformational Preferences of Nucleosides

*Mihai Burai Patrascu, Elise Malek-Adamian, Masad J. Damha and Nicolas Moitessier**

Department of Chemistry, McGill University, Montréal, Quebec, Canada, H3A 0B8

nicolas.moitessier@mcgill.ca

Table of Content

I. QM/MM MD Protocol

QM/MM MD Study

All calculations were carried out in AMBER12[1], using SCC-DFTB[2] for the QM region (ligand) and the ff99SB force field[3] for the MM region (solvent). Compound **5** was initially optimized using GAMESS-US[4],[5] at the HF/6-31G* level of theory in order to generate the electrostatic potential, which was subsequently used to generate the RESP charges in AMBER12. The system was then constructed by solvating **5** in a pre-equilibrated box of TIP3P[6] water molecules. Following this step the topology and coordinate files for the system were generated. Visual inspection of the final system was undertaken in order to ensure its integrity. With the solvated system in hand, 2000 steps of energy minimization were performed (500 steps of steepest descent energy minimization, followed by 1500 steps of conjugate gradient energy minimization). Then, an equilibration run of 200 ps was performed (NPT, 303K) using a Langevin thermostat,[7] with a collision frequency γ = 2.0 ps^{-1} and a step of 2 fs. The SHAKE[8] algorithm was used in order to fix all the hydrogen-containing bonds to equilibrium values; periodic boundary conditions were used, with a cutoff of 8 Å for non-bonded interactions (including the QM method/TIP3P electrostatic interactions). The particle mesh Ewald (PME)[9,10] method was used to control the long-range electrostatic interactions. The equilibration run was followed by a production run of 500 ns (NVT, 303 K, 1 atm) where the same conditions as above were applied. System integrity was ensured at the end of each step described above.

The structure of **5** obtained after the production run was used to generate starting conformations in which the dihedral angle O4′-C4′-C5′-O5′ had values of 60, 120, 180, 240, 300 and 360°. The structures for each angle were subjected to the same protocol described above. The angles were restrained by applying a harmonic biasing potential, with a force constant of 200 kcal/mol/rad^2. For each structure, the angle distribution was plotted to ensure that the dihedral angle was restrained at the correct value.

Umbrella Sampling Simulations

The reaction coordinate for the umbrella sampling simulations was chosen to be the exocyclic torsion angle O4′-C4′-C5′-O5′.Using the above described structures 72 windows were built (window width of 5° each) to cover the 0-360° range for the O4′-C4′-C5′-O5′ dihedral angle– each window was subjected to a slightly modified protocol than the one describe above: the parameters were identical to the ones described above but the equilibration run was 100ps and the production run was 1ns, leading to a total of 72 ns of production runs. Once the simulations were finished, the systems were visually inspected to ensure integrity was maintained. Data obtained from umbrella sampling simulations can be readily analyzed by using the weighted histogram analysis method (WHAM)[11], which provides the free energy of the reaction as a function of a chosen reaction coordinate (also known as potential of mean force or PMF) and average population distribution for the chosen reaction coordinate. Thus the PMF and average distributions for the pseudorotation phase angle P were obtained using the WHAM software.[12] For the pseudorotational angle P, a force constant of 0 was used to unbias the angle, since no biasing was applied in the first place.

II. NBO, MO and NMR Prediction Analysis Protocol.

Natural Bond Orbital (NBO) Analysis: The 2 minima obtained following the QM/MM calculations for **5** were subjected to an NBO analysis using ORCA v.3.0.3[13] and the NBO6 program[14,15] at the M06L/def2 TZVP[16,17] level of theory. The threshold for the E_2 energies (hyperconjugation and anomeric effects) was 0.05 kcal/mol.

Molecular Orbital Analysis: The molecular orbitals were built for the *North* and *South* conformations of **5** using the Molekel[18] software. The molecular orbitals used in building the maps were obtained following the natural bond orbital analysis described above.

Note: The protocols detailed in Sections S1 and S2 were applied to all nucleosides used in this study; the protocol in Section S1 was applied to the monosaccharides used in this study.

NMR Prediction Calculations:

The 2 minima obtained following the QM/MM calculations for **5** were subjected to geometry optimization of the base (in the gas phase at the PBE0[19]/pcSseg-2[20] level of theory) followed by NMR spectrum prediction at the same level of theory. TMS was used as a reference for all NMR predictions (optimized at the same level of theory).

For the computations involving 6 conformations, representative structures were extracted from the QM/MM calculations at P values of 18, 58.5, 100, 168.5, 233 and 305° and the bases were optimized at the PBE0/pcSseg-2 level of theory. NMR prediction calculations were performed at the same level of theory. TMS was used as a reference for all NMR predictions (optimized at the same level of theory).

III. Benchmark Study

DFT Calculations

Recently one of our laboratories carried out the synthesis and conformational preferences of nucleoside **5.**[21] Oligonucleotides containing this nucleoside analogue exhibit favorable binding properties, increased nuclease resistance, and perform well in RNAi gene knockdown experiments. This nucleoside is of interest as the methoxy substituent C-4′ alters the stereoelectronics and subsequently the sugar conformation. Therefore, we decided to start our benchmark study with DFT calculations using ORCA v.3.0.3[13] focused on a number of envelope conformations that **5** could adopt in solution. We constrained one dihedral angle for each (to maintain the envelope conformation) and optimized these conformations using an implicit solvent. The energies we obtained for each conformation are plotted in Figure S1. Moreover, we subjected all envelope conformations to frequency calculations in order to ensure that they were energy minima. Furthermore, we collected the puckering parameters for each envelope conformation using PROSIT[22] and the results are presented in Table S1.

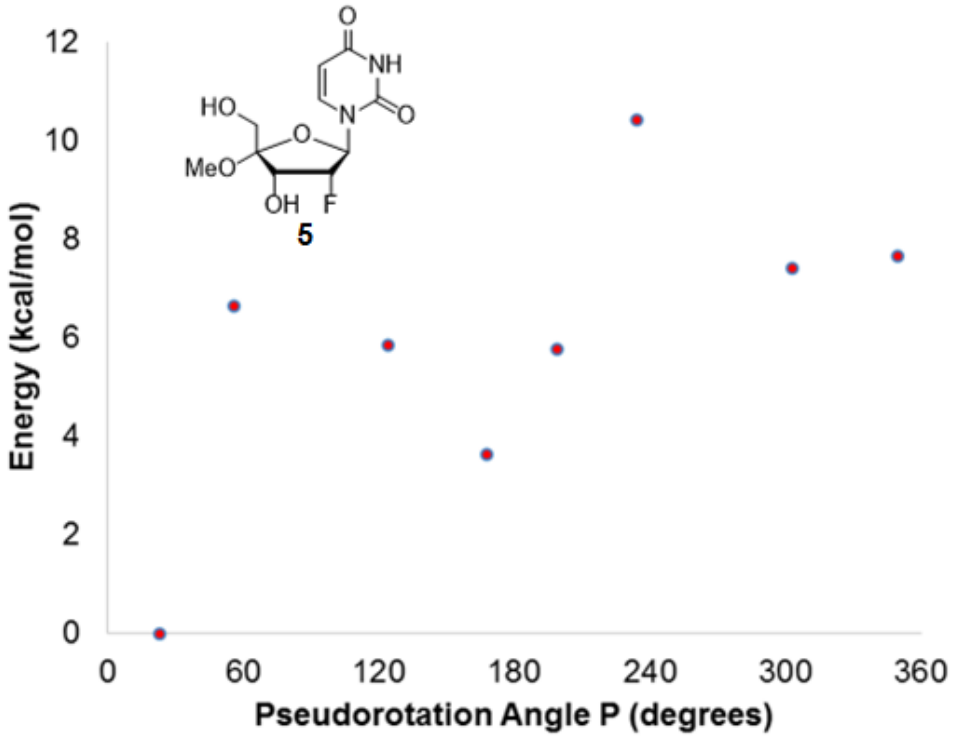


Figure S1. Energy curve obtained for compound **5** using M06L/def2-TZVP[16-17] and the COSMO[23] solvent model (water).

As shown in Figure S1 our computed data suggested a significantly large *North* (*N*) preference with an energy difference well above 3.5 kcal/mol relative to the *South* (*S*) conformer, hence a *N/S* ratio greater than 100:1. However, the experimental *North/South* ratio obtained by NMR experiments at 303 K for **5**[21] revealed a preference for the *Northern* conformation in the ratio *N/S* 87:13 ($\Delta E_{(North/South)}$ = 1.15 kcal/mol).

In order to explain such a deviation, one must look at the underpinnings of DFT. First of all, although DFT is able to describe the hyperconjugation effects, it only provides static conformations of the molecule, meaning that the dynamic behavior of sugar puckering would not be described fully. Moreover, the solvent (water) used in DFT-solution phase calculations is implicit. As continuum solvation does not consider individual water molecules, the possibility of hydrogen-bonding between the sugar hydroxyls and water is non-existent. Utilizing explicit water molecules would add a high degree of complexity (e.g., location, orientation of water molecules) and CPU time to the calculations, thus making the calculations intractable. In addition, a decision would have to be made as to how many and where the water molecules would be placed. Poor placement (and estimate on the number) of water molecules will likely yield incorrect results. Therefore, we reasoned that the cause of this inaccurate prediction of the *N/S* ratio likely resides in the improper computation of intermolecular hydrogen bond strengths, as well as the lack of explicit waters which would modulate the stereoelectronic effects.

Table S1. Data obtained following DFT calculations on **5**.

Envelope Conformation	P* (°)	ϕ_m**(°)	Pucker Type	Energy (kcal/mol)	Energy (Eh)
^{1}E	303	26.11	C1′-endo	7.40	-1049.8754
^{2}E	168	38.20	C2′-endo	3.64	-1049.8814
^{3}E	23	33.71	C3′-endo	0.00	-1049.8872
^{4}E	234	41.90	C4′-endo	10.42	-1049.8706
$_1$E	124	43.88	C1′-exo	5.84	-1049.8779
$_2$E	350	34.59	C2′-exo	7.66	-1049.8750
$_3$E	199	32.51	C3′-exo	5.77	-1049.8780
$_4$E	56	45.37	C4′-exo	6.65	-1049.8766

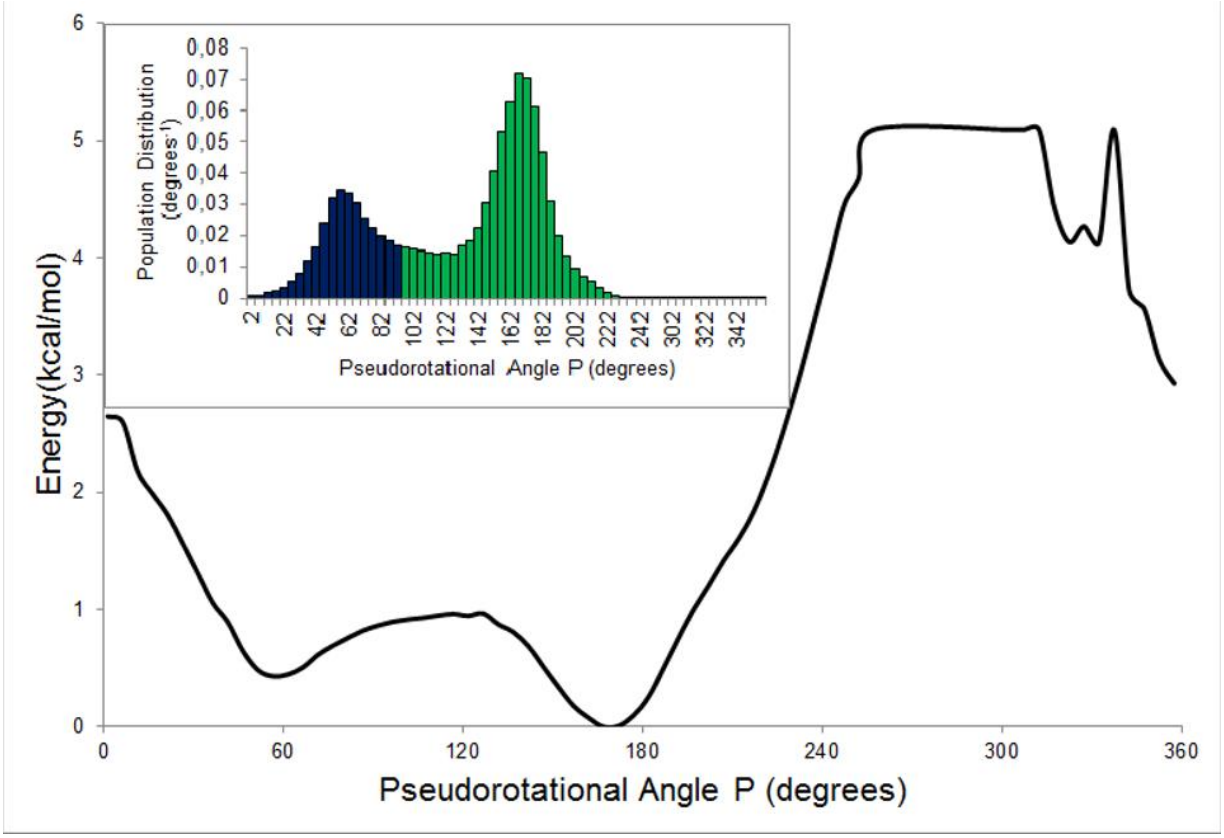
*P = pseudorotational phase angle; **ϕ_m = puckering amplitude

Molecular Mechanics

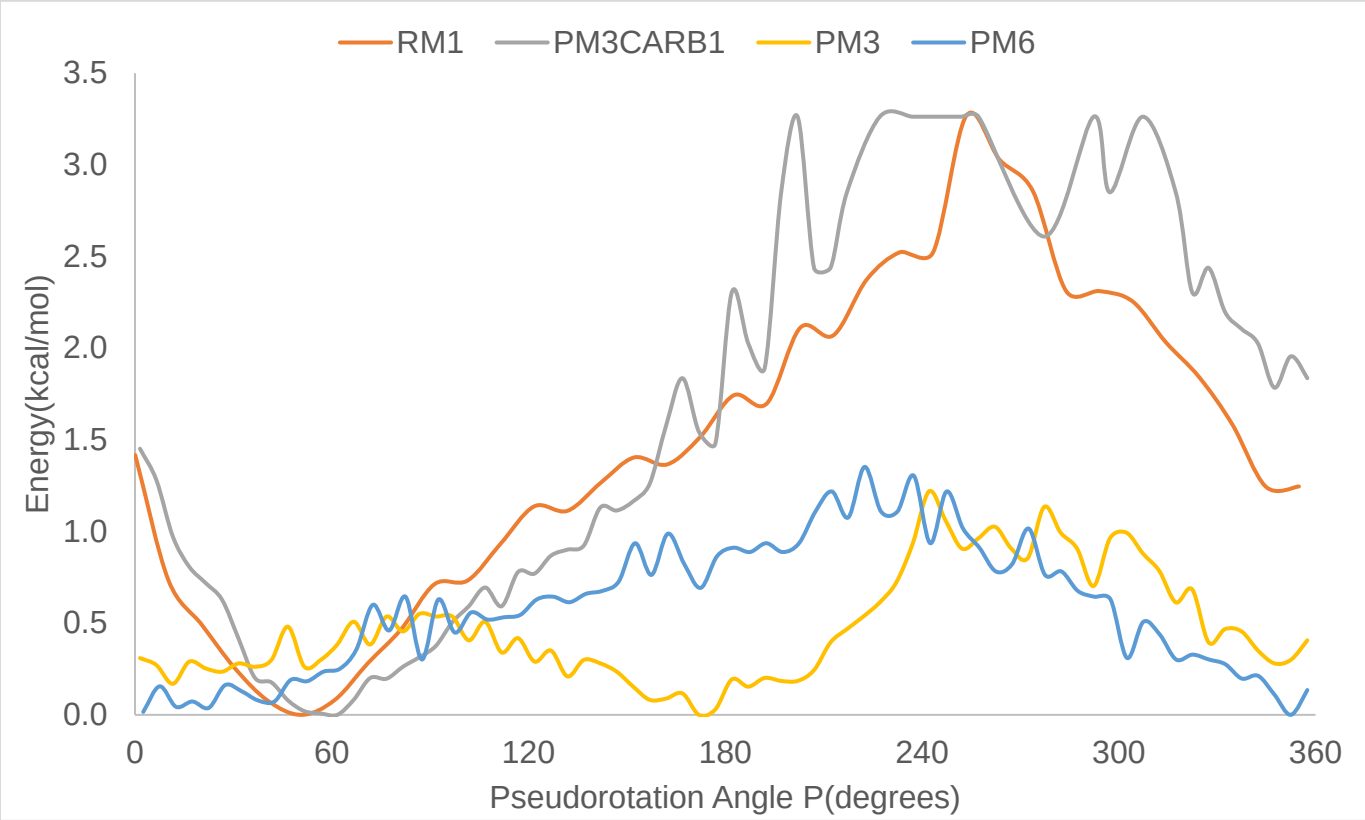
Recently, Roy and co-workers[24] suggested umbrella sampling simulations with the GLYCAM parameter set were suitable for treating furanosides, and thus we turned our attention to this approach in our study as well. Nonetheless, such a simulation requires a specific reaction coordinate in the context of which sugar puckering can be analyzed. Roy and co-workers showed that the rotation about the exocyclic C4-C5 bond of both compound **6** and **7** (Chart S1) occurs on a faster time scale than sugar puckering, thus making the analysis of puckering in the context of an exocyclic torsion reasonable. We decided to use the same exocyclic torsion as our reaction coordinate (Figure S2d). Another advantage of using umbrella sampling simulations was the significantly lower CPU-time requirement (since they rely on MM), which enabled the use of explicit solvents that we believe may be critical for optimal predictions.

The free energy (also known as Potential of Mean Force – PMF) of sugar puckering as well as average sugar pucker population distributions could be easily obtained from our simulations. This data is shown in Figure S2a.

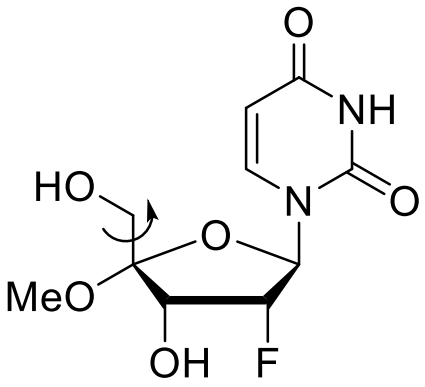
a)


Energy(kcal/mol)
Pseudorotational Angle P (degrees)
Population Distribution (degrees-1)
Pseudorotational Angle P (degrees)

b)


RM1
PM3CARB1
PM3
PM6
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Energy(kcal/mol)
0
60
120
180
240
300
360
Pseudorotation Angle P(degrees)

c)


O
NH
N
O
HO
O
MeO
OH
F

d)

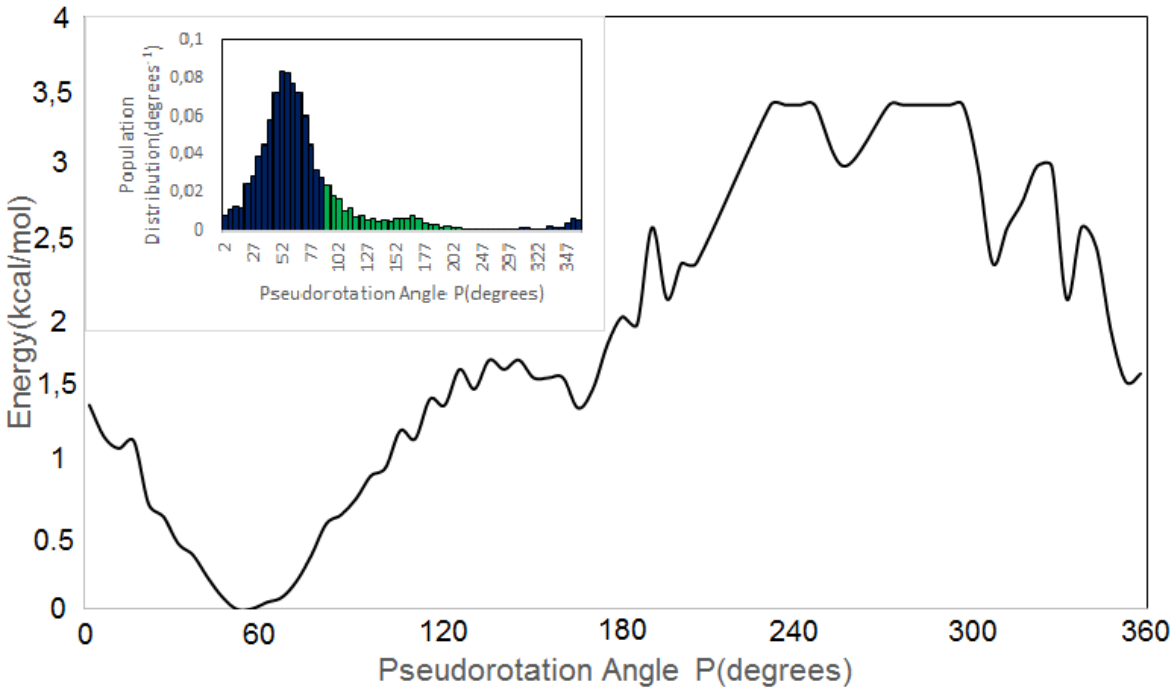


Figure S2. a) PMF curve along the pseudorotational phase angle for **5** using GLYCAM. Inset shows the sugar pucker distribution. (blue– *Northern* hemisphere, green – *Southern* hemisphere). b) PMF curves obtained using the RM1, PM6, PM3 and PM3CARB1. c) Exocylic C4-C5 bond rotation of **5** chosen as coordinate for the umbrella sampling simulations. d) PMF curve of **5** using SCC-DFTB. Inset shows the sugar pucker distribution (blue – *Northern* hemisphere, green – *Southern* hemisphere).

Although this method identifies two distinct minima, it assigns the global energy minimum for **5** (and the most populated) to the *Southern* conformation. This is contradicted by experiment, which assigns the *Northern* region as the most populated one, thus rendering the pure MM approach with the GLYCAM parameter set unviable. A closer look at the previous study shows that the systems previously treated with the AMBER/GLYCAM approach focuses on mono- and oligosaccharides lacking strong electron-withdrawing substituents in the 2′ and 4′ positions (such as -F and –OMe present in **5**). It is known that MM methods are not explicitly parametrized to describe strong hyperconjugation effects such as the anomeric or gauche effect; the GLYCAM parameter set does not have specific fluorine parameters, meaning that generic parameters (the ff99SB force field) were applied to our system. Since these electronic effects play a major role in determining sugar puckering,[25] it is reasonable to assume that the failure of this method stems from this improper description of these effects. As an additional limitation, since MM methods explicitly consider nuclei but not electrons, most commonly used force fields used in MM (including the ff99SB force field used for these calculations) are non-polarizable and thus molecular polarizability is not accounted for any given system.

QM/MM Investigations Using RM1[26], PM3[27] , PM3CARB1[28] and PM6[29].

In AMBER12 an electrostatic embedding scheme is implemented, meaning that there is an interaction between the MM point charges and the QM electrons and the atomic cores (for semi-empirical methods). Moreover, using a cutoff of 8Å between the QM and MM regions allows for the long-range electrostatic interactions between the QM and MM regions to be described using a full multipole treatment. In this framework, hydrogen bonds between nucleosides (QM region) and solvent (MM region) are considered as a combination of van der Waals (Lennard-Jones potential) and electrostatic interactions (full multipole treatment). Between these four semi-empirical methods, PM6, RM1 and PM3CARB1 correctly described the global minimum in the *Northern* region while PM3 failed to describe this minimum altogether. However, none of these methods correctly describe a minimum in the *South* region. Moreover, the ratios obtained from these methods are inconsistent with experimental data. The inability of these methods to model the dynamics of these sugar structures has also been described by Naidoo and Barnett,[30] who argued that these semi-empirical methods produce sugar rings having inaccurately low conformational free energy barriers.

IV. PSEUROT[31]

This program requires starting *North* and *South* conformations (a two conformation system), which are then optimized; over multiple runs involving various starting *North/South* conformations the *N/S* ratio as well as the puckering parameters (P and ϕ_m) are determined and the set of conformers that best replicates experimental $^3J_{H,H}$ data is selected. However, it is important to note that more than one set of conformers can replicate the experimental coupling-constant data and thus a decision has to be made with regards to choosing one over others.[32] Moreover, the data obtained between each run may differ significantly[32] which might introduce an inherent error vis-à-vis the pseudo-experimental *N/S* ratios. Hence we decided to call the experimental data determine with PSEUROT pseudo-experimental data.

V. Nucleoside Puckering

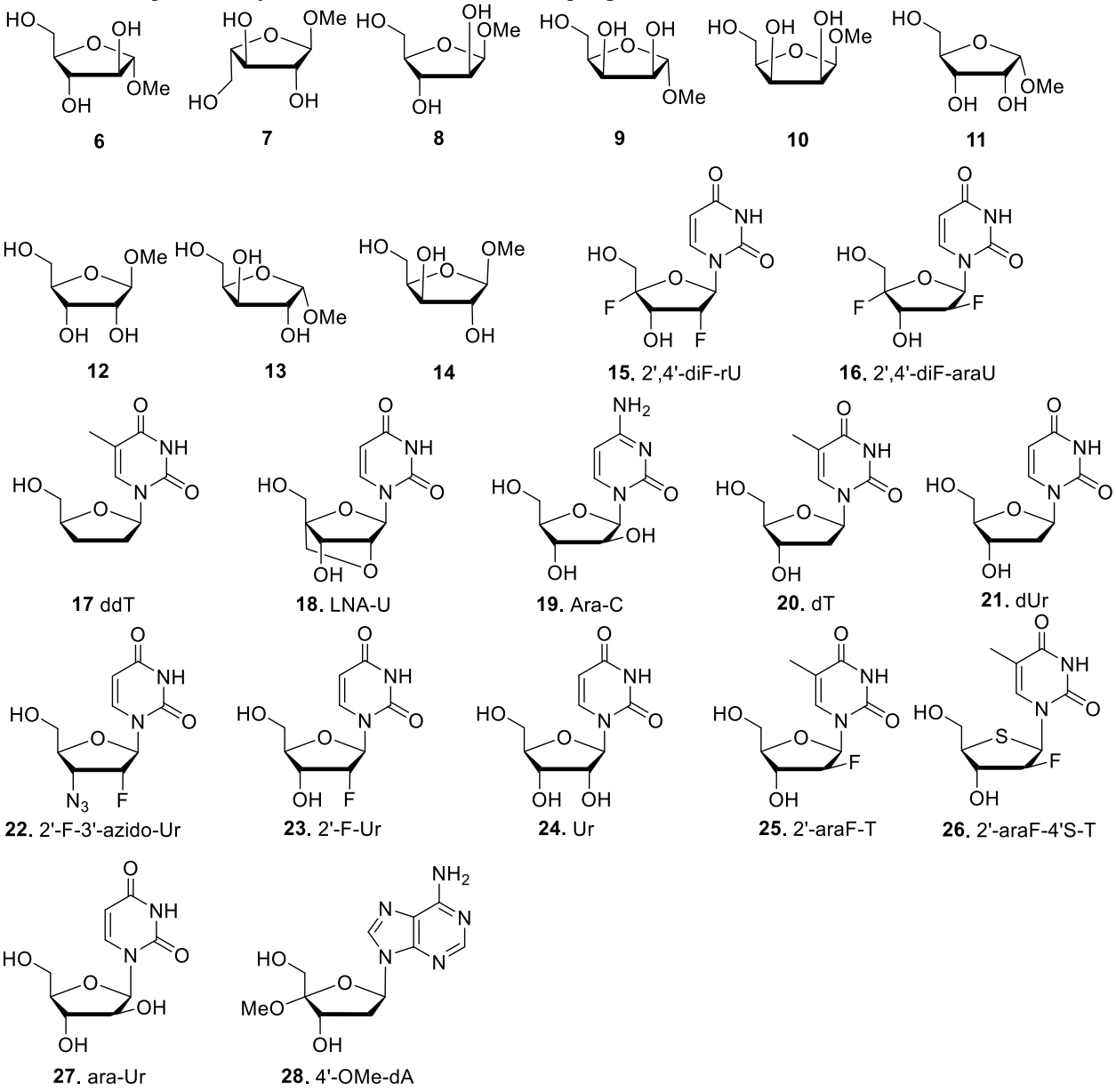
Crystal structures

Crystal structures were obtained from the CCSD with the following codes – monosaccharide **6** (CCSD112946), **9** (CCSD112951), **12** (CCSD112949) and **13** (CCSD112950). The crystal structure for monosaccharide **8** was taken from the SI of Evdokimov *et al*.[33]

Testing set

Monosaccharides **6-14** together with nucleosides and nucleosides analogues **1**, **5**, **15-28** were investigated.

Chart S1. Compounds subjected to QM/MM umbrella sampling simulations



The sugar puckering data was collected using the program PROSIT[34]. The definition of the angles in Tables S2 and S3 is the one given by PROSIT for a sugar as the one presented in Figure S3.

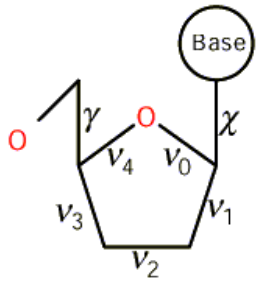


Figure S3. Angle definition presented by the PROSIT program for a furanose ring.

The pseudorotational angle P and the puckering amplitude ϕ_m are calculated by the program PROSIT according to the equations S1 and S2 developed by Altona and Sundaralingam[35]. If P lies between 0-89° and 271°-359° then it is part of the *Northern* hemisphere (encompassing the *NW*, *N* and *NE* regions) while if it lies between 91°-269° then it is part of the *Southern* hemisphere (encompassing the *SE*, *S* and *SW* regions).

$$tanP = \frac{(v_4 + v_1) - (v_3 + v_0)}{2v_2(\sin 36 + \sin 72)} \qquad eq\ (S1)$$

$$\phi\mathrm{m} = \frac{v_2}{\cos P} \qquad eq\ (S2)$$

Table S2. Angle information on the lowest energy *North* conformations for the nucleosides used in this study.

Entry	Nucleoside	V_0 (°)	V_1 (°)	V_2 (°)	V_3 (°)	V_4 (°)	γ (°)	χ (°)
1	1	14.93	-8.05	27.84	-36.52	31.78	21.72	-136.57
2	5	-28.57	2.62	21.89	-40.15	42.31	112.01	69.20
3	15	-21.63	-0.53	21.33	-34.03	35.30	114.17	-122.12
4	16	-15.04	-5.96	23.31	-32.10	29.38	113.66	42.59
5	17	-0.25	-9.59	16.34	-17.56	10.83	46.03	-112.44
6	18	-2.25	-32.50	53.43	-56.79	35.76	-33.62	-168.66
7	19	-7.41	-3.18	11.63	-16.17	15.08	120.85	35.72
8	20	-16.89	1.80	13.51	-24.13	26.00	156.95	-123.16
9	21	-8.37	-2.43	12.86	-18.32	16.35	156.58	-165.26
10	22	-17.62	-3.26	21.71	-31.66	31.57	175.19	75.60
11	23	-27.42	5.82	16.72	-34.06	38.13	-132.54	70.57
12	24	-26.35	-0.02	24.83	-41.64	41.17	-88.80	63.29
13	25	-18.47	-3.23	22.08	-33.24	31.65	123.31	67.31
14	26	-24.91	5.61	23.12	-39.95	39.09	163.86	-2.37
15	27	-11.54	-5.40	19.07	-25.38	23.35	131.15	58.72

16	28	-4.83	-15.91	31.73	-34.48	24.76	109.19	55.99

Table S3. Angle information on the lowest energy *South* conformations for the nucleosides used in this study.

Entry	Nucleoside	V_0 (°)	V_1 (°)	V_2 (°)	V_3 (°)	V_4 (°)	γ (°)	χ (°)
1	1	-12.78	32.56	-38.88	31.75	-12.48	96.23	-127.14
2	5	-15.88	29.15	-30.58	21.60	-3.87	59.28	-142.56
3	15	-19.25	34.39	-35.61	22.89	-1.53	58.64	-114.31
4	16	-23.59	24.70	-18.47	5.04	11.85	-68.32	-133.62
5	17	-11.19	26.16	-31.23	24.93	-7.56	108.37	-120.33
6	18	-	-	-	-	-	-	-
7	19	-19.68	30.46	-31.51	19.82	-0.08	151.07	52.86
8	20	-12.20	17.96	-18.28	10.66	1.22	116.75	-143.97
9	21	-15.34	22.64	-23.07	14.72	0.30	125.43	-119.38
10	22	-7.78	21.93	-27.35	22.40	-9.60	-55.30	179.87
11	23	-25.75	41.19	-42.47	26.42	0.39	95.89	-125.28
12	24	-6.76	9.49	-8.89	4.81	1.41	-110.09	-123.60
13	25	-27.74	33.13	-25.94	8.30	12.47	-97.86	-144.34
14	26	-20.94	42.86	-50.19	33.67	-7.08	161.64	-31.60
15	27	-14.69	22.22	-22.52	13.73	1.07	133.92	-145.11
16	28	-9.89	26.79	-34.23	28.98	-11.62	153.75	56.80

Table S4. Puckering information obtained following the umbrella sampling simulations

Entry	Nucleoside	P_N* (°)	ϕ_m**(°)	P_S* (°)	ϕ_m**(°)
1	1	41.3	37.0	179.5	38.9
2	5	59.3	43.0	168.3	31.2
3	15	54.0	36.3	165.1	36.9
4	16	44.5	32.7	135.5	25.7
5	17	20.7	17.5	177.1	31.3
6	18	20.8	57.1	0.0	0.0
7	19	44.8	16.4	162.7	33.0
8	20	58.9	26.1	159.8	19.5

9	21	45.7	18.4	161.7	24.3
10	22	49.3	33.3	181.6	27.4
11	23	64.0	38.1	162.6	44.5
12	24	55.0	43.3	154.8	9.8
13	25	49.7	34.1	140.8	33.5
14	26	57.0	42.5	171.5	50.8
15	27	43.1	26.1	160.7	23.9
16	28	26.3	35.4	182.1	34.3

*P = pseudorotational phase angle; **ϕ_m = puckering amplitude;

Table S5. Data obtained following the NBO analysis carried out at the M06L/def2-TZVP level of theory. (+) denotes a predominant N effect while (-) denotes a predominant S effect.

Entry	Nucleoside	$\Delta\sigma_{C3'H} - \sigma^*_{C4'R}$ (kcal/mol)	$\Delta\sigma_{C3'H} - \sigma^*_{C2'R}$ (kcal/mol)	$\Delta\sigma_{C2'H} - \sigma^*_{C3'R}$ (kcal/mol)	$\Delta\sigma_{C3'C4'} - \sigma^*_{C2'R}$ (kcal/mol)	$\Delta\sigma_{C1'2'} - \sigma^*_{C3'R}$ (kcal/mol)	$\Delta n_{O4'} - \sigma^*_{C4'R}$ (kcal/mol)
1	1	+1.8	+1.5	-4.7	-2.5	+1.8	+1.8
2	5	+2.7	+2.4	-1.8	-2.3	+0.2	+1.9
3	15	+5.2	+2.3	-3.4	-3.3	+2.2	+7.7
4	16	+2.8	-0.3	+0.3	+2.2	+1.3	+5.0
5	17	+2.1	+0.3	-1.6	+1.1	+1.2	+1.1
6	18	+2.8	+2.5	0.0	+0.1	+5.8	+7.5
7	19	+1.2	+0.8	+0.5	+1.2	+1.0	-0.6
8	20	+0.4	+0.3	-0.8	-0.3	+0.4	+3.6
9	21	+1.8	+0.7	-1.4	-0.6	+1.1	+2.2
10	22	+1.9	+2.2	-2.3	-2.1	+1.37	+1.1
11	23	+1.8	+3.1	-2.9	-2.3	+1.5	+1.5
12	24	+1.9	+2.0	-1.1	-1.0	+2.2	+1.2
13	25	+0.8	+0.0	+0.3	+4.0	+1.9	+0.3
14	26	+3.3	+1.4	+0.5	+1.8	+1.4	+2.2
15	27	+0.9	+0.2	+0.4	+1.7	+1.6	-0.1
16	28	+3.7	+0.0	-3.5	+1.3	+1.8	+0.8

VI. Intramolecular Hydrogen Bonding

The molecular orbitals depicting intramolecular hydrogen bonding (wireframe) for the *North* and *South* conformations were determined using the molecular orbitals obtained following the natural bond orbital analysis. The structures were built in Molekel, using an isosurface value of 0.01.

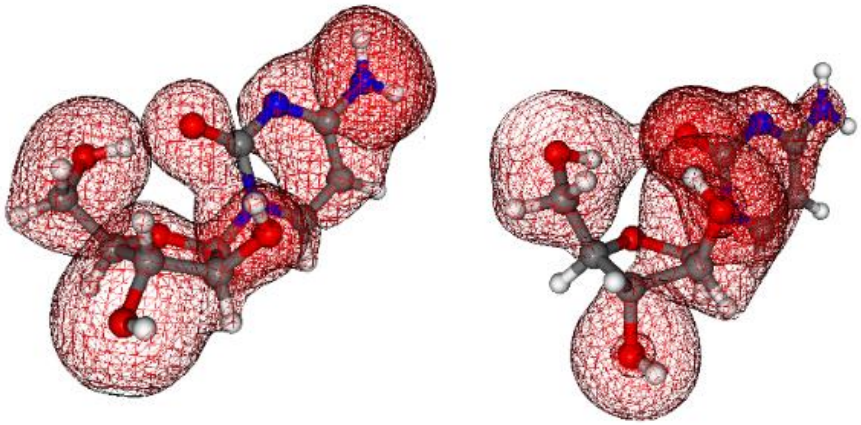


Figure S4. Intramolecular hydrogen bonds for the *North* and *South* puckers for **19.** Left – *North* conformation – hydrogen bond between C=O and C5′-OH. Right – *South* conformation – hydrogen bonds between 5′-OH···C=O and C=O···2′-OH.

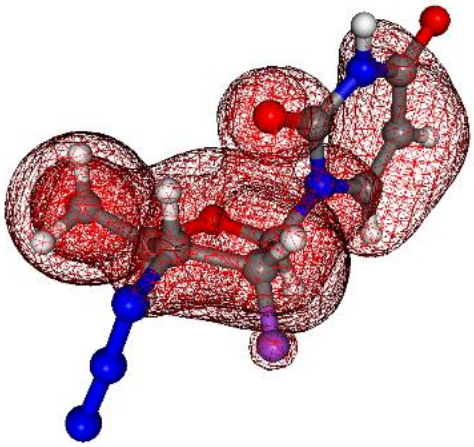


Figure S5. Intramolecular hydrogen bonds for the *North* pucker for **22.** Hydrogen bonds depicted are C5′-OH---O4′ and C=O---H2′ and C=O---H3′.

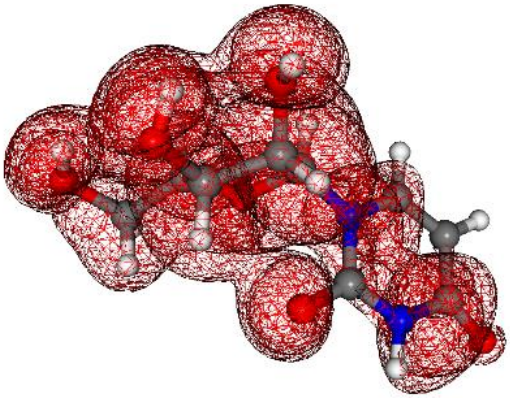


Figure S6. Intramolecular hydrogen bonds for the *North* pucker for **24.** Hydrogen bonds between 5′-OH···3′-OH, 2′-OH···3′-O and C=O···2′H.

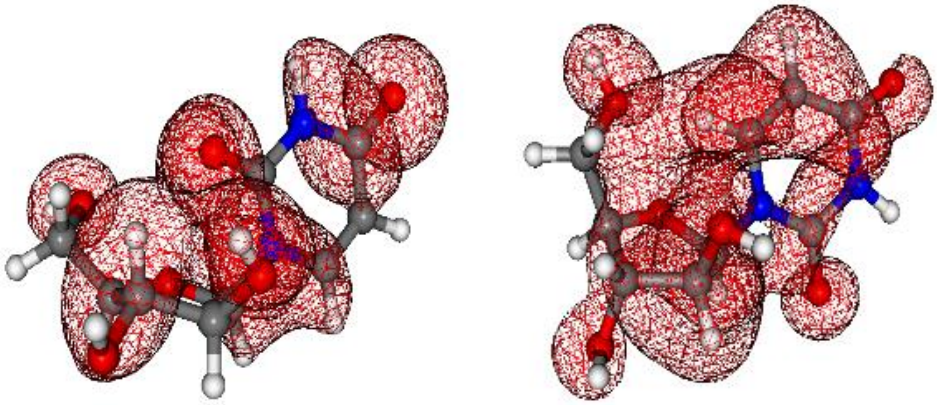


Figure S7. Intramolecular hydrogen bonds for the *North* pucker (left) and *South* pucker (right) for **27.** Left – hydrogen bonding between C=O···H2′ and C=O···2′OH. Right – hydrogen bonding between O5···H(base).

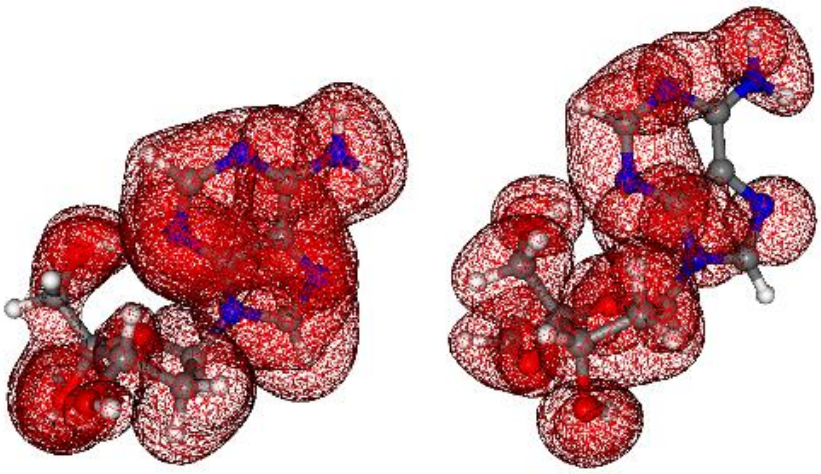


Figure S8. Intramolecular hydrogen bonds for the *North* pucker (left) and *South* pucker (right) for **28.** Left – hydrogen bonding between O5···H5-N(base). Right – hydrogen bonding between O5···H5-N(base).

VII. Coordinates for the lowest energy conformations of the nucleosides used in this study.

1 – North (DFTBESCF Energy = -5.88416 Eh; Total Energy of the System (incl. water) = -18.677 Eh)

O	16.08500	11.33200	17.04400
H	15.77000	10.85500	16.27300
C	15.07900	12.40100	17.46100
H	14.15800	11.84600	17.63900
H	15.46000	12.77700	18.41100
C	14.84400	13.52400	16.56200
O	15.72900	14.68800	16.99600
H	13.79000	13.79900	16.59400
C	15.17500	13.41200	15.05800
N	14.04000	12.73600	14.34500
N	14.53500	12.17400	13.34100
N	14.70900	11.53500	12.41900
H	16.10200	12.85200	14.93200
C	15.37700	14.89300	14.56000
H	14.43200	15.30400	14.20500
H	16.11800	14.98300	13.76500
C	15.85700	15.63700	15.82100
H	15.17100	16.46800	15.98800
N	17.24700	16.17400	15.81400
C	18.34500	15.33600	15.72000
H	18.08700	14.28800	15.73700
C	19.60000	15.87900	15.53100
C	20.78800	15.05700	15.14900

H 20.50300 14.01100 15.03300
H 21.50800 15.20600 15.95400
H 21.27700 15.32300 14.21200
C 19.75900 17.32800 15.44500
O 20.88200 17.88200 15.45800
N 18.56900 18.10000 15.55500
H 18.80300 19.07400 15.43400
C 17.31200 17.54900 15.56700
O 16.33000 18.25900 15.24700

1 – South (DFTBESCF Energy = -5.81257 Eh; Total Energy of the System (incl. water) = -18.6945 Eh)

O 18.43000 15.92400 13.83300
H 18.94700 15.64300 13.07500
C 17.95600 17.27800 13.69700
H 18.65400 18.01300 14.09800
H 17.82900 17.45800 12.62900
C 16.56100 17.35800 14.37100
O 15.90100 16.07200 14.46700
H 15.89100 17.99100 13.79000
C 16.68000 17.86800 15.78200
N 15.39500 18.54500 16.07300
N 15.38800 19.74400 16.42600
N 15.37500 20.84900 16.65200
H 17.45000 18.62500 15.93000
C 16.82800 16.58200 16.55600
H 16.71100 16.62800 17.63900
H 17.84800 16.27500 16.32800
C 15.90400 15.68400 15.86800
H 14.89300 15.82900 16.24900
N 16.20900 14.24100 16.08900
C 16.83600 13.53500 15.10100
H 17.38500 14.11600 14.37400
C 16.69500 12.19400 14.97700
C 17.53000 11.41900 14.02400
H 18.19600 12.00300 13.38900
H 16.88800 10.78500 13.41300
H 18.14900 10.67300 14.52200
C 15.86000 11.41600 15.94700
O 15.72900 10.18500 16.01200
N 15.21100 12.27000 16.86800
H 14.49900 11.94800 17.50800
C 15.30100 13.59700 16.84700
O 14.39900 14.17700 17.51700

5 – North (DFTBESCF Energy = -5.80798 Eh; Total Energy of the System (incl. water) = -36.39147 Eh)

C	20.3930	20.3120	20.2600
O	21.3810	21.3420	19.7950
C	21.8620	21.9430	21.0130
C	22.0620	20.7190	21.9640
H	22.9580	20.1930	21.6350
O	22.0870	21.2620	23.2940
H	22.5560	20.6780	23.8940
C	23.0020	22.8470	20.7960
O	23.4630	22.7020	19.3600
H	23.3900	21.7900	19.0690
H	22.8420	23.9190	20.9090
H	23.8310	22.5990	21.4590
O	20.7120	22.6900	21.6060
C	19.9960	23.4270	20.6020
H	18.9370	23.1690	20.6230
H	20.3520	23.1560	19.6080
H	20.1260	24.4810	20.8450
N	20.2630	19.2060	19.2950
C	21.3450	18.3380	19.1430
O	22.4710	18.5890	19.7190
N	21.1930	17.3000	18.2890
H	21.9690	16.6530	18.2510
C	20.0040	17.0140	17.5870
O	19.8610	15.9150	16.8940
C	18.9910	18.0460	17.5830
C	19.1180	19.0910	18.4860
H	18.4210	19.8930	18.6790
H	18.1490	18.0650	16.9070
H	19.4360	20.8340	20.2460
C	20.8320	19.8620	21.6370
F	19.8160	20.0910	22.5300
H	21.0410	18.7970	21.7380

5 – South (DFTBESCF Energy = -5.75236 Eh; Total Energy of the System (incl. water) = -36.06479 Eh)

C	20.3860	20.3050	20.2530
O	21.8080	20.4350	20.5050
C	22.5530	19.5730	19.6140
C	21.5880	18.7600	18.7890
H	21.9760	18.7560	17.7710
O	21.3640	17.4170	19.1540
H	20.7580	17.3510	19.8960

C 23.6160 20.3390 18.8400
O 22.8170 21.2340 17.9520
H 23.4600 21.4940 17.2880
H 24.3110 20.9450 19.4220
H 24.0800 19.5570 18.2400
O 23.2770 18.6140 20.5480
C 23.7240 19.1500 21.8130
H 24.3100 20.0540 21.6440
H 24.2320 18.3140 22.2940
H 22.8450 19.4130 22.4020
N 19.7970 21.6350 20.1910
C 18.5540 21.7650 20.7340
O 17.9400 20.7550 21.1620
N 18.0850 23.0850 20.7310
H 17.2290 23.2610 21.2370
C 18.8150 24.1960 20.4930
O 18.2190 25.2900 20.7200
C 20.1140 24.0160 19.8970
C 20.5440 22.7500 19.7480
H 21.5210 22.4040 19.4430
H 20.7430 24.8120 19.5280
H 19.8810 19.6870 20.9950
C 20.3080 19.6150 18.8810
F 19.0690 18.8300 18.8770
H 20.2330 20.3470 18.0770

15– North (DFTBESCF Energy = -5.13582 Eh; Total Energy of the System (incl. water) = -17.78459 Eh)

O 18.17100 15.08800 16.81700
C 17.45800 16.02900 17.23900
N 17.76200 16.76000 18.43200
H 18.59200 16.44400 18.91300
C 17.14800 17.88800 18.82200
O 17.56000 18.47200 19.87500
C 15.99100 18.32600 17.95600
H 15.63800 19.32900 18.14500
C 15.66400 17.59100 16.86100
H 14.82900 17.73100 16.19100
N 16.26700 16.38900 16.63800
C 15.71000 15.50000 15.63700
O 15.56500 16.28700 14.37300
H 16.32200 14.62400 15.42300
C 14.25700 14.98900 15.98000
F 14.24300 13.58200 16.06900
H 13.97900 15.39600 16.95300
C 13.32300 15.54500 14.85600

O 12.25000 14.76300 14.47700
H 12.50000 13.95500 14.02200
H 12.92800 16.50600 15.18500
C 14.29600 15.84400 13.67800
F 14.73800 14.54100 13.14000
C 13.89400 16.88100 12.61400
H 14.15700 16.41300 11.66600
H 12.84500 17.14900 12.74200
O 14.85600 17.98500 12.76900
H 14.53100 18.70900 12.22900

15 – South (DFTBESCF Energy = -5.08217 Eh; Total Energy of the System (incl. water) = -17.50093 Eh)

O 15.15100 12.82600 15.46600
C 15.48000 13.26400 16.60200
N 15.58300 12.39500 17.62500
H 15.36400 11.43000 17.42400
C 15.78700 12.84500 18.90300
O 15.78800 12.01500 19.85600
C 16.04200 14.24000 19.06300
H 16.31200 14.57600 20.05400
C 16.05900 15.10800 18.04300
H 16.35000 16.13100 18.22700
N 15.67800 14.62700 16.81000
C 15.76000 15.54900 15.68700
O 17.14400 16.12000 15.63500
H 15.63100 14.90900 14.81400
C 14.91800 16.88700 15.90500
F 13.57800 16.68300 15.49000
H 14.96000 17.05900 16.98000
C 15.62400 17.99000 15.14300
O 15.34700 17.95500 13.71900
H 16.01400 18.46500 13.25500
H 15.58400 18.99700 15.55800
C 17.11700 17.60200 15.33200
F 17.76600 17.70400 14.12800
C 17.93800 18.33600 16.47800
H 18.96000 17.96600 16.55600
H 17.99600 19.39100 16.20800
O 17.20700 18.03100 17.69000
H 17.77200 18.20500 18.44700

16 – North (DFTBESCF Energy = -5.15772 Eh; Total Energy of the System (incl. water) = -16.76468 Eh)

O 15.53300 17.81300 14.71900

C 15.49300 17.04200 13.78800
N 15.57500 17.47700 12.44300
H 15.85200 18.44500 12.36000
C 15.59200 16.66600 11.31800
O 15.72400 17.13400 10.15800
C 15.49600 15.20600 11.64200
H 15.79700 14.47000 10.91100
C 15.45000 14.81400 12.93300
H 15.53400 13.74500 13.06600
N 15.43300 15.68300 13.98200
C 15.32100 15.09000 15.29800
O 14.35900 15.79900 16.17300
H 15.01700 14.04500 15.22700
C 16.66300 15.12600 16.05100
F 17.59400 15.61000 15.07100
H 17.06700 14.14200 16.28800
C 16.41200 16.00200 17.27000
O 17.03300 15.44900 18.39200
H 16.89500 15.99000 19.17300
H 16.72100 17.04200 17.16500
C 14.89800 15.94300 17.50900
F 14.57800 14.78300 18.34400
C 14.20600 17.16300 18.10900
H 13.66800 16.83300 18.99700
H 14.92900 17.88100 18.49500
O 13.31100 17.73900 17.13800
H 12.50400 17.95700 17.61100

16– South (DFTBESCF Energy = -5.16228 Eh; Total Energy of the System (incl. water) = -16.78699 Eh)

O 17.49000 13.42200 15.10500
C 16.38900 13.03500 14.62400
N 16.20900 11.90100 13.90700
H 17.06500 11.38100 13.77900
C 15.08300 11.37800 13.35100
O 15.12200 10.32300 12.69300
C 13.88200 12.10300 13.63900
H 12.87900 11.80300 13.37400
C 13.93500 13.23400 14.37800
H 13.15100 13.83200 14.81800
N 15.15400 13.70800 14.78800
C 15.31000 15.07900 15.28700
O 14.29100 15.12800 16.35500
H 16.31000 15.15300 15.71400
C 14.96100 16.27600 14.35400
F 14.07900 15.93500 13.28100

H 15.80200 16.66100 13.77900
C 14.21000 17.23400 15.27400
O 15.04700 18.23800 15.82300
H 15.61400 18.57200 15.12400
H 13.44200 17.77200 14.71800
C 13.71800 16.46100 16.51800
F 14.32100 17.01800 17.64800
C 12.22200 16.23600 16.75900
H 11.70600 15.80400 15.90200
H 12.11400 15.52800 17.58000
O 11.67700 17.54700 17.05200
H 11.03200 17.52700 17.76200

17 – North (DFTBESCF Energy = -5.42163 Eh; Total Energy of the System (incl. water) = -18.379 Eh)

O 14.96100 14.22400 13.50100
C 14.85300 13.72400 14.63800
N 14.45500 12.32200 14.73100
H 14.11100 11.88500 13.88800
C 14.44300 11.64900 15.97200
O 13.94600 10.51200 16.06700
C 15.20300 12.22900 17.01900
C 15.75600 11.49200 18.21000
H 16.44400 12.12300 18.77200
H 16.34900 10.68200 17.78500
H 14.90800 11.14400 18.80000
C 15.59100 13.57900 16.90400
H 16.26000 14.03700 17.61800
N 15.20100 14.32700 15.78500
C 15.82000 15.72100 15.77300
O 17.18000 15.49200 16.22500
H 15.79700 16.06300 14.73800
C 15.05500 16.71400 16.77100
H 14.78900 17.68500 16.35400
H 14.06400 16.34900 17.03800
C 16.05900 16.80000 17.94200
H 15.57700 16.18300 18.70000
H 16.15700 17.82000 18.31400
C 17.37500 16.30800 17.50800
H 18.10900 17.08900 17.31500
C 18.19700 15.48700 18.47300
H 18.70500 16.05100 19.25600
H 19.06400 15.03200 17.99500
O 17.35500 14.52000 19.05900
H 17.65500 14.05200 19.84200

17 – South (DFTBESCF Energy = -5.41879 Eh; Total Energy of the System (incl. water) = -18.30092 Eh)

O 14.44400 14.12800 17.56100
C 15.67700 14.07400 17.61600
N 16.30800 13.38500 18.68000
H 15.71500 12.88900 19.33000
C 17.67400 13.12900 18.75600
O 18.15500 12.45100 19.71300
C 18.45700 13.75400 17.72500
C 19.94200 13.44400 17.74600
H 20.07600 12.40000 18.02700
H 20.48300 14.04700 18.47500
H 20.28200 13.54600 16.71500
C 17.85500 14.60400 16.79000
H 18.43200 14.97500 15.95600
N 16.51300 14.76800 16.72300
C 15.85600 15.75700 15.80900
O 16.13300 15.33100 14.42500
H 14.78100 15.73200 15.98400
C 16.45200 17.22900 15.91600
H 15.80600 17.90200 16.48100
H 17.46900 17.16600 16.30200
C 16.40500 17.72300 14.52000
H 17.09800 18.55700 14.40600
H 15.40200 18.12900 14.39000
C 16.59700 16.50800 13.60700
H 15.90700 16.56400 12.76600
C 17.99300 16.36000 13.11700
H 18.55400 17.23600 13.44300
H 18.04800 16.41500 12.02900
O 18.59100 15.21000 13.79800
H 17.81500 14.73500 14.10500

18 – North (DFTBESCF Energy = -5.50680 Eh; Total Energy of the System (incl. water) = -17.47384 Eh)

O 13.37600 15.57900 10.84400
C 14.42100 14.90000 10.84700
N 14.87700 14.37900 9.69000
H 14.32100 14.45300 8.85000
C 16.17000 13.81900 9.57800
O 16.53500 13.22800 8.53600
C 17.07000 14.13700 10.69200
H 18.13100 13.94800 10.62000
C 16.59100 14.73800 11.82900

H 17.18800 14.83400 12.72400
N 15.27500 15.08500 11.96800
C 14.72700 15.58300 13.15200
O 15.74800 15.97500 14.13800
H 14.12200 16.44700 12.87900
C 13.82400 14.56400 13.90800
C 14.86600 13.96700 14.77800
O 14.37200 12.93800 15.76200
H 14.94800 13.00500 16.52800
H 15.70300 13.57700 14.19900
H 13.24200 13.95600 13.21700
O 13.05100 15.33400 14.86800
C 13.99500 15.81200 15.86100
H 13.87300 16.89300 15.79100
H 13.79800 15.45000 16.87000
C 15.38200 15.25500 15.36800
C 16.60500 15.21700 16.24800
H 16.49900 15.65300 17.24200
H 16.85800 14.18400 16.48800
O 17.64100 15.84100 15.56300
H 17.93200 15.30000 14.82600

19 – North (DFTBESCF Energy = -5.38381 Eh; Total Energy of the System (incl. water) = -18.17343 Eh)

O 15.07600 17.95900 17.19100
C 16.07500 17.94800 16.39100
N 16.82200 19.06200 16.27300
C 18.01900 18.97600 15.71700
N 18.82400 20.07300 15.74200
H 18.67900 20.88900 16.32000
H 19.63200 20.13500 15.13900
C 18.45100 17.86600 14.88400
H 19.37100 17.85300 14.31800
C 17.66300 16.73800 15.00800
H 17.82000 15.84900 14.41500
N 16.52600 16.76600 15.70500
C 15.75700 15.54600 15.68400
O 14.36600 15.82400 15.63000
H 16.00000 14.98000 14.78500
C 15.98700 14.68700 17.03200
O 17.01100 15.39100 17.80500
H 16.74000 15.70800 18.66900
H 16.36100 13.69300 16.78600
C 14.60500 14.62200 17.70100
O 14.32400 13.23200 17.92700
H 14.89100 13.00400 18.66700

H 14.55500 15.14700 18.65500
C 13.59800 15.13500 16.65300
H 13.06700 14.35600 16.10600
C 12.43800 16.01200 17.21700
H 12.59600 16.27500 18.26300
H 11.46800 15.55500 17.02100
O 12.57000 17.29400 16.51800
H 13.48800 17.56200 16.59700

19 – South (DFTBESCF Energy = -5.34819 Eh; Total Energy of the System (incl. water) = -18.01247 Eh)

O 17.38800 17.61700 14.64800
C 16.56300 17.12100 13.82700
N 16.61500 17.38800 12.49900
C 15.70600 16.86800 11.60000
N 15.70100 17.31200 10.30800
H 16.29100 18.03000 9.91300
H 15.15500 16.74500 9.67400
C 14.77400 15.84500 12.04400
H 14.07000 15.38100 11.36900
C 14.77600 15.56500 13.39100
H 14.18500 14.82200 13.90500
N 15.70500 16.08800 14.26200
C 15.70600 15.49700 15.63300
O 15.57500 16.72000 16.59900
H 14.86400 14.81600 15.76000
C 17.03800 14.84000 16.13300
O 18.16300 15.34300 15.37300
H 18.14300 16.30100 15.43900
H 16.88200 13.76200 16.09200
C 17.07300 15.18600 17.67600
O 16.26700 14.11100 18.27800
H 16.84900 13.35300 18.18400
H 18.06300 15.13100 18.12900
C 16.37400 16.51800 17.77100
H 15.74400 16.54300 18.66000
C 17.44300 17.58700 17.90600
H 18.41700 17.25300 17.54900
H 17.67600 17.77300 18.95500
O 17.03000 18.87100 17.35400
H 16.51400 18.61600 16.58500

20– North (DFTBESCF Energy = -5.57262 Eh; Total Energy of the System (incl. water) = -18.21008 Eh)

O 14.80900 16.25600 18.25800

C 15.81500 16.84800 17.76300
N 16.52100 17.80000 18.53100
H 16.26100 17.96000 19.49300
C 17.77500 18.36900 18.20200
O 18.43400 18.90900 19.05800
C 18.26500 18.08700 16.85800
C 19.60200 18.69700 16.43200
H 19.78800 18.52500 15.37200
H 20.41400 18.24300 16.99900
H 19.51000 19.75800 16.66400
C 17.52900 17.22500 16.05700
H 17.90100 16.87600 15.10500
N 16.33200 16.66900 16.51400
C 15.76400 15.55300 15.69000
O 16.84200 14.64500 15.28900
H 15.15500 15.03600 16.43200
C 15.10000 16.03000 14.40100
H 14.05100 15.75500 14.29300
H 15.21600 17.11400 14.37000
C 15.83000 15.30900 13.26900
H 16.29600 16.04300 12.61200
O 14.86900 14.47200 12.43100
H 14.07200 14.97200 12.24100
C 16.61300 14.24000 13.91100
H 16.03800 13.31500 13.88200
C 17.96500 13.99100 13.34300
H 18.45000 14.90200 12.99300
H 17.82300 13.44800 12.40800
O 18.82000 13.35600 14.31700
H 18.30800 12.70100 14.79700

20– South (DFTBESCF Energy = -5.58651 Eh; Total Energy of the System (incl. water) = -18.26107 Eh)

O 13.38000 16.62400 16.73900
C 14.50400 16.90000 17.32600
N 14.51300 17.73800 18.40500
H 13.60700 18.08300 18.68800
C 15.59500 17.92100 19.16900
O 15.44600 18.61300 20.20600
C 16.87000 17.34300 18.69400
C 18.10000 17.76900 19.40500
H 18.80300 18.36100 18.81800
H 18.66600 16.86400 19.62300
H 17.92300 18.28800 20.34600
C 16.91900 16.58600 17.56700
H 17.78200 16.09700 17.14000

N 15.75900 16.38800 16.86100
C 15.74200 15.53100 15.66900
O 16.76200 14.50200 15.85000
H 14.71400 15.17200 15.61600
C 16.21100 16.30500 14.42200
H 15.39300 16.47000 13.72100
H 16.51900 17.30500 14.72900
C 17.21700 15.39500 13.71000
H 18.08400 15.99500 13.43400
O 16.77100 14.85000 12.40100
H 16.06600 15.46100 12.17500
C 17.70400 14.34100 14.67500
H 17.65100 13.29300 14.38100
C 19.18700 14.54800 15.22400
H 19.70800 15.37700 14.74500
H 19.76100 13.69100 14.86900
O 19.32400 14.76400 16.64700
H 19.05300 13.92300 17.02300

21– North (DFTBESCF Energy = -5.13391 Eh; Total Energy of the System (incl. water) = -16.57345 Eh)

O 17.65700 13.80000 14.23200
C 17.44000 14.98300 13.84100
N 18.24100 15.62700 12.95700
H 19.15500 15.22400 12.80700
C 18.04100 16.97900 12.50700
O 18.86600 17.38900 11.64300
C 16.91200 17.63300 13.10400
H 16.84400 18.70500 12.99000
C 16.07300 16.97600 13.97200
H 15.19800 17.42600 14.41800
N 16.27000 15.66500 14.30300
C 15.35600 15.03700 15.22200
O 14.10900 15.86100 15.19300
H 14.99900 14.06200 14.89200
C 15.91000 14.97200 16.67200
H 15.88200 13.96300 17.08100
H 16.88400 15.43300 16.84000
C 14.85600 15.91200 17.42100
H 15.37800 16.84300 17.64100
O 14.46100 15.24600 18.69000
H 13.65800 14.74800 18.52200
C 13.69200 16.18100 16.58000
H 12.91400 15.46400 16.84300
C 13.05500 17.55200 16.64800
H 13.89200 18.24700 16.71800

H 12.34300 17.57600 17.47200
O 12.34100 17.77000 15.40700
H 12.40500 18.72400 15.32600

21– South (DFTBESCF Energy = -5.11368 Eh; Total Energy of the System (incl. water) = -16.58301 Eh)

O 13.56100 16.69700 13.90300
C 14.59100 16.32100 13.26500
N 14.89400 16.77800 12.02900
H 14.33000 17.58300 11.79800
C 16.04400 16.38400 11.30600
O 16.26500 16.97700 10.17500
C 16.90500 15.40700 11.99200
H 17.71200 14.95700 11.43300
C 16.60700 15.01700 13.24200
H 17.24400 14.44200 13.89900
N 15.52300 15.47900 13.89600
C 15.34800 15.02900 15.21400
O 16.54100 15.53900 16.04400
H 14.52800 15.60900 15.63800
C 15.27000 13.49500 15.36900
H 14.20800 13.25900 15.42800
H 15.65900 12.94700 14.51100
C 16.03800 13.29300 16.72700
H 16.60200 12.36500 16.63600
O 15.13600 13.17700 17.85200
H 15.28800 12.29000 18.18600
C 16.96100 14.46400 16.98100
H 16.74300 14.91400 17.94900
C 18.49200 14.23700 16.75800
H 18.61700 13.22100 16.38400
H 18.96700 14.30300 17.73700
O 18.99400 15.23800 15.83300
H 19.91500 14.96700 15.83600

22 – North (DFTBESCF Energy = -5.35624 Eh; Total Energy of the System (incl. water) = -18.41565 Eh)

O 16.25900 13.01200 15.30800
C 15.29000 13.35700 15.97000
N 14.45700 12.39800 16.52600
H 14.74600 11.44300 16.37100
C 13.23400 12.67900 17.13200
O 12.41400 11.78100 17.48600
C 13.01500 14.09800 17.43500
H 12.16400 14.37400 18.04100

C 13.94800 15.03900 17.16300
H 13.90900 16.06200 17.50900
N 15.04800 14.69500 16.34500
C 15.94200 15.72100 15.90600
O 15.56400 15.87900 14.50400
H 15.82400 16.67300 16.42400
C 17.45800 15.34000 15.92500
F 18.22100 16.30100 16.57300
H 17.70400 14.39600 16.41200
C 17.88900 15.21100 14.40900
N 19.31000 15.58000 14.21500
N 19.59600 16.77000 14.08100
N 19.99500 17.83600 13.92200
H 17.83500 14.16000 14.12500
C 16.81100 16.05100 13.65700
H 17.05200 17.10500 13.79500
C 16.64700 15.81700 12.20000
H 16.57000 14.76100 11.94000
H 17.56800 16.24000 11.79800
O 15.55500 16.56600 11.64500
H 15.10800 17.00400 12.37300

22 – South (DFTBESCF Energy = -5.37803 Eh; Total Energy of the System (incl. water) = -18.4539 Eh)

O 18.27000 16.62600 17.15600
C 17.55800 15.80300 17.83500
N 17.92000 15.48100 19.09800
H 18.79100 15.84900 19.45100
C 17.30900 14.50200 19.83400
O 17.73800 14.20000 20.98500
C 16.18600 13.84800 19.20500
H 15.73200 12.98700 19.67100
C 15.75100 14.29500 17.98000
H 14.79700 13.98900 17.57600
N 16.44200 15.29700 17.29300
C 15.95300 15.73200 15.91700
O 14.77900 15.01500 15.66000
H 15.78200 16.80700 15.85100
C 16.95200 15.29300 14.81000
F 17.74300 16.36600 14.49000
H 17.60400 14.47100 15.10600
C 16.07500 14.74500 13.65000
N 15.71800 15.83400 12.76500
N 16.27800 15.78100 11.63600
N 16.85600 15.87500 10.65100
H 16.48100 13.87600 13.13200

C 14.81100 14.25000 14.35600
H 13.91500 14.52900 13.80100
C 14.78200 12.80100 14.81300
H 13.79500 12.55300 15.20400
H 15.54900 12.70400 15.58100
O 15.04900 11.93500 13.71700
H 14.64400 11.11200 13.99700

23– North (DFTBESCF Energy = -5.10235 Eh; Total Energy of the System (incl. water) = -17.10571 Eh)

O 15.23300 17.40800 17.48300
C 15.72100 16.27200 17.64100
N 16.22800 15.97300 18.88100
H 16.05900 16.63900 19.62100
C 16.94900 14.87200 19.28800
O 17.25400 14.76800 20.48700
C 17.13000 13.84300 18.18500
H 17.90800 13.10600 18.31800
C 16.73200 14.18300 16.93600
H 17.07500 13.60500 16.09100
N 15.89300 15.24800 16.69700
C 15.37200 15.49600 15.35800
O 16.28000 16.54800 14.75200
H 15.55300 14.59500 14.77100
C 13.92500 16.03300 15.15600
F 13.16200 15.03600 14.48600
H 13.29200 16.06700 16.04300
C 14.04800 17.29000 14.31900
O 13.01700 17.44900 13.37400
H 12.35800 16.78600 13.59300
H 14.05800 18.17600 14.95400
C 15.47000 17.20600 13.70800
H 15.40400 16.60500 12.80100
C 16.03100 18.57400 13.49300
H 16.83700 18.64900 14.22300
H 15.24400 19.30400 13.68200
O 16.60000 18.72000 12.18700
H 17.01500 19.58500 12.13800

23 – South (DFTBESCF Energy = -5.14449 Eh; Total Energy of the System (incl. water) = -17.08818 Eh)

O 12.95600 16.62500 15.10400
C 12.97400 15.52700 15.68400
N 11.80900 14.85600 16.17100
H 10.92500 15.32600 16.30500

C 11.83500 13.59300 16.69800
O 10.76300 13.15700 17.13500
C 13.12600 12.96500 16.87400
H 13.15600 11.97700 17.30900
C 14.27100 13.55700 16.40900
H 15.21000 13.02300 16.38800
N 14.13300 14.85200 15.84900
C 15.36700 15.49000 15.35200
O 16.04800 14.53600 14.41800
H 15.07400 16.42600 14.87600
C 16.59600 15.87500 16.24100
F 16.42800 17.12300 16.82100
H 16.74900 15.00500 16.88000
C 17.70600 16.01800 15.22100
O 17.50600 17.24800 14.48800
H 18.32700 17.29100 13.99200
H 18.63100 16.00600 15.79600
C 17.47600 14.84200 14.31700
H 17.68600 15.10500 13.28100
C 18.21000 13.55500 14.81800
H 19.08600 13.86600 15.38700
H 18.54300 13.03200 13.92100
O 17.33900 12.74000 15.57400
H 17.72900 11.98500 16.02200

24 – North (DFTBESCF Energy = -5.25459 Eh; Total Energy of the System (incl. water) = -17.07703 Eh)

O 14.98500 17.55600 17.02500
C 14.26700 17.42700 16.01200
N 13.39400 18.36600 15.70100
H 13.35400 19.11700 16.37400
C 12.44800 18.30500 14.67100
O 11.57600 19.16900 14.64900
C 12.56800 17.13600 13.82600
H 11.89200 17.06600 12.98600
C 13.47900 16.16300 14.12000
H 13.51200 15.34400 13.41600
N 14.35100 16.25700 15.12600
C 15.32900 15.20200 15.31500
O 14.99000 14.61300 16.62900
H 15.22300 14.38300 14.60500
C 16.84500 15.69300 15.41300
O 17.55100 14.88300 14.51900
H 18.31900 15.37300 14.21700
H 16.95600 16.75500 15.19300
C 17.17800 15.29700 16.84100

O 18.60400 14.94900 16.97600
H 18.91000 14.52900 16.16900
H 17.01100 16.10300 17.55600
C 16.23300 14.14900 17.20700
H 16.61300 13.24200 16.73600
C 16.08000 13.85000 18.70600
H 15.10500 13.43900 18.96900
H 16.01400 14.79100 19.25200
O 17.07700 12.89300 19.17400
H 17.88600 13.01000 18.67100

24 – South (DFTBESCF Energy = -5.21783 Eh; Total Energy of the System (incl. water) = -17.0085 Eh)

O 16.32900 13.75700 17.37300
C 15.15700 13.72000 17.22400
N 14.31700 12.91100 17.99200
H 14.71800 12.42500 18.78200
C 13.00100 12.67100 17.72100
O 12.44200 11.66100 18.21000
C 12.37100 13.55200 16.78700
H 11.29600 13.55900 16.68600
C 13.09500 14.54800 16.24000
H 12.61900 15.33800 15.67700
N 14.48500 14.56800 16.30700
C 15.33200 15.20500 15.31800
O 15.13700 16.71500 15.41900
H 16.35500 14.98800 15.62600
C 14.98400 14.91900 13.76900
O 16.13900 14.10900 13.23600
H 16.58900 14.73700 12.66600
H 14.05800 14.35300 13.66800
C 14.85100 16.30400 13.02200
O 15.88800 16.49100 12.10600
H 16.66800 16.46800 12.66500
H 13.93000 16.28400 12.43900
C 14.88700 17.33600 14.13500
H 15.66300 18.04500 13.84700
C 13.62000 18.21700 14.34100
H 13.26500 18.29000 15.36800
H 12.74900 17.96200 13.73600
O 14.01600 19.55000 13.91300
H 14.68100 19.85500 14.53400

25 – North (DFTBESCF Energy = -5.62155 Eh; Total Energy of the System (incl. water) = -18.58298 Eh)

O 16.19000 15.41900 13.05400
C 17.07300 14.93900 13.88400
N 18.16200 14.29300 13.40300
H 18.20600 14.05200 12.42300
C 19.30000 13.90600 14.17900
O 20.27000 13.38500 13.65300
C 19.24900 14.23300 15.60900
C 20.38400 13.54000 16.45900
H 21.25200 14.18300 16.60200
H 20.68300 12.65400 15.89900
H 19.82300 13.16900 17.31700
C 18.02400 14.65200 16.07400
H 17.93000 14.85900 17.13000
N 17.00200 15.08800 15.23000
C 15.87800 15.77900 15.83100
O 14.73900 14.88800 15.56800
H 16.02600 15.91100 16.90300
C 15.62100 17.15100 15.19500
H 15.53900 17.92000 15.96300
F 16.74400 17.55300 14.40600
C 14.31900 17.06500 14.45700
O 13.56300 18.25900 14.57100
H 14.02200 18.87300 13.99300
H 14.46000 16.74200 13.42600
C 13.59300 15.82300 15.18600
H 13.22800 16.29100 16.10000
C 12.42600 15.27500 14.35500
H 12.38900 15.79300 13.39700
H 11.50400 15.40600 14.92200
O 12.56300 13.87000 14.05800
H 12.54400 13.44500 14.91800

25 – South (DFTBESCF Energy = -5.60983 Eh; Total Energy of the System (incl. water) = -18.58298 Eh)

O 13.57600 15.83700 14.15700
C 14.70100 15.95900 13.63000
N 14.83100 15.89400 12.24500
H 14.03500 15.89000 11.62300
C 16.05000 15.93500 11.60000
O 16.09400 15.56500 10.39400
C 17.12600 16.45300 12.45100
C 18.56500 16.62100 11.85700

H 18.52300 16.84700 10.79100
H 18.95600 15.60900 11.96600
H 19.05600 17.43600 12.38800
C 17.05300 16.46200 13.78300
H 17.80800 16.83800 14.45800
N 15.86300 16.11800 14.40500
C 15.82800 15.72900 15.78100
O 16.66500 16.67700 16.53700
H 14.80600 15.73200 16.15800
C 16.47900 14.41800 16.08300
H 15.82300 13.55300 15.98700
F 17.50800 14.24100 15.13300
C 17.01500 14.50600 17.47400
O 16.05200 14.05600 18.47200
H 15.24200 14.54900 18.32200
H 17.88400 13.85600 17.57100
C 17.24600 15.96000 17.68900
H 16.67300 16.36400 18.52400
C 18.71800 16.36300 17.76900
H 18.78700 17.34500 17.30100
H 19.38300 15.67500 17.24700
O 19.05300 16.45700 19.10100
H 18.89300 15.58500 19.47000

26– North (DFTBESCF Energy = -5.63511 Eh; Total Energy of the System (incl. water) = -18.33279 Eh)

O 13.71700 16.65700 17.47400
C 14.80300 17.25700 17.43400
N 14.96100 18.48900 18.08700
H 14.09700 18.70100 18.56600
C 16.12600 19.23200 18.18400
O 16.25000 20.13100 19.04500
C 17.11700 18.96000 17.18300
C 18.28900 19.89300 17.07400
H 18.20800 20.60200 17.89700
H 18.19300 20.44400 16.13800
H 19.20400 19.30100 17.06000
C 16.89700 17.82600 16.45700
H 17.63000 17.58000 15.70200
N 15.86000 16.93300 16.61200
C 15.80800 15.70900 15.74300
S 14.24700 14.67800 15.94600
H 15.85200 16.13800 14.74100
C 16.97600 14.66500 15.98600
H 17.63900 14.84100 15.13900
F 17.62200 15.10300 17.16700

C 16.49000 13.18000 16.15900
O 17.39100 12.19800 15.66000
H 17.99900 12.50300 14.98300
H 16.24100 13.04900 17.21300
C 15.13300 13.05900 15.49200
H 15.19500 12.84300 14.42500
C 14.09500 11.99100 16.02700
H 14.12700 11.84000 17.10600
H 14.38200 11.05000 15.55900
O 12.84400 12.38500 15.47400
H 12.73800 12.08800 14.56700

26– South (DFTBESCF Energy = -5.61948 Eh; Total Energy of the System (incl. water) = -18.59732 Eh)

O 14.52400 13.10900 15.97100
C 14.28800 13.97800 16.84000
N 13.37400 13.72600 17.85500
H 13.08800 12.76100 17.94200
C 12.85100 14.68600 18.68800
O 11.90900 14.36400 19.49100
C 13.23200 16.10100 18.41400
C 12.78700 17.15900 19.35600
H 11.72300 17.00700 19.53300
H 13.06500 18.14900 18.99400
H 13.16500 17.09700 20.37600
C 14.26000 16.33300 17.51300
H 14.67800 17.32100 17.38500
N 14.77600 15.29100 16.76900
C 15.77200 15.67400 15.70700
S 16.83500 14.33000 15.26900
H 15.09200 15.88800 14.88300
C 16.68000 16.80000 16.15700
H 16.23200 17.78800 16.25800
F 17.08900 16.50800 17.47500
C 17.80800 16.79300 15.13900
O 17.19600 17.21600 13.91000
H 17.68000 17.78900 13.31100
H 18.65300 17.42100 15.42100
C 18.32900 15.36500 15.01800
H 18.84400 15.18000 14.07500
C 19.26500 15.11900 16.13300
H 18.86300 15.51900 17.06400
H 20.22500 15.61300 15.98100
O 19.40200 13.69600 16.26400
H 19.79100 13.25100 15.50700

27– North (DFTBESCF Energy = -5.29563 Eh; Total Energy of the System (incl. water) = -16.86349 Eh)

O 16.98500 16.20000 17.32400
C 16.18700 16.97500 16.77300
N 16.17500 18.30200 17.09000
H 16.71500 18.66400 17.86200
C 15.40200 19.28300 16.41900
O 15.44600 20.49300 16.74100
C 14.58100 18.75400 15.39100
H 13.98100 19.48800 14.87400
C 14.67900 17.47700 14.99600
H 14.16100 17.11400 14.12100
N 15.34600 16.57400 15.75700
C 15.32400 15.19700 15.31000
O 14.96800 14.38000 16.50500
H 14.49700 15.06100 14.61300
C 16.68700 14.58900 14.81500
O 17.63000 15.61600 14.64200
H 18.47600 15.45200 15.06400
H 16.42000 14.00200 13.93600
C 17.11400 13.53100 15.84300
O 17.72300 12.39700 15.11600
H 18.67100 12.48300 15.23900
H 17.90300 13.94200 16.47300
C 15.81700 13.17100 16.63500
H 15.38500 12.35200 16.05900
C 15.97400 12.82600 18.11900
H 16.98600 13.13200 18.38600
H 15.83400 11.75200 18.24000
O 14.98800 13.59500 18.92500
H 14.44600 12.99700 19.44500

27– South (DFTBESCF Energy = -5.27233 Eh; Total Energy of the System (incl. water) = -16.94157 Eh)

O 13.72300 13.73900 16.99300
C 13.33800 14.89100 16.62300
N 12.06300 15.38600 16.98800
H 11.66700 14.65400 17.56100
C 11.73400 16.73700 17.04800
O 10.68200 17.07700 17.66600
C 12.66400 17.55800 16.27500
H 12.32000 18.55700 16.05200
C 13.74800 17.05100 15.60100
H 14.52800 17.64500 15.14700

N 14.08000 15.69600 15.78700
C 15.29900 15.17200 15.28500
O 16.23700 16.36200 15.20700
H 15.74800 14.49200 16.00800
C 15.31100 14.66600 13.79000
O 14.18700 15.24400 13.11200
H 13.47100 14.60700 13.07300
H 15.34600 13.57800 13.73300
C 16.60600 15.20000 13.17400
O 17.64200 14.24300 13.30700
H 17.46100 13.71300 12.52700
H 16.43700 15.41400 12.11900
C 17.01200 16.42400 14.04600
H 18.04800 16.34500 14.37400
C 16.72300 17.71600 13.28300
H 16.04300 17.63900 12.43500
H 17.69800 18.08000 12.95600
O 16.07800 18.56500 14.19800
H 16.15200 19.46900 13.88300

28– North (DFTBESCF Energy = -6.32656 Eh; Total Energy of the System (incl. water) = -19.37341 Eh)

O 19.30900 16.80000 17.28000
H 18.82800 17.40100 16.70600
C 19.50800 15.59800 16.61600
H 20.00900 15.78500 15.66600
H 20.13500 14.93600 17.21300
C 18.22000 14.86300 16.24100
O 17.08600 15.73100 16.81300
O 18.17900 13.48500 16.74800
C 17.74900 13.63300 18.14700
H 18.44200 14.26600 18.70000
H 17.73400 12.68400 18.68300
H 16.75500 14.07000 18.23800
C 17.93500 14.86900 14.76700
O 18.43300 13.62300 14.14500
H 18.22900 13.68800 13.20900
H 18.34500 15.77600 14.32100
C 16.35600 14.89600 14.62700
H 16.03900 13.85500 14.69000
H 15.99200 15.19400 13.64300
C 15.92300 15.70200 15.88600
H 15.09600 15.21300 16.40200
N 15.69500 17.09800 15.51500
C 14.40100 17.53400 15.19800
H 13.51700 16.91500 15.24400

N 14.45100 18.68000 14.52700
C 15.79900 19.04200 14.47200
C 16.64500 17.97800 14.93500
N 17.97400 18.02300 14.94400
C 18.50200 19.04200 14.24900
H 19.57400 19.12300 14.34600
N 17.79700 20.09100 13.76900
C 16.46800 20.12600 13.95200
N 15.83400 21.15000 13.40900
H 14.82600 21.10100 13.37400
H 16.39700 21.90100 13.03600

28– South (DFTBESCF Energy = -6.35809 Eh; Total Energy of the System (incl. water) = -19.38934 Eh)

O 16.03700 12.80400 18.44600
H 16.83300 13.26900 18.17600
C 15.31800 12.56600 17.20700
H 16.03700 12.28800 16.43700
H 14.54700 11.79900 17.28300
C 14.59600 13.84300 16.77900
O 15.30400 15.04000 17.11000
O 13.25600 13.98700 17.41600
C 13.30400 13.99100 18.90400
H 14.29500 14.10100 19.34500
H 12.86200 13.05700 19.25100
H 12.73800 14.83000 19.30900
C 14.43400 13.80000 15.31800
O 13.21400 14.44900 14.93100
H 13.22400 15.28300 15.40800
H 14.37700 12.77500 14.95200
C 15.61800 14.55000 14.77700
H 15.49000 14.92300 13.76100
H 16.48800 13.89400 14.75300
C 15.85700 15.63800 15.82100
H 15.33500 16.57300 15.61900
N 17.22200 15.95400 16.04600
C 17.78500 17.28900 16.12600
H 17.19500 18.17300 15.93300
N 19.09600 17.21200 16.37400
C 19.36000 15.84300 16.58300
C 18.18400 15.08200 16.45100
N 18.14700 13.78300 16.70200
C 19.30900 13.24800 17.04400
H 19.24200 12.21100 17.33900

N	20.47400	13.85400	17.12400
C	20.57600	15.19000	16.91800
N	21.73800	15.83700	17.10700
H	21.78700	16.84300	17.03800
H	22.62400	15.35500	17.16800

VIII. Puckering distributions and PMFs for the nucleosides used in this study.

The puckering distributions and PMFs were obtained using the weighted histogram analysis method. For an optimal resolution 72 bins were selected for constructing the graphs and curves.

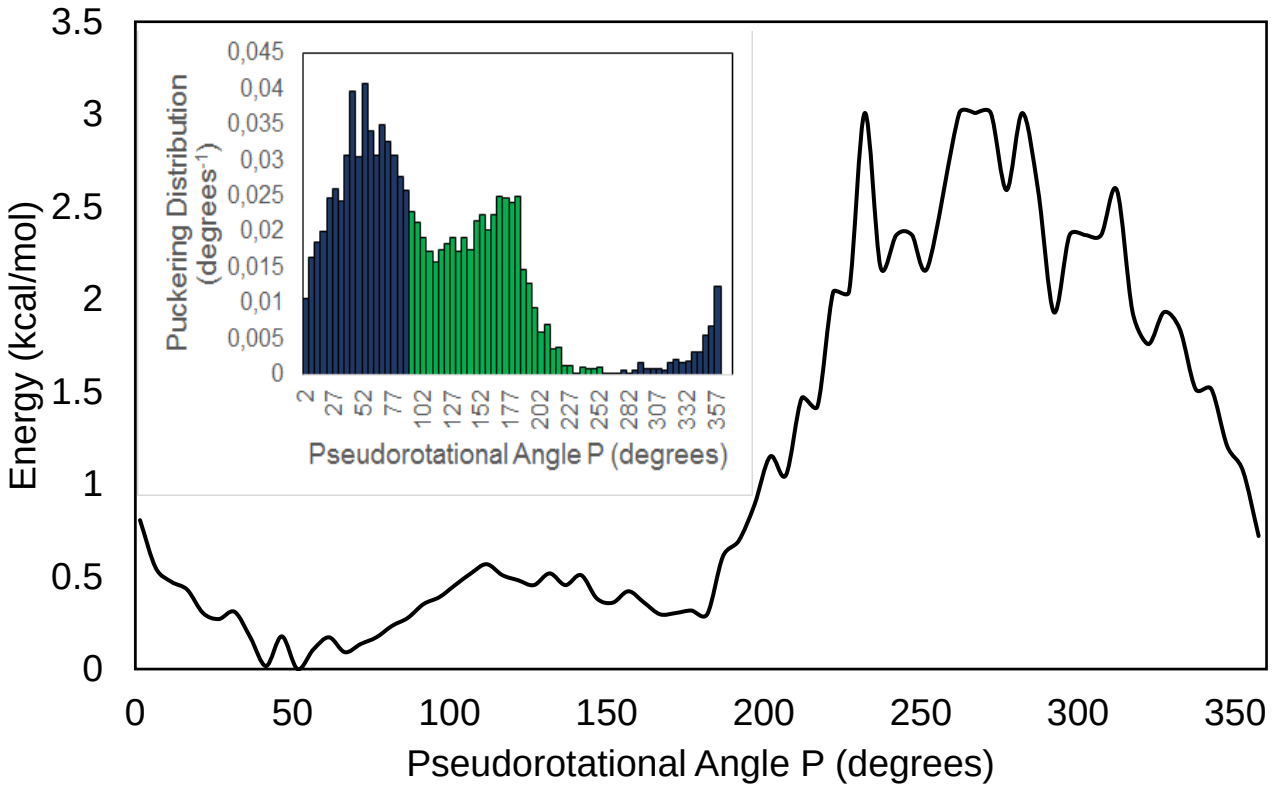


Figure S9. PMF curve for **1** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

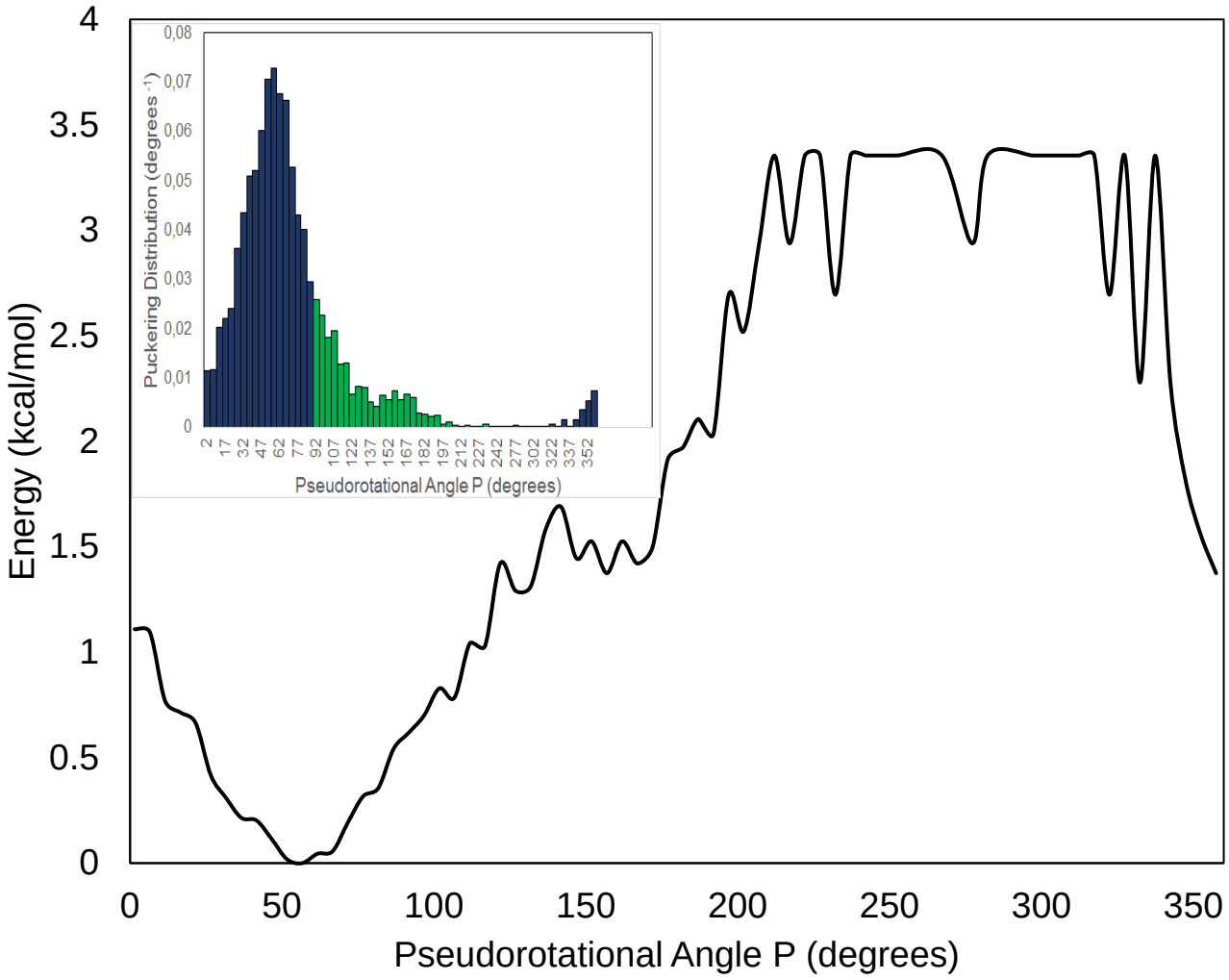


Figure S10. PMF curve for **15** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

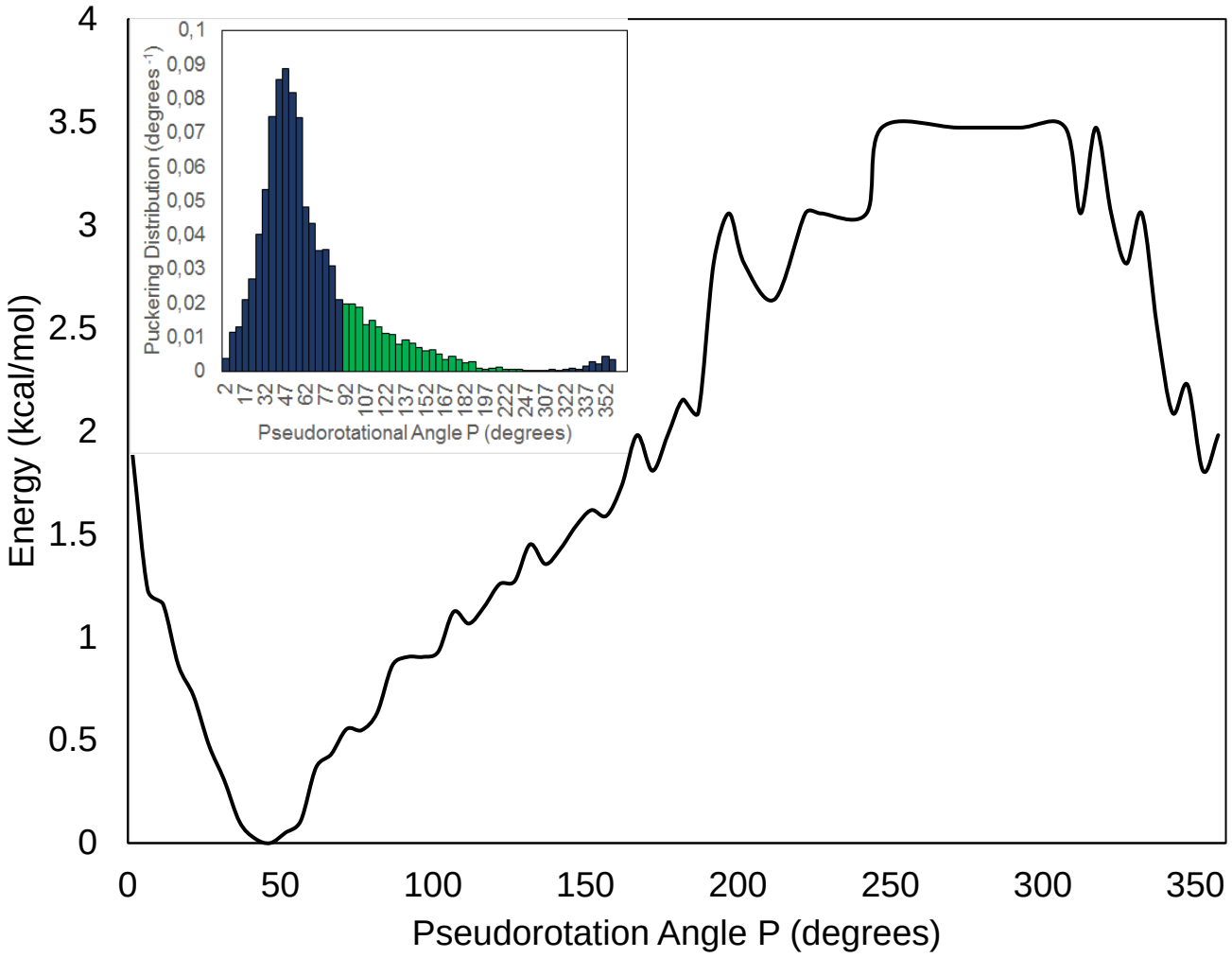


Figure S11. PMF curve for **16** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

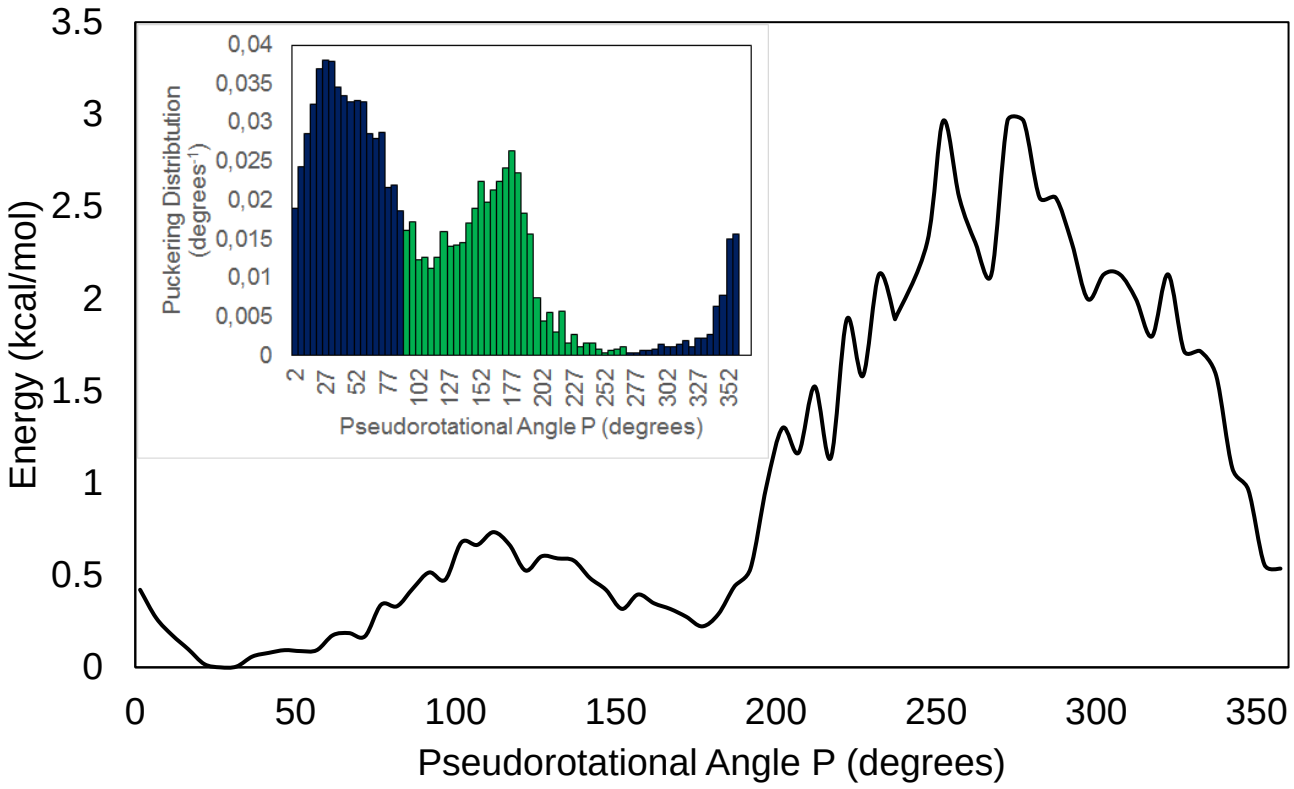


Figure S12. PMF curve for **17** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

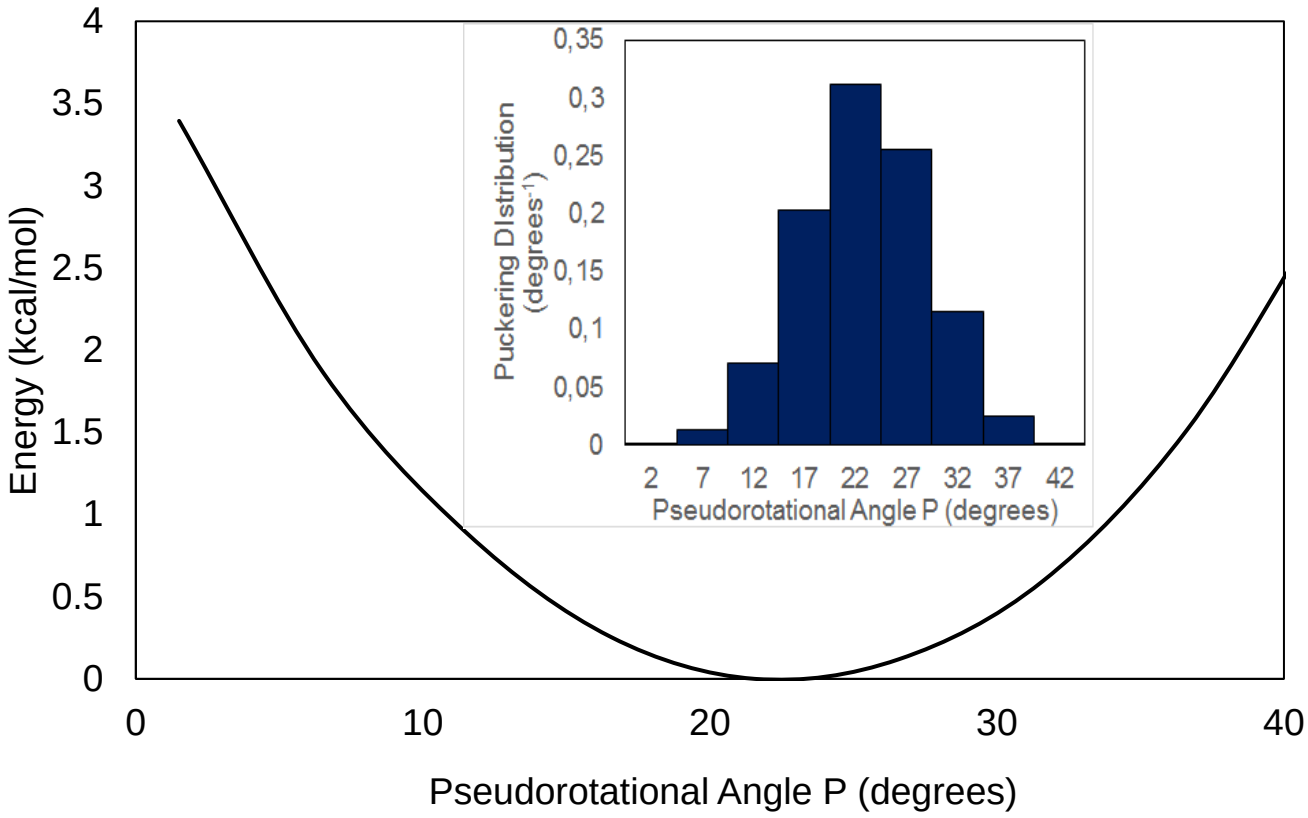


Figure S13. PMF curve for **18** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation.

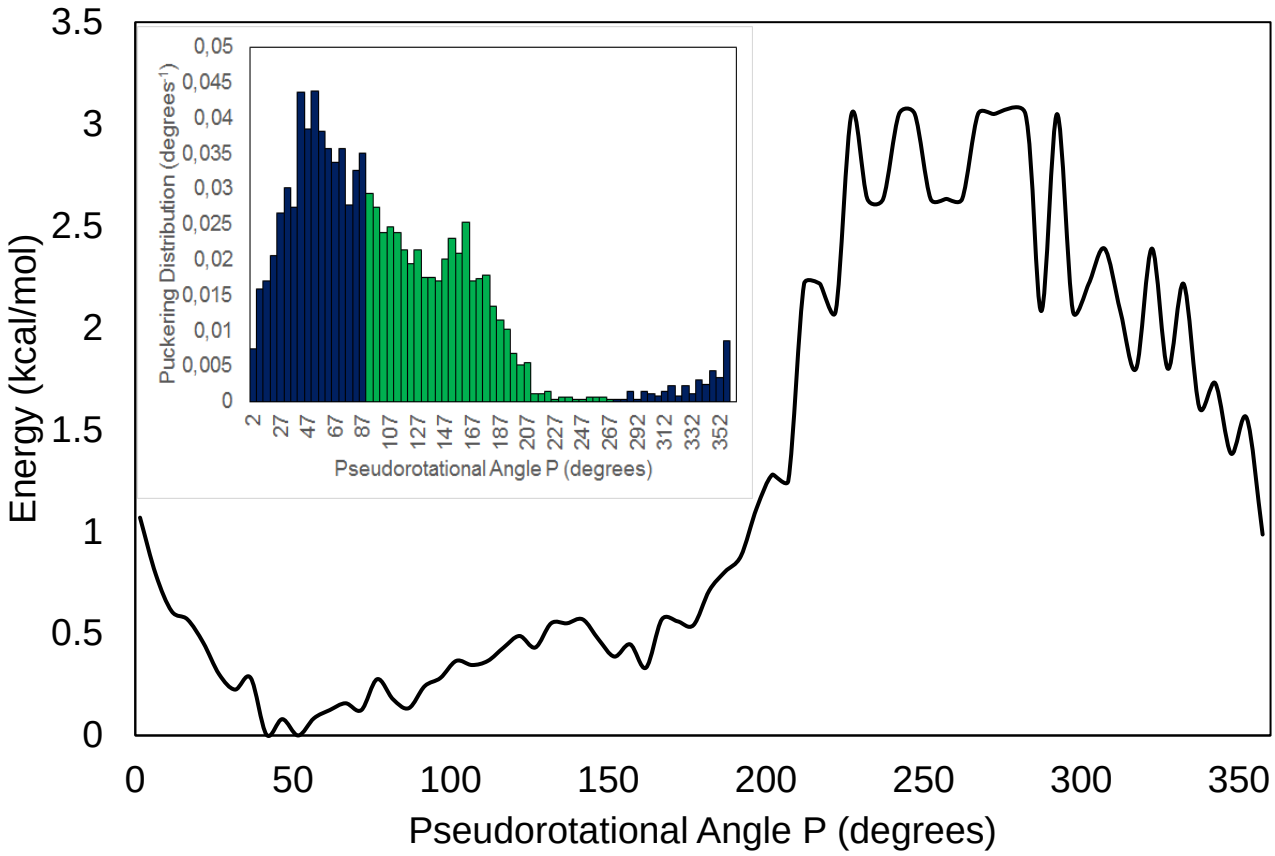


Figure S14. PMF curve for **19** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

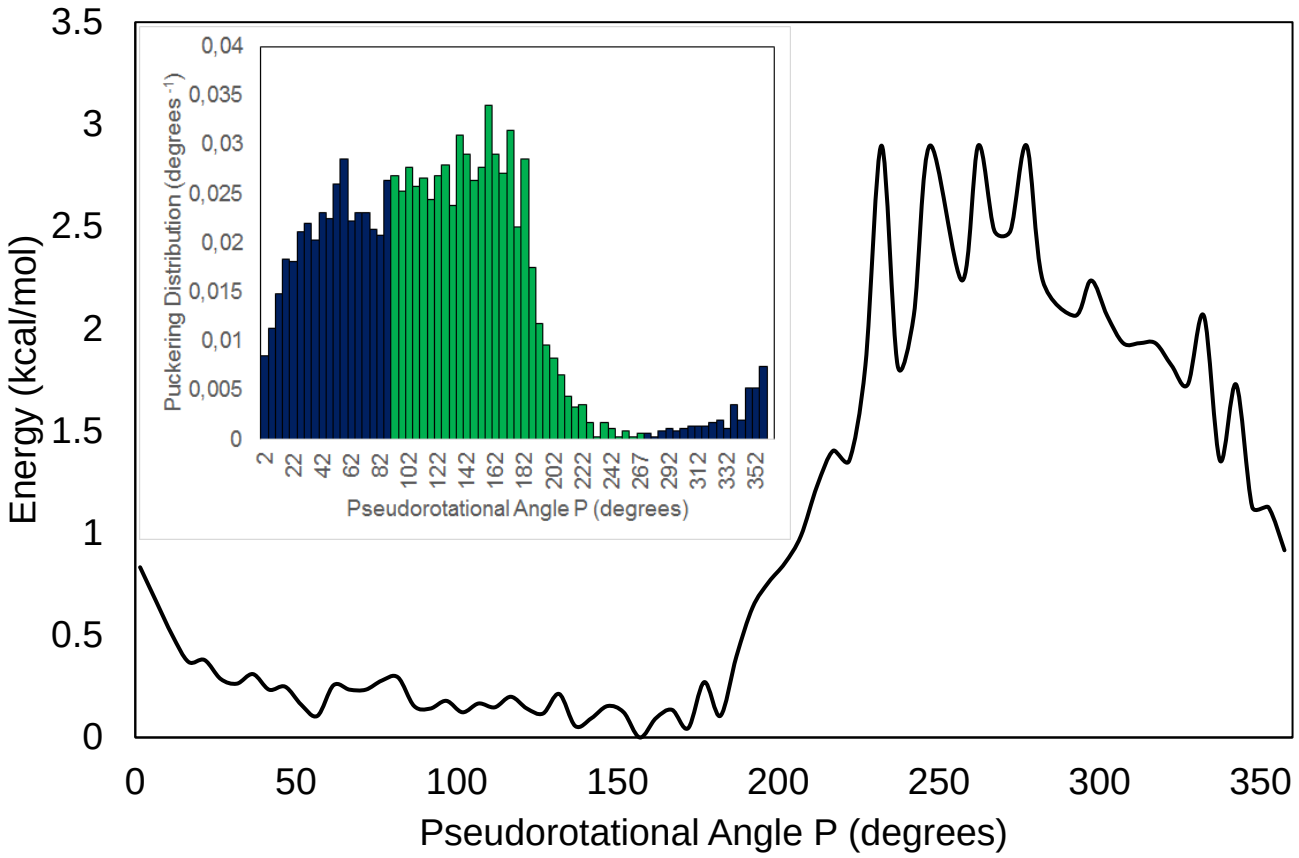


Figure S15. PMF curve for **20** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

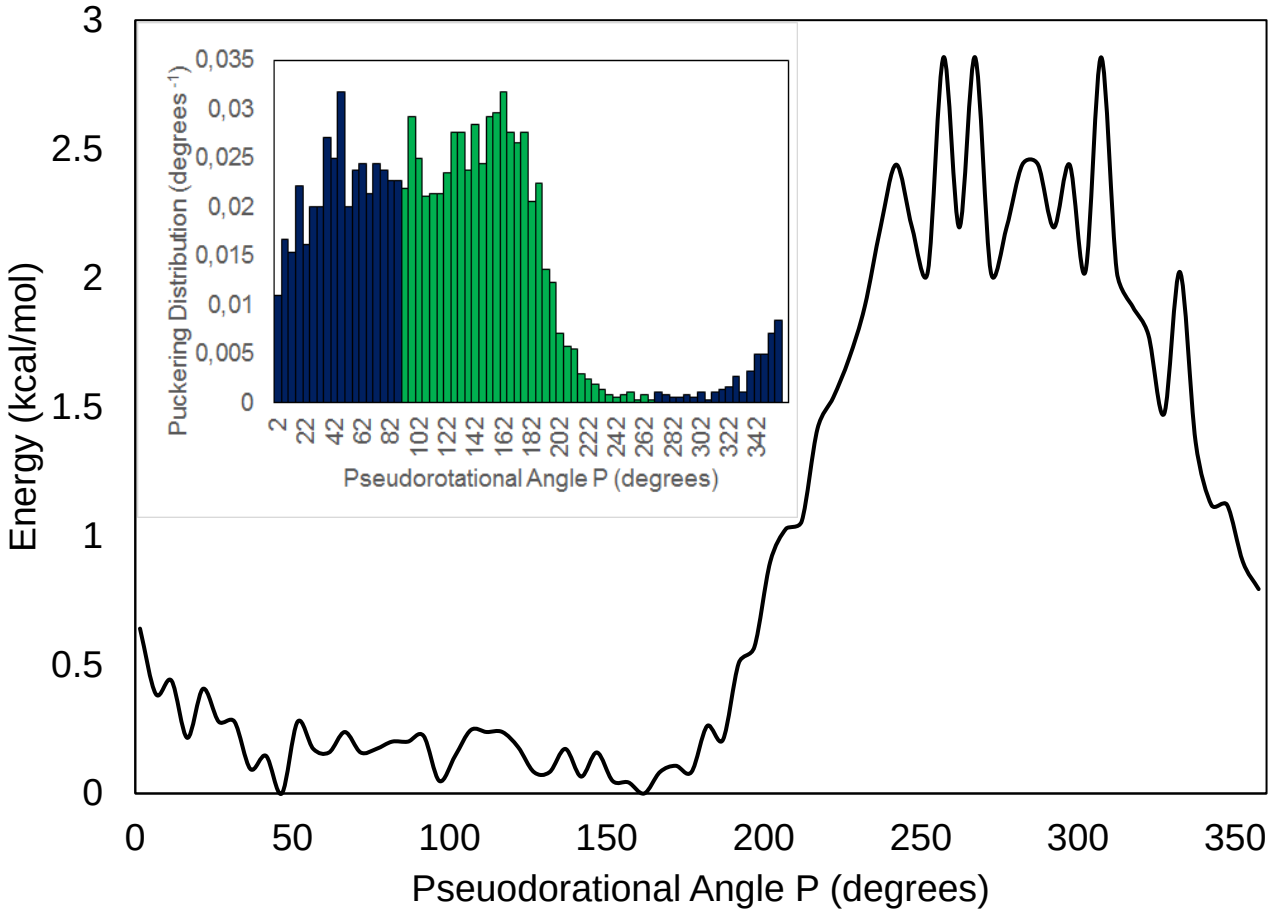


Figure S16. PMF curve for **21** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

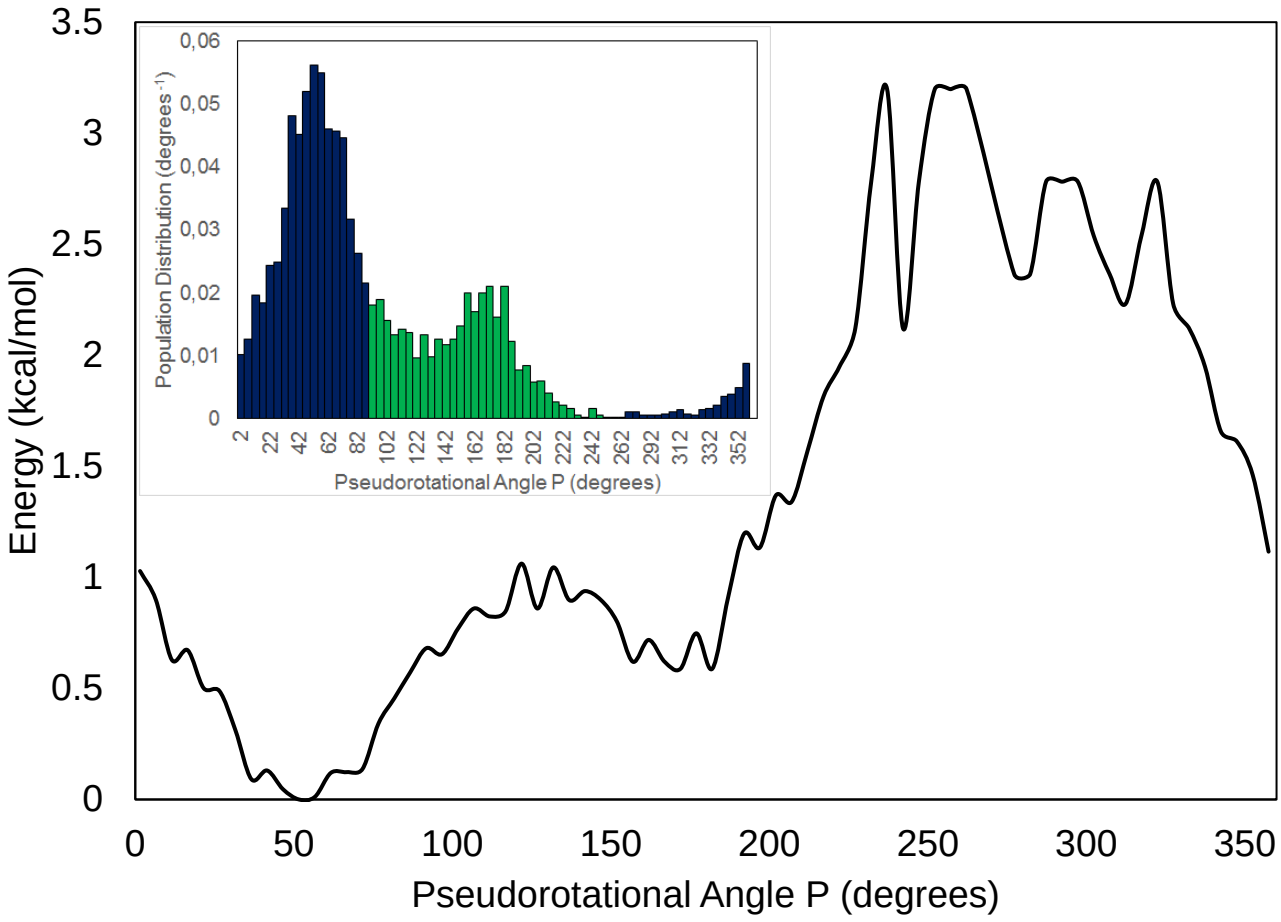


Figure S17. PMF curve for **22** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

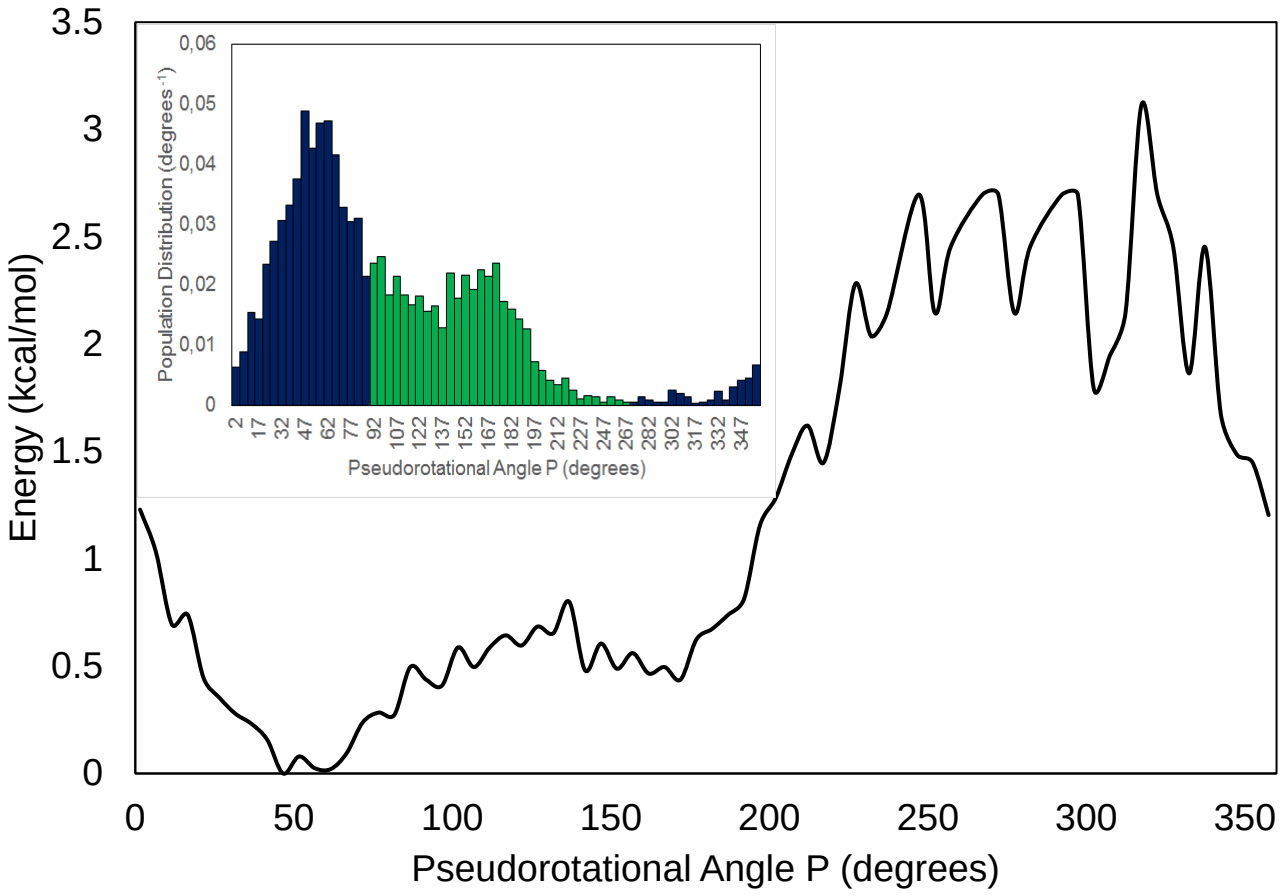


Figure S18. PMF curve for **23** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

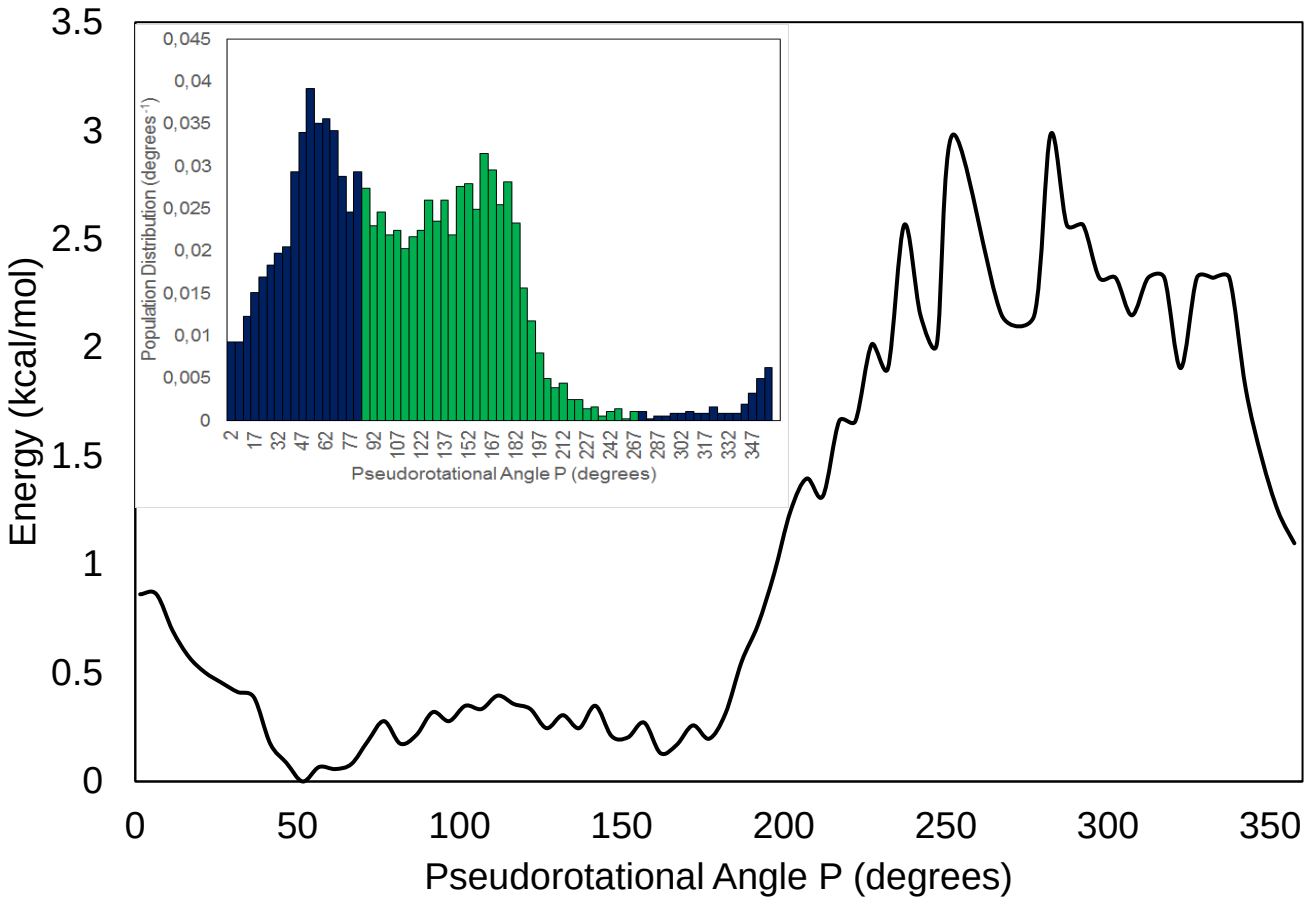


Figure S19. PMF curve for **24** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

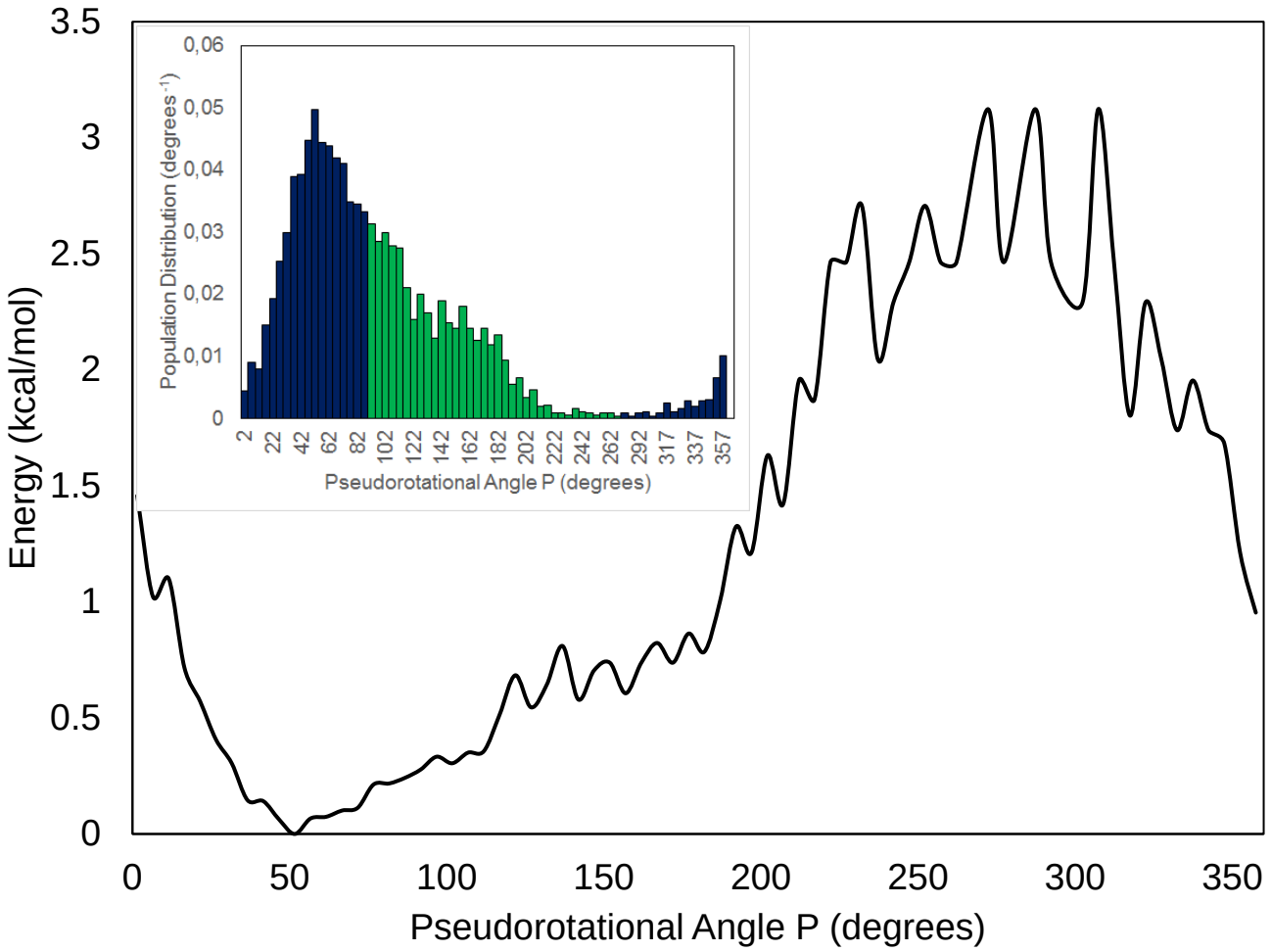


Figure S20. PMF curve for nucleoside **25**. Inset shows the sugar puckering distribution along the pseudorotational angle P.

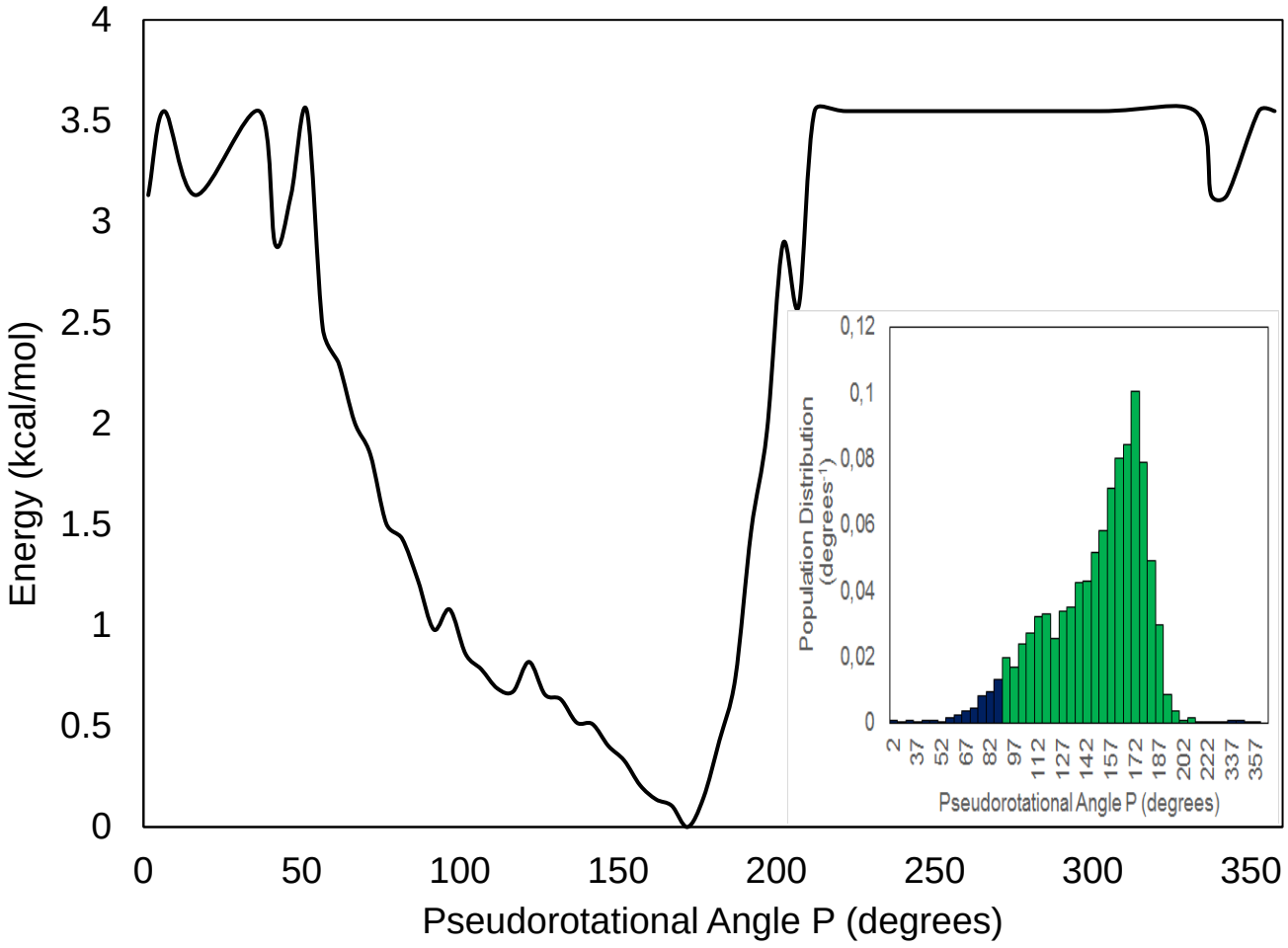


Figure S21. PMF curve for **26** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

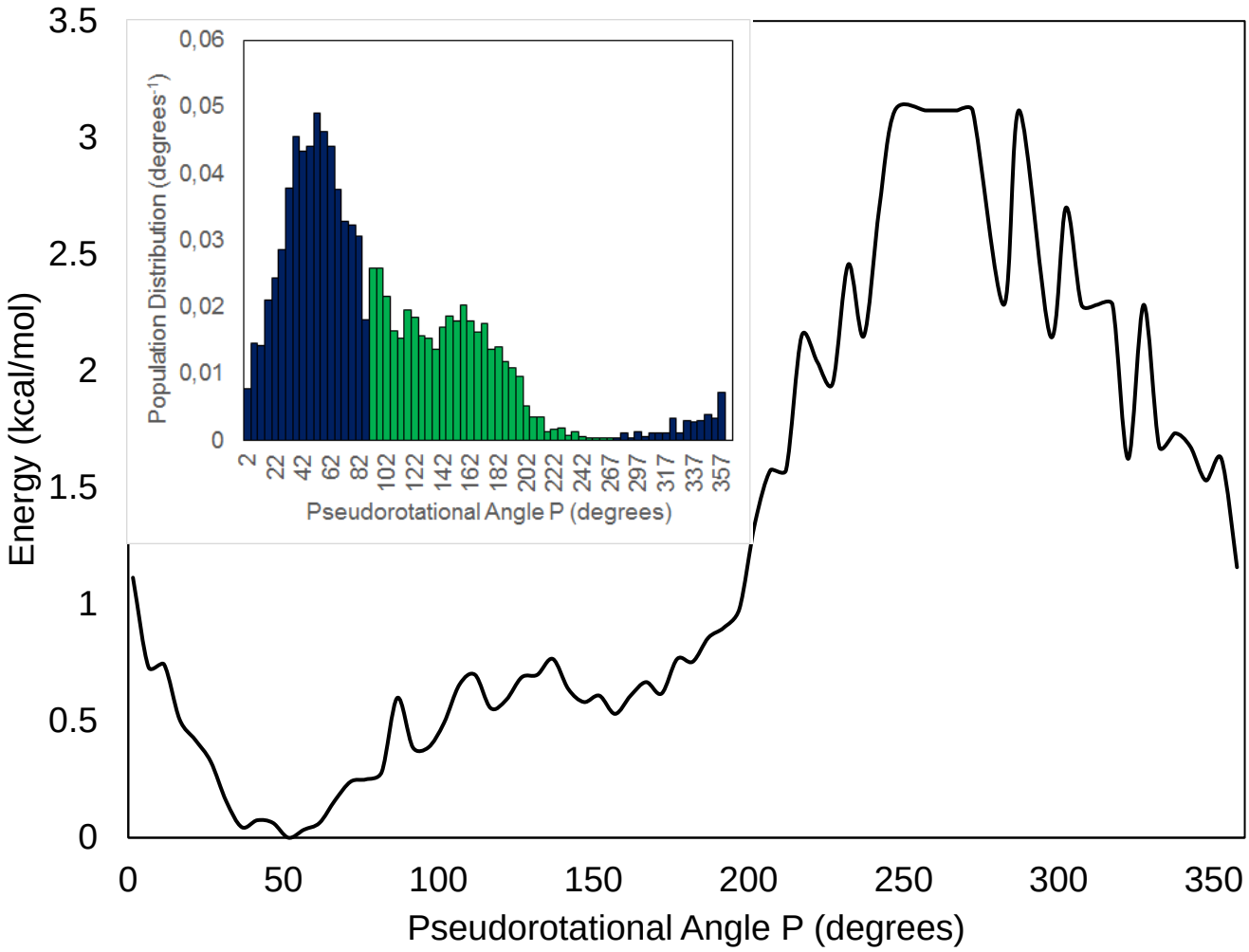


Figure S22. PMF curve for **27** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

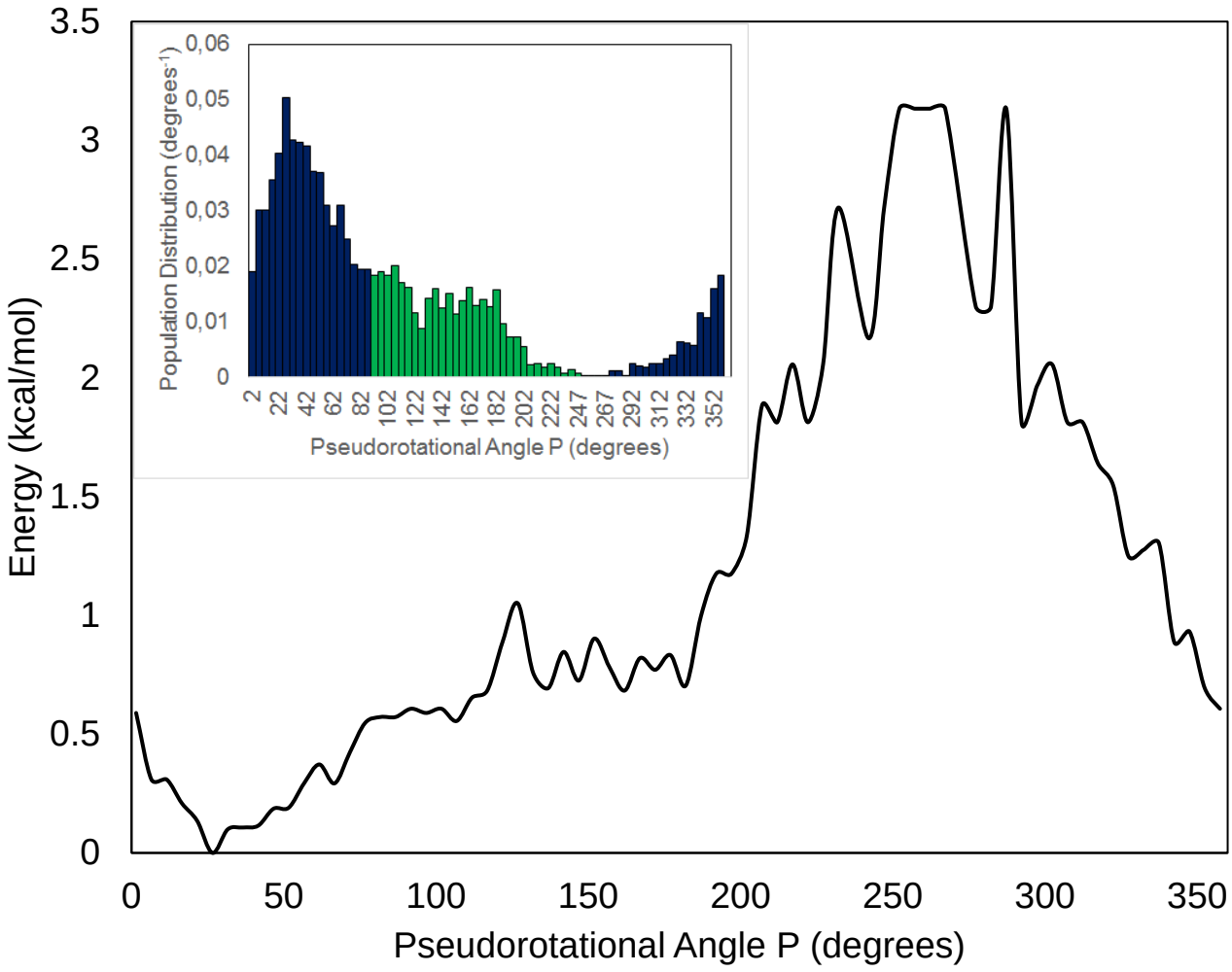


Figure S23. PMF curve for **28** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

IX. Simulation time vs *N/S* ratio for the Nucleosides Used in this Study

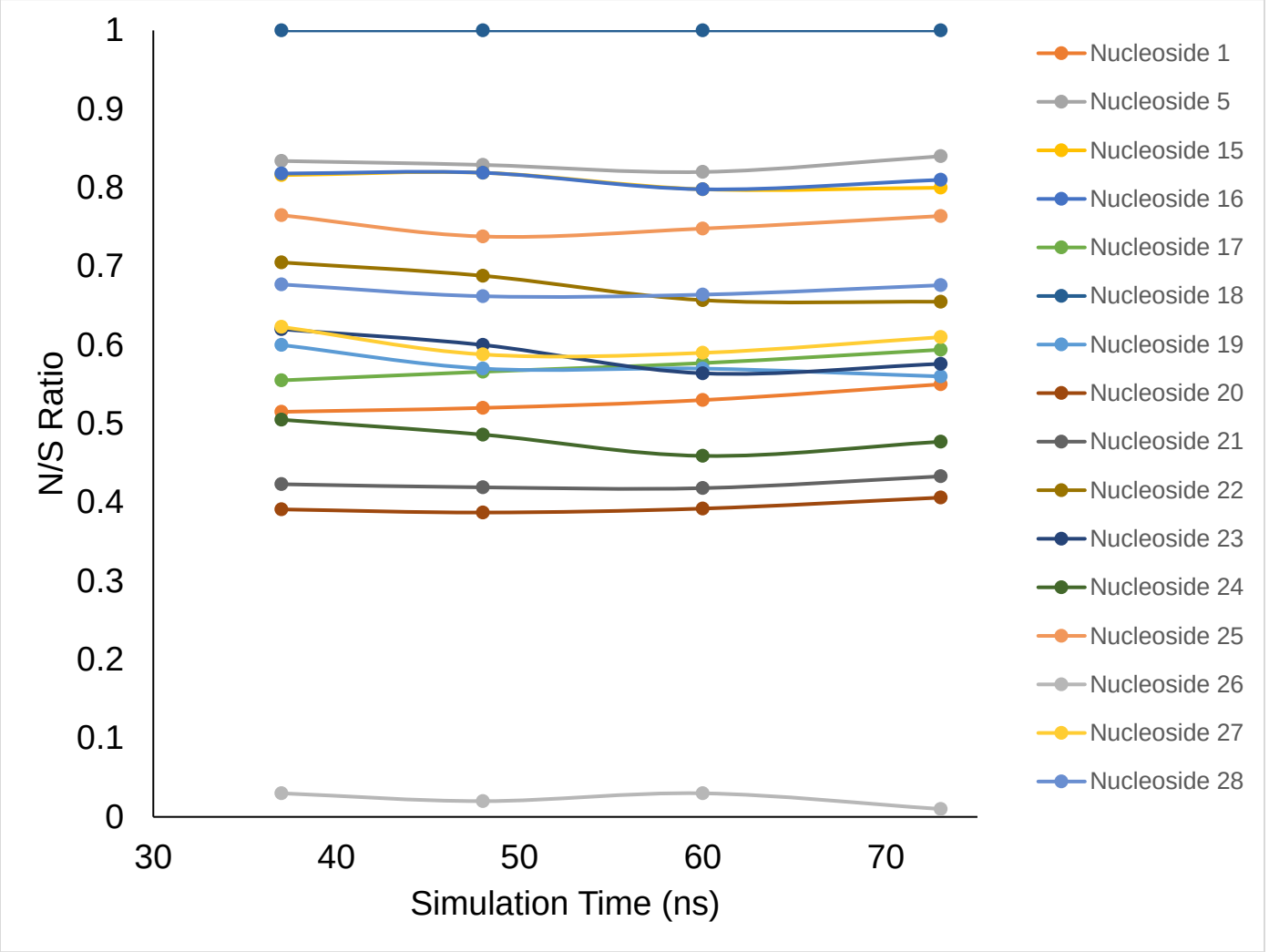


Figure S24. Plot of the change in *N/S* ratio (by Boltzmann population distribution) as a function of the simulation time. The selected simulation times were 37, 48, 60 and 72ns.

X. PMFs and pucker distributions for monosaccharides 6-14.

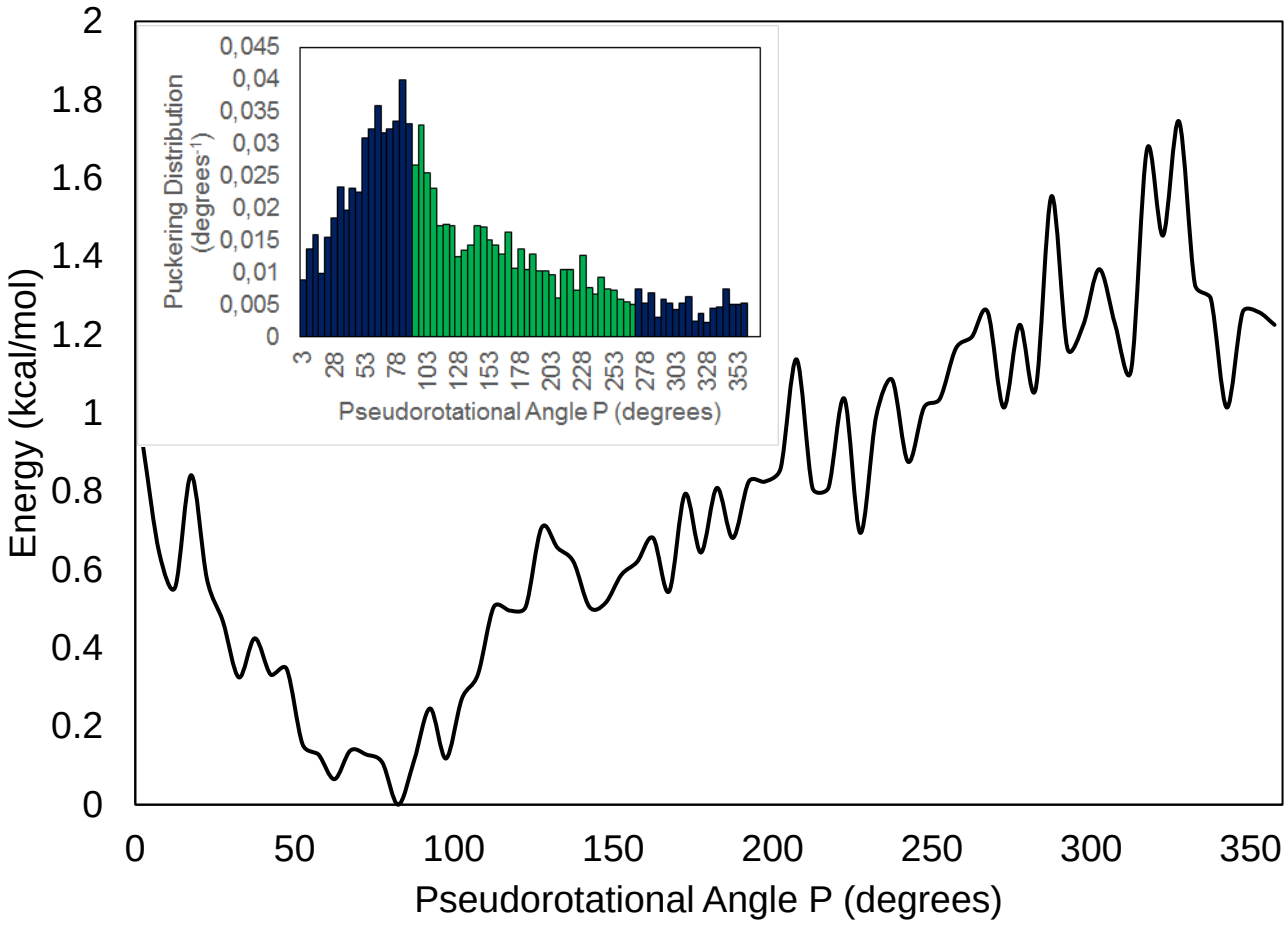


Figure S25. PMF curve for **6** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

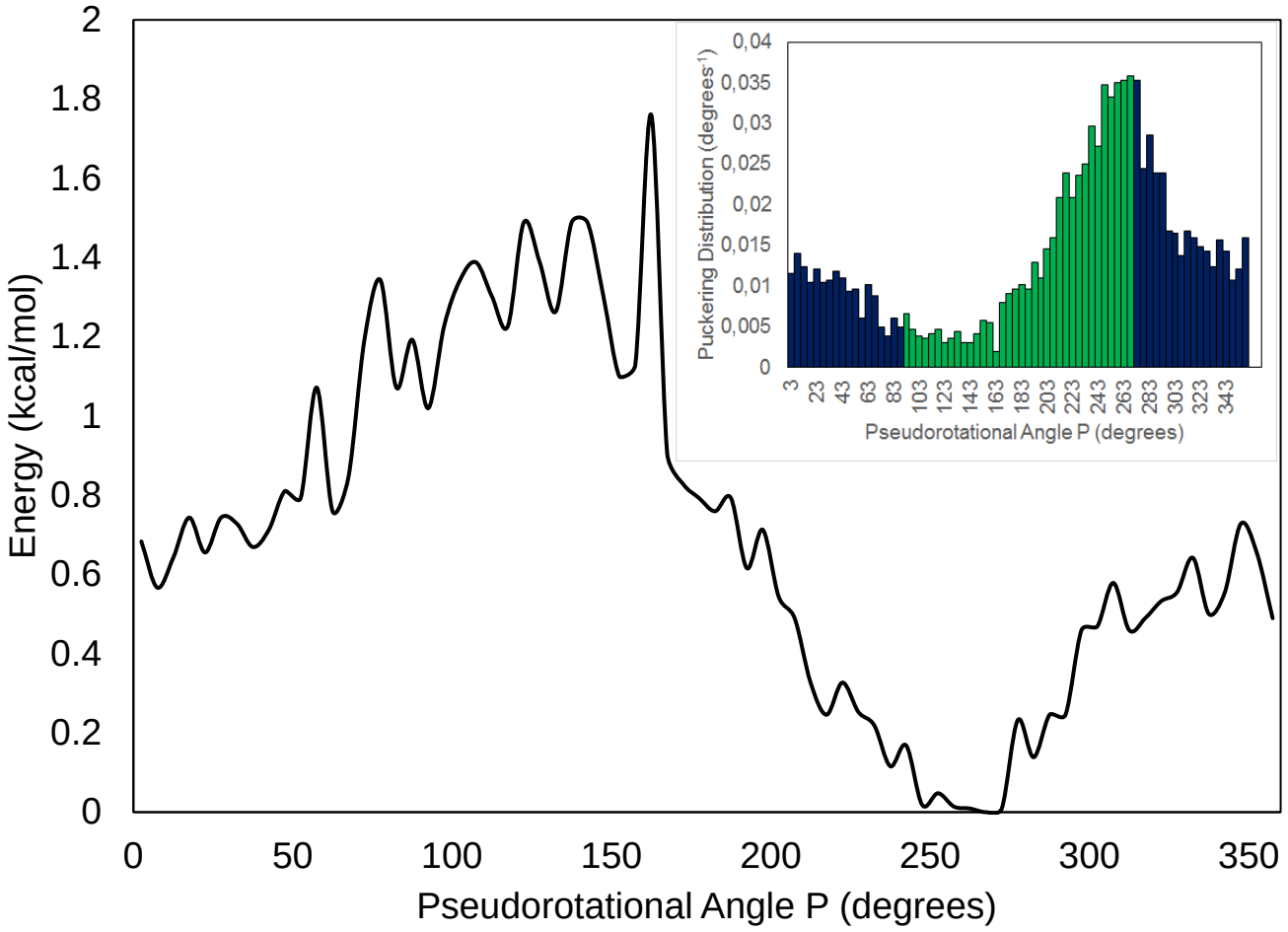


Figure S26. PMF curve for **7** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

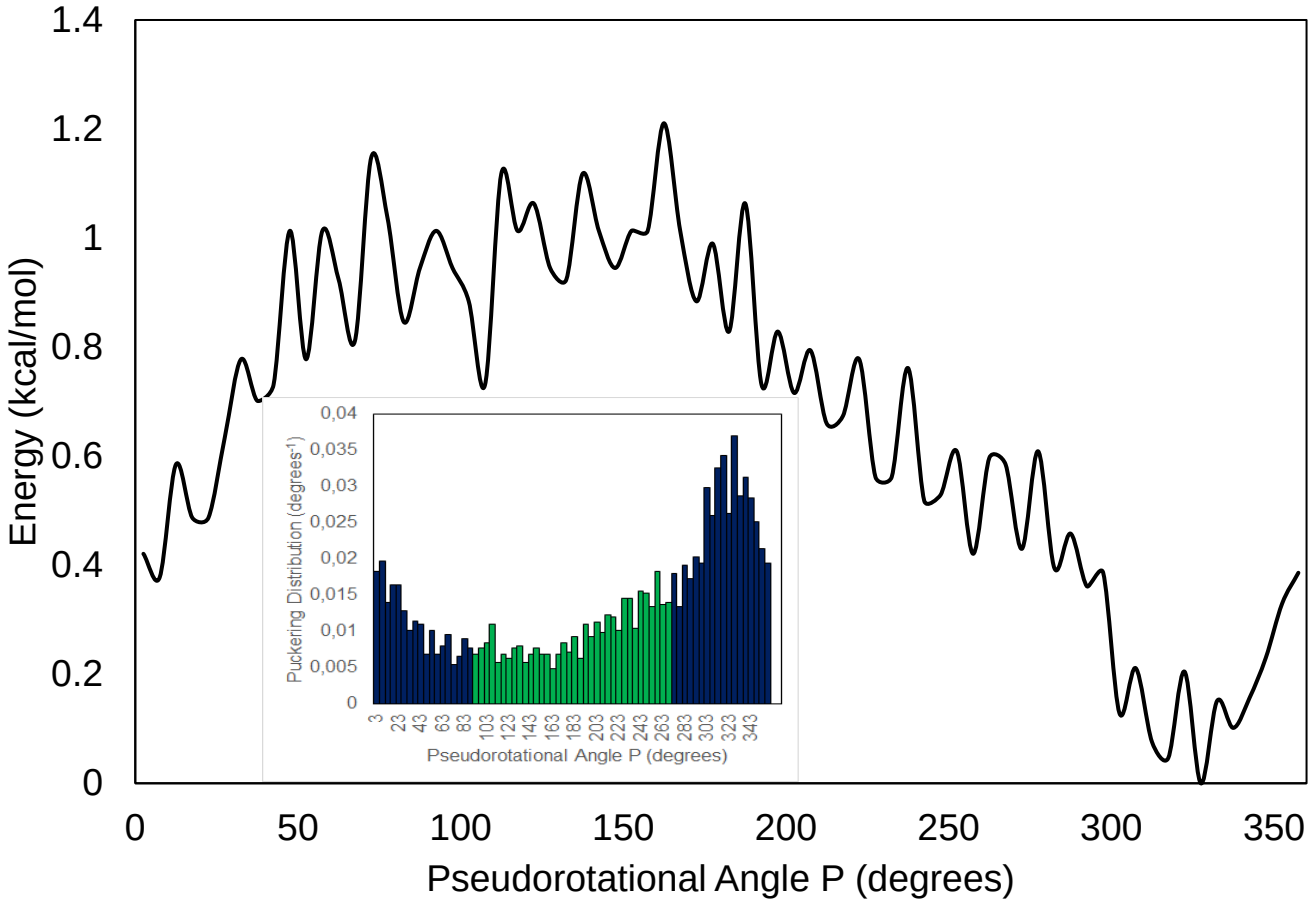


Figure S27. PMF curve for **8** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

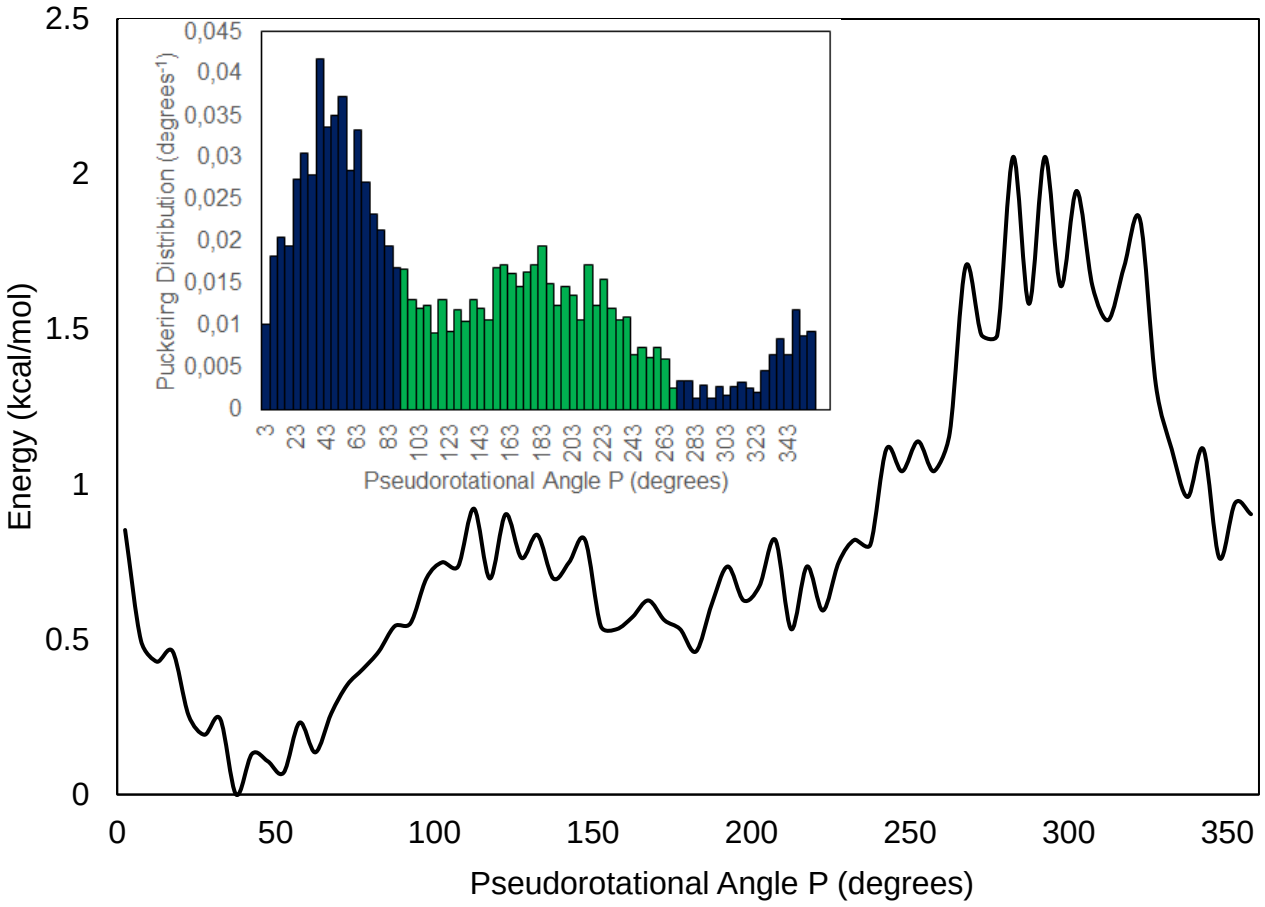


Figure S28. PMF curve for **9** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

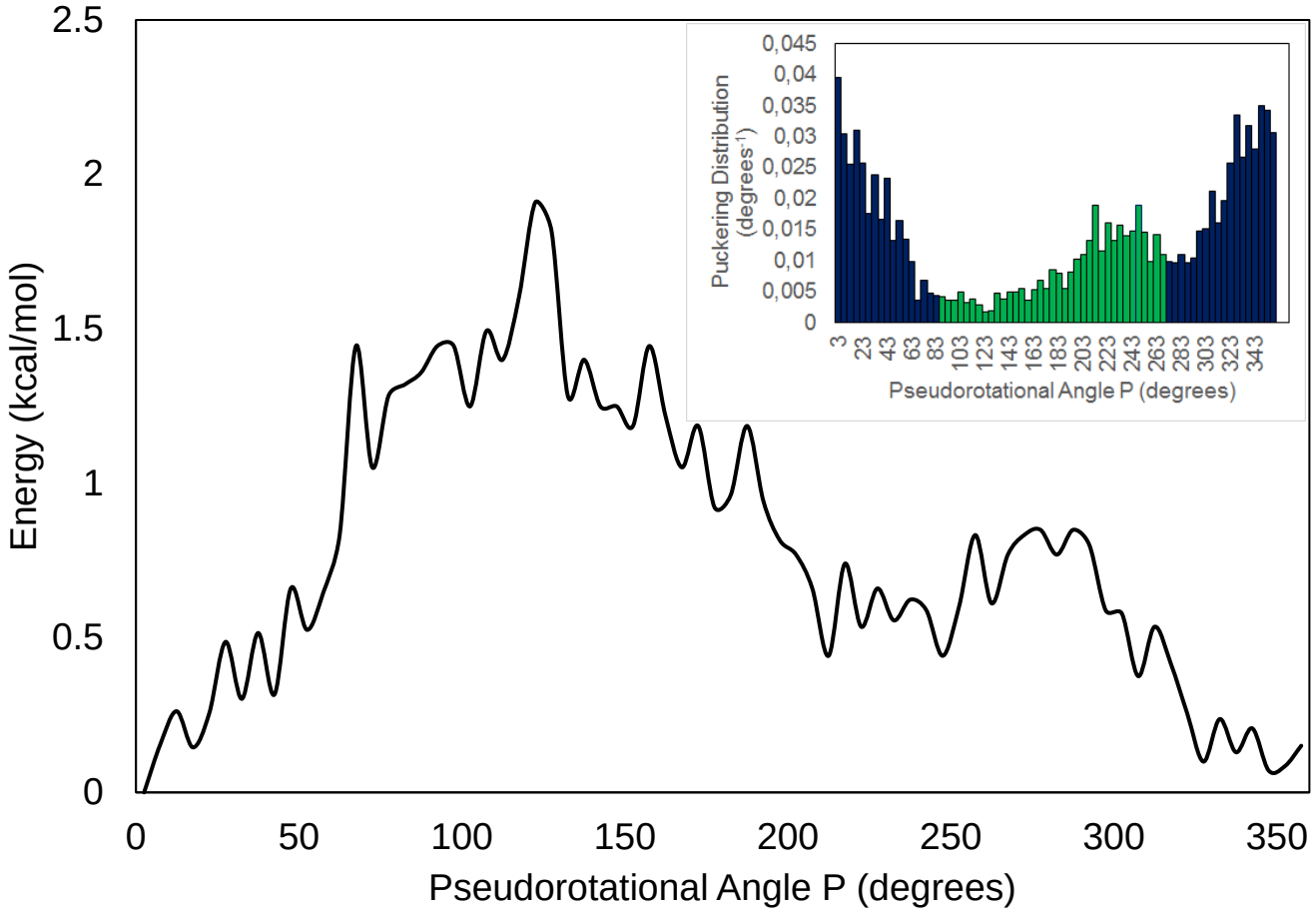


Figure S29. PMF curve for **10** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

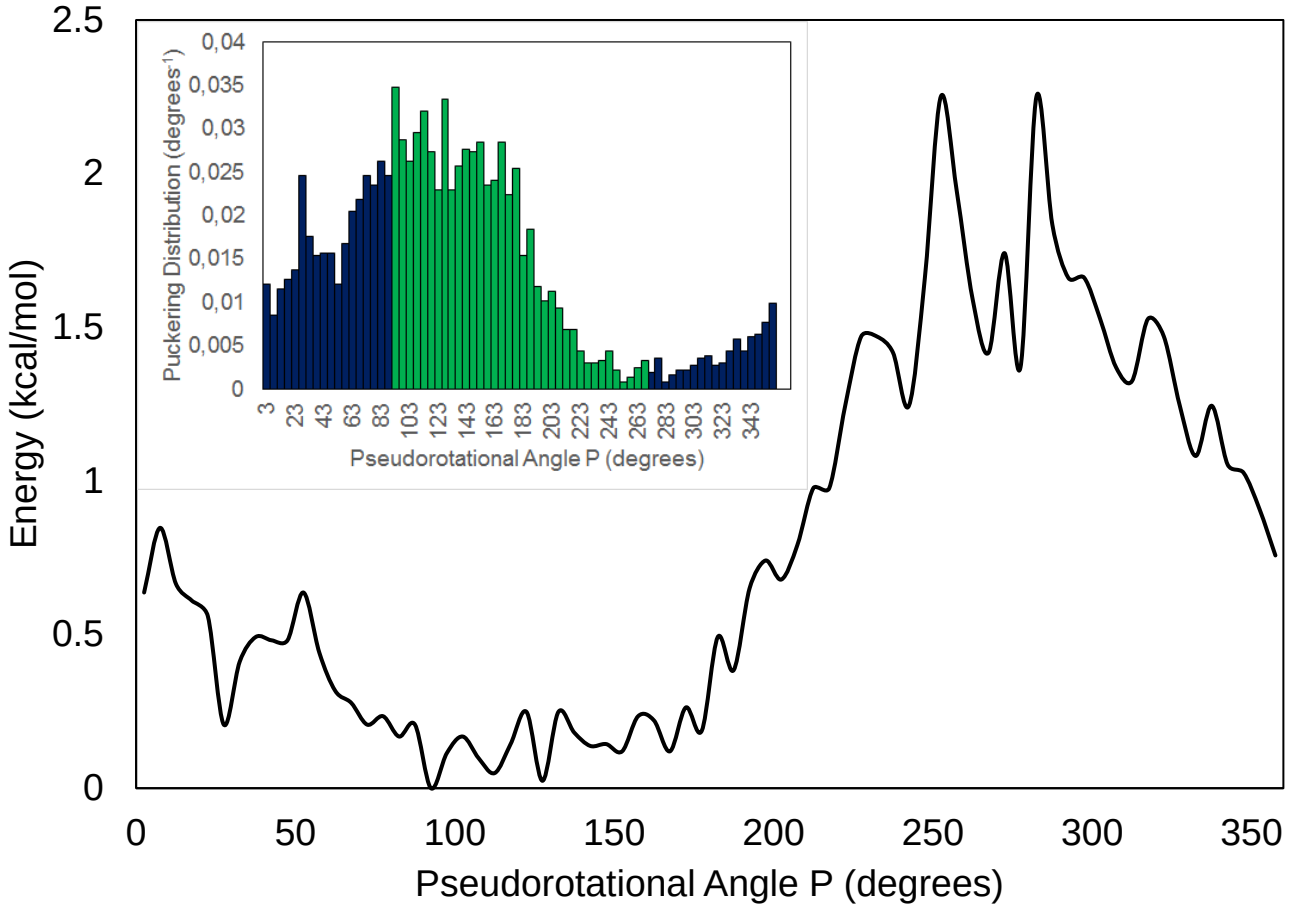


Figure S30. PMF curve for **11** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

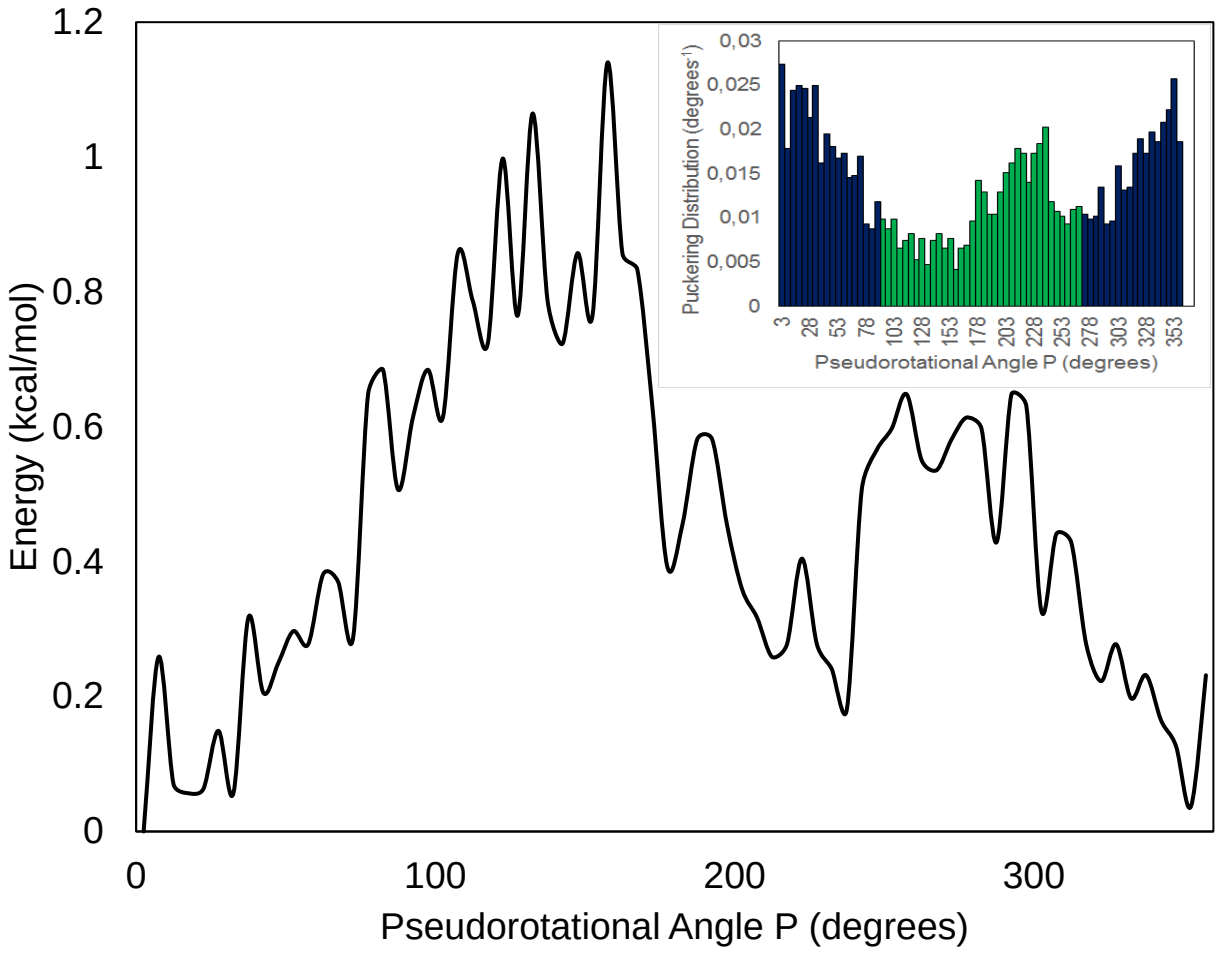


Figure S31. PMF curve for **12** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

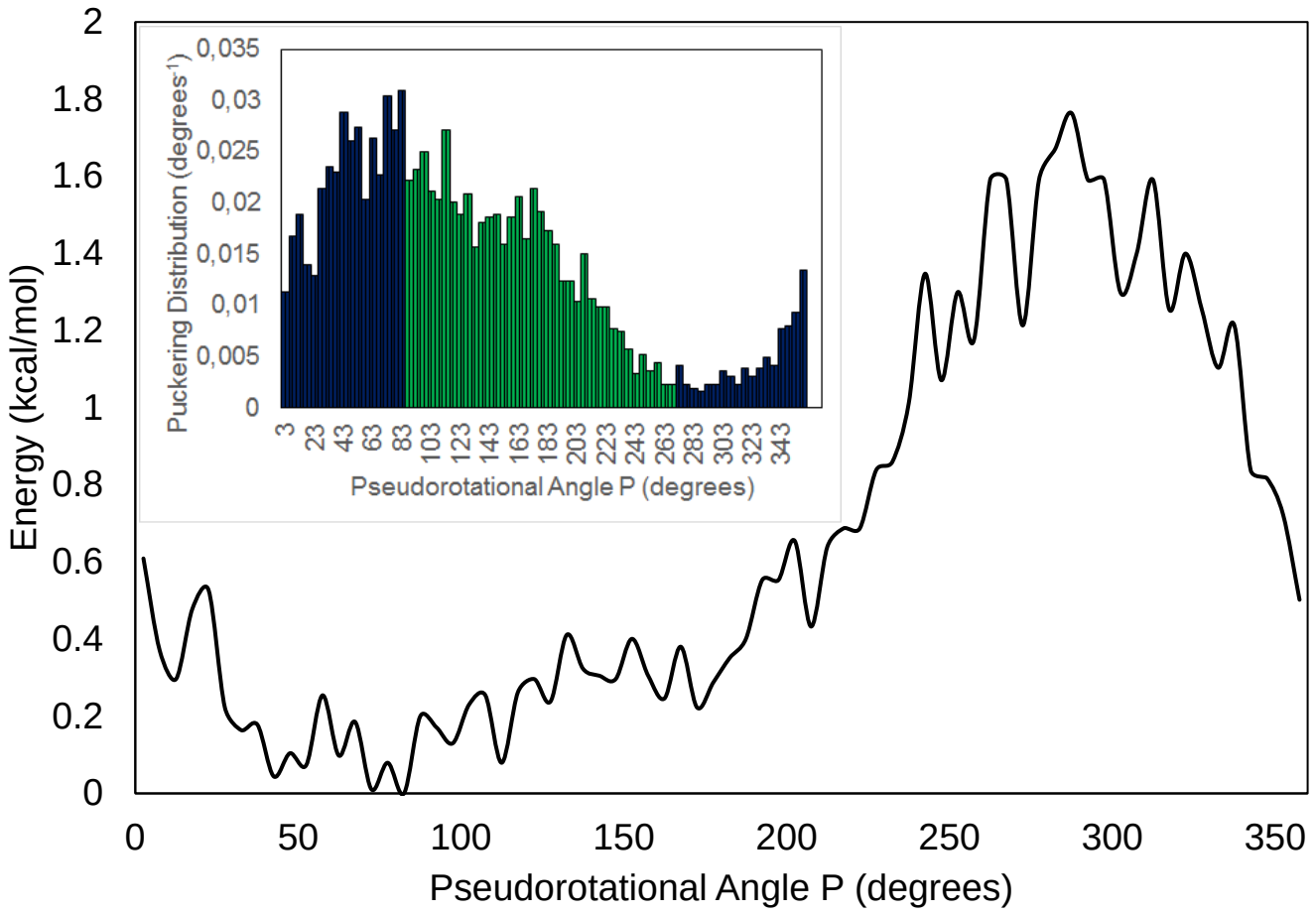


Figure S32. PMF curve for **13** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

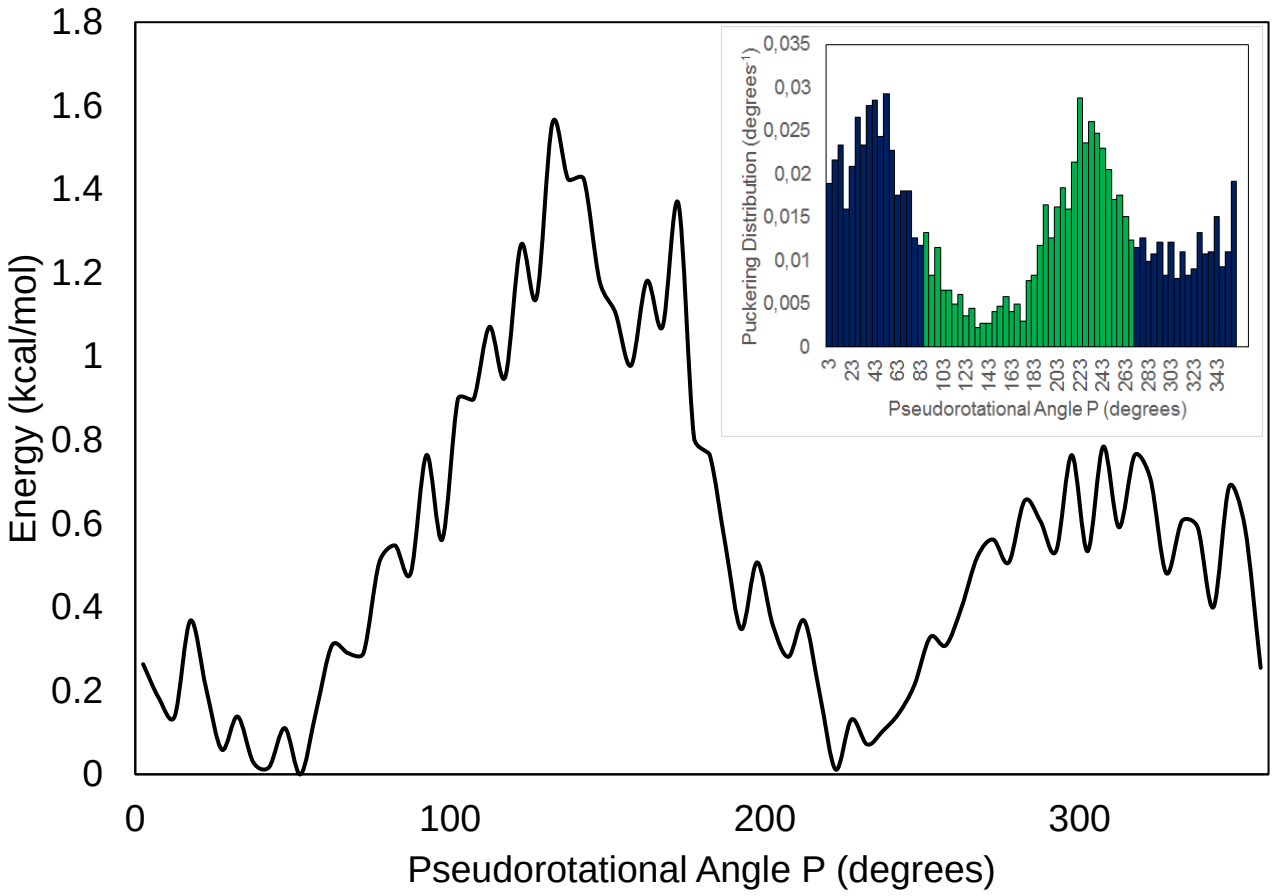


Figure S33. PMF curve for **14** along the pseudorotational angle P. Inset shows the puckering distribution around the same angle. Blue – *Northern* conformation, Green – *Southern* conformation.

XI. Lowest energy conformers for monosaccharides 6-14.

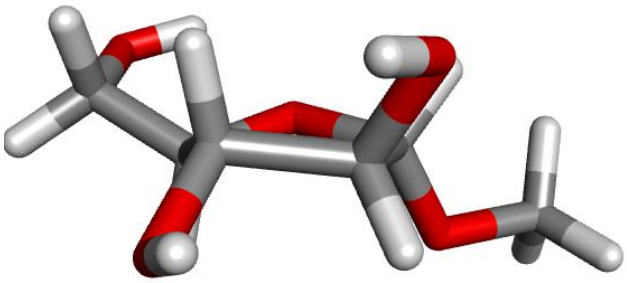


Figure S34. E_4 conformer – monosaccharide **6**. (DFTBESCF Energy = -3.82067 Eh; Total Energy of the System (incl. water) = -14.00488 Eh)

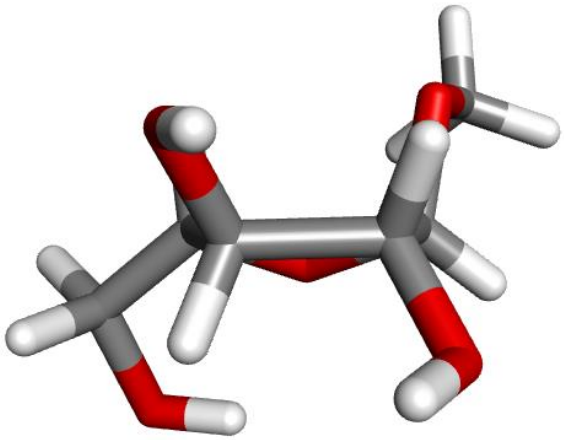


Figure S35. ^4T_O conformer – monosaccharide **7**. (DFTBESCF Energy = -3.83101 Eh; Total Energy of the System (incl. water) = -13.01318 Eh)

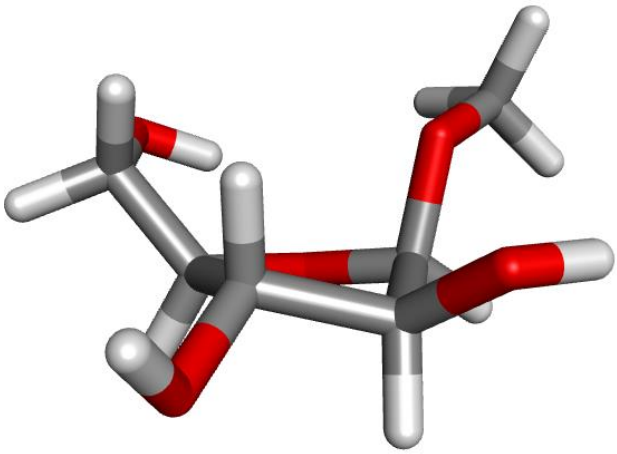


Figure S36. $^{1}T_{2}$ conformer – monosaccharide **8**. (DFTBESCF Energy = -3.80083 Eh; Total Energy of the System (incl. water) = -12.94960 Eh)

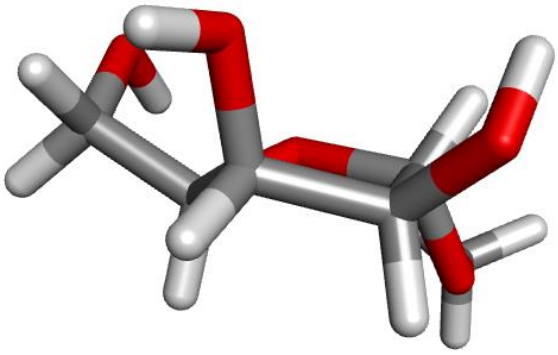


Figure S37. ^{3}E conformer – monosaccharide **9**. (DFTBESCF Energy = -3.80729 Eh; Total Energy of the System (incl. water) = -13.85668 Eh)

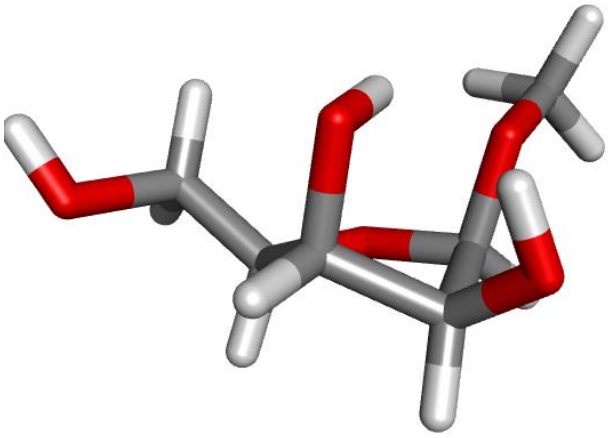


Figure S38. $^{3}T_{2}$ conformer – monosaccharide **10**. (DFTBESCF Energy = -3.82095 Eh; Total Energy of the System (incl. water) = -12.92458 Eh)

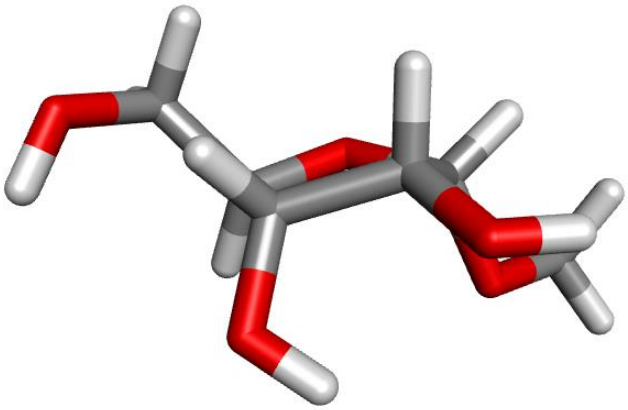


Figure S39. E_1 conformer – monosaccharide **11**. (DFTBESCF Energy = -3.80835 Eh; Total Energy of the System (incl. water) = -13.83692 Eh)

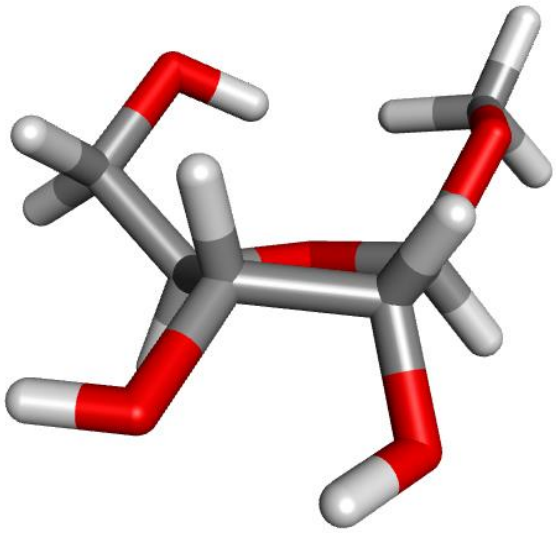


Figure S40. $^{3}T_2$ conformer – monosaccharide **12**. (DFTBESCF Energy = -3.7965 Eh; Total Energy of the System (incl. water) = -12.65191 Eh)

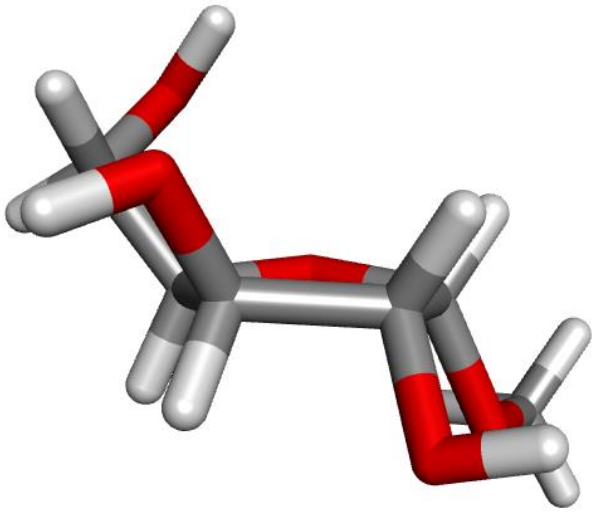


Figure S41. ^{O}E conformer – monosaccharide **13**. (DFTBESCF Energy = -3.78863 Eh; Total Energy of the System (incl. water) = -13.35899 Eh)

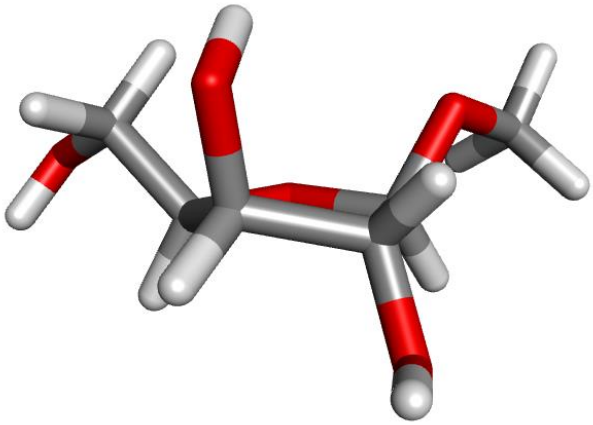


Figure S42. $^{3}T_{4}$ conformer – monosaccharide **14**. (DFTBESCF Energy = -3.82471 Eh; Total Energy of the System (incl. water) = -13.32617 Eh)

XII. Determination of *N/S* Ratio for Nucleoside 28.[36]

The *North* conformer population was estimated for 4′-methoxy-2′-deoxyadenosine by applying the following equation: *North* (%) = 100 - 10 × ($J_{H1'H2'}$) on their calculated $J_{H1'H2'}$ value of 7.6 Hz, as retrieved on page 16 of their supporting information. From this computation we obtained a ratio of 76:24 for *N/S* contribution, as opposed to their reported >90% within their main text.

XIII. Experimental vs Predicted *N/S* Ratios for Monosaccharides and Nucleosides; Crystal Structure vs Predicted Conformers for Monosaccharides.

Table S6. Comparison between the *N/S* ratios obtained for monosaccharides.

Entry	Monosacch.	Pseudo-Experimental *N/S* ratio[24, 32, 37 38]	Crystal Structure[33, 39-40] Conformer	Method	Predicted *N/S* ratio[a]	Predicted Conformer
1	**6**	39:61	E_4	DFT[c,32,38]	nd[d]	E_4 - E_1
2		83:17[b]		GLYCAM[37]	79:21	nd
3		83:17[b]		SCC-DFTB[37]	91:9	nd
4				our method	53:47	E_4 - ^{2}E
5	**7**	83:17[b]	nd	GLYCAM[37]	77-84:23-16	nd
6				our method	49:51	$^{3}T_2$ - $^{4}T_O$
7	**8**	87:13	$^{1}T_2$	DFT[32,38]	nd	E_2 - ^{4}E
8				our method	64:36	$^{1}T_2$ - E_3
9	**9**	65:35	^{3}E	DFT[32,38]	nd	^{O}E - E_3

10				our method	56:44	$^3T_4$ - ^2E
11	**10**	77:23	nd	DFT[32,38]	nd	E_2 - ^4E
12				our method	70:30	$^3T_2$ - $^4T_3$
13	**11**	8:92	nd	DFT[32,38]	nd	E_4 - ^2E
14				our method	39:61	$^3T_4$ - E_1
15	**12**	86:14	E_2	DFT[32,38]	nd	^1E - ^4E
16				our method	61:39	$^3T_2$ - ^2E
17	**13**	4:96	^2E	DFT[32,38]	nd	^3E - E_1
18				our method	48:52	$^3T_4$ - ^OE
19	**14**	78:22	nd	DFT[32,38]	nd	E_2 - ^4E
20				our method	58:42	$^3T_4$ - $^4T_3$

[a] For our method - at 303K based on the Boltzmann population distribution. [b] Not determined with PSEUROT. [c] The DFT predicted *N/S* conformers are determined in solution using a solvent model. [d] nd: not determined.

Table S7. The predicted *N/S* ratios obtained for the nucleosides used in this study.

Entry	Nucleoside	Experimental *N/S* ratio	Predicted *N/S* ratio[a]	Predicted *N/S* ratio[b]
1	**1**	50:50[41]	55:45	58:42
2	**5**	87:13[21]	84:16	86:14
3	**15**	100:0 [41]	80:20	91:9
4	**16**	80:20[42]	81:19	86:14
5	**17**	75:25[41]	60:40	58:42
6	**18**	100:0[43]	100:0	-[c]
7	**19**	54:46[44]	56:44	53:37
8	**20**	37:63[41]	41:59	44:56
9	**21**	39:61[41]	43:57	45:55
10	**22**	85:15[41]	66:34	76:24

11	**23**	87:13[41]	58:42	70:30
12	**24**	58:42[41]	48:52	56:44
13	**25**	37:63[45]	60:40	72:28
14	**26**	77:23[45]	1:99	0:100
15	**27**	46:54[41]	61:39	69:31
16	**28**	91:9[36](76:24)[d]	68:32	68:32

[a]At 303 K; based on the Boltzmann population distribution; [b]At 303 K, based on the differences in energy between the *North* and *South* minima; [c]For entry 6 only one minimum in the *North* conformation was observed; [d]according to our own calculations based on the data provided in the reference.

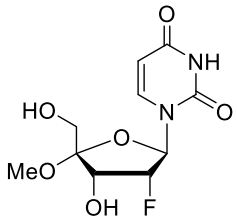
Figure 2 (main text) was made using the data in Tables S6 and S7. More specifically, the R^2 values for nucleosides and monosaccharides are as follows:

R^2 (nucleosides, after removal of 26) = 0.6234

R^2 (monosaccharides) = 0.5538.

The R^2 values in our case are computed based on the correlation between our predicted data and pseudo-experimental data, which means that we are effectively accounting for two different errors (since the pseudo-experimental data comes with an associated error).

XIV. NMR data for compound 5.



5

^{1}H NMR (500 MHz, Methanol-d_4) δ 7.83 (d, J = 8.1 Hz, 1H), 6.04 (dd, J = 19.1, 1.4 Hz, 1H), 5.68 (d, J = 8.1 Hz, 1H), 5.08 – 4.95 (m, 1H), 4.54 (dd, J = 22.6, 5.8 Hz, 1H), 3.90 (d, J = 11.9 Hz, 1H), 3.65 (d, J = 11.9 Hz, 1H), 3.42 (s, 3H).^{13}C NMR (126 MHz, MeOD) δ 166.14, 151.89, 143.13, 107.93, 102.86, 94.51, 93.00, 91.35, 91.06, 71.09, 70.96, 60.33, 50.04. ^{19}F NMR (471 MHz, MeOD) δ -197.93, -197.97, -197.98, -198.02, -198.05, -198.09, -198.09, -198.13. HRMS (ESI$^+$) m/z calcd for $C_{10}H_{13}FN_2NaO_6$ $[M + Na]^+$ 299.0650, found 299.0647.

Table S8. NMR *N/S* ratios obtained at 303K for compound **5.**

Temperature (K)	N/S Ratio
303	87.0

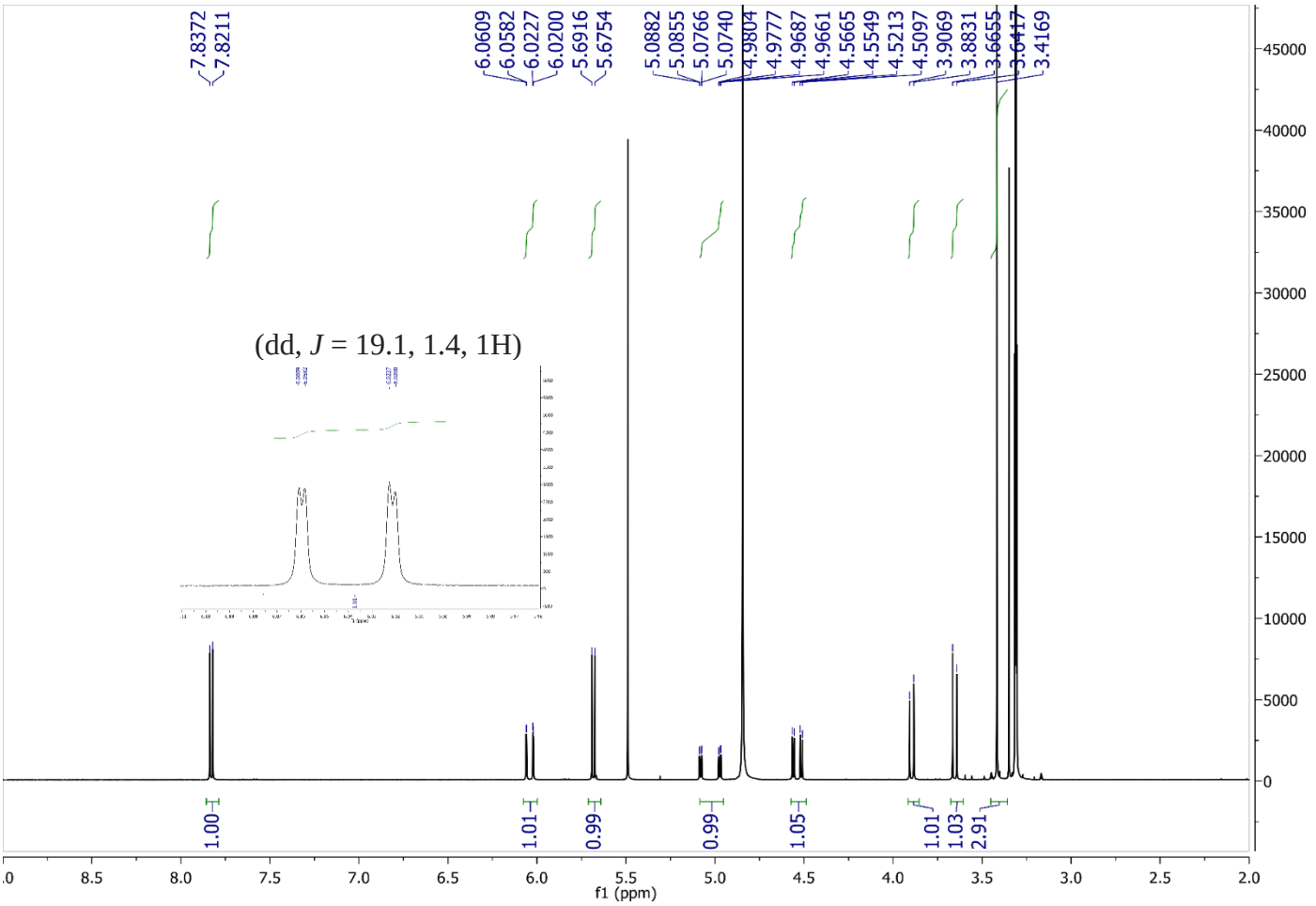


Figure S43. NMR spectrum of **5** (in MeOD) at 500MHz.

XV. NMR Prediction Data:

We employed a DFT/GIAO approach to predict the NMR spectra of our lowest energy conformations (obtained from QM/MM) for **5** to assess the accuracy of our modelling. Our chosen level of theory was PBE0/pcSseg-2 in the gas phase, which has been shown to give reliable results in the literature. The pcSseg-2 basis set is specifically optimized for calculating NMR shieldings. Since the bases of the structures obtained during QM/MM are not fully optimized, we optimized them while constraining the rest of the system to preserve the sugar conformation and substituent positions. In addition, 6 representative structures were extracted from the QM/MM calculations at P values of 18, 58.5, 100, 168.5, 233 and 305° and the bases were optimized at the PBE0/pcSseg-2 level of theory, followed by NMR prediction.

Table S9. Absolute energies of the conformations for which the base was optimized. Level of theory: PBE0/pcSseg-2.

Conformation	P	Energy (Eh)
1	18	-1048.82662
2	58.5	-1048.81854
3	100	-1048.82283
4	168.5	-1048.82475
5	233	-1048.80999
6	305	-1048.80980
TMS	-	-448.94231

We tested whether there was a significant difference between the predicted NMR spectrum using a linear combination of the two lowest-energy conformations (Figure S44) or whether using a linear combination of more conformations (six) would lead to a better prediction (Figure S45). In our case, the NMR spectrum predicted using the two lowest-energy conformations had a lower RMSE than the six conformations one, however the difference between the two RMSE values is only 0.05 ppm.

The % populations of each conformation used to predict the spectra in Figures S44-45 are given in Table S10. The absolute chemical shifts (ppm) are given in Table S11 for all 6 conformations and TMS, while the relative chemical shifts (ppm) are given in Table S12 for all 6 conformations and TMS.

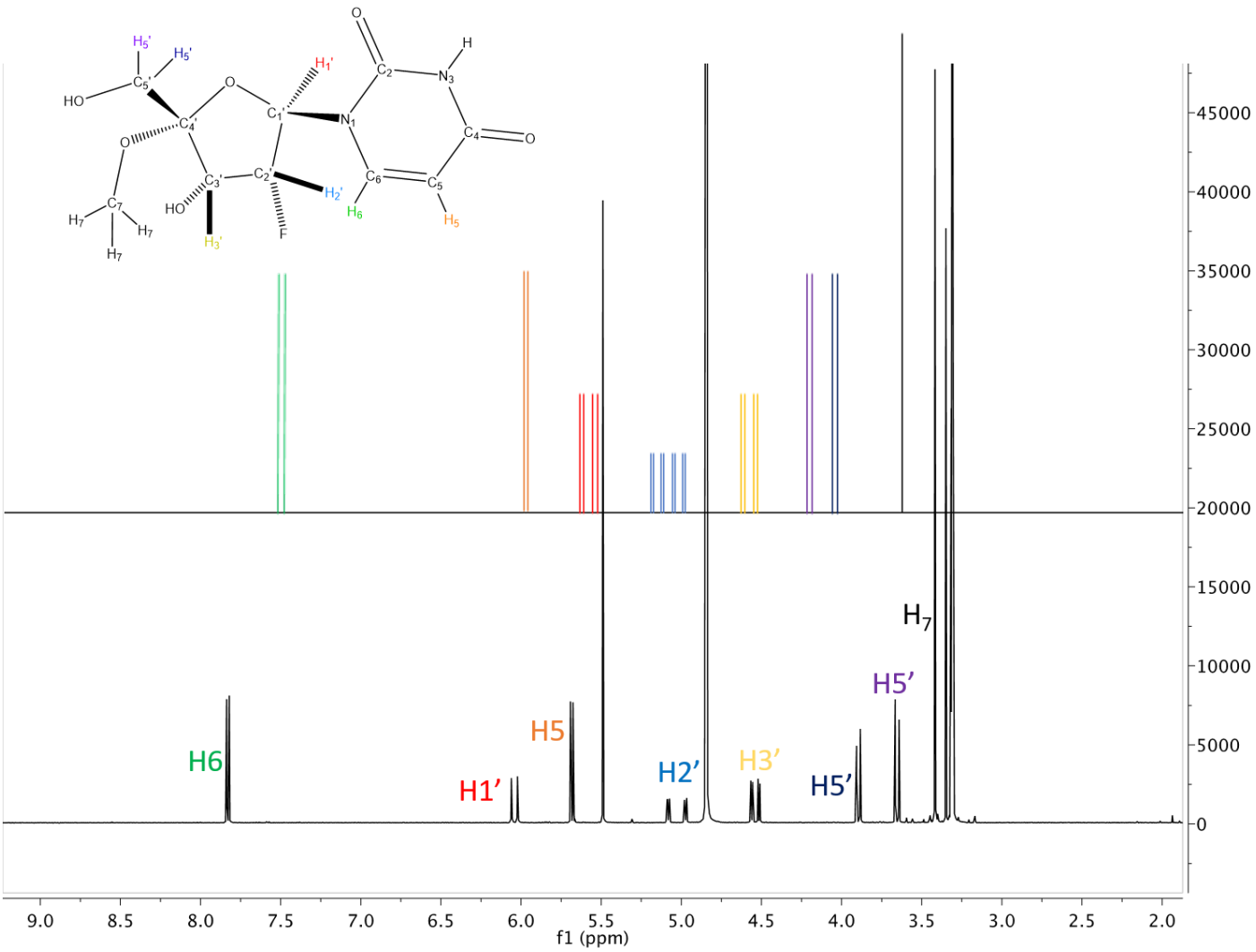


Figure S44. Superposition of the predicted NMR spectrum for the 84% *N* + 16% *S* linear combination (top) and the experimental NMR for **5**. RMSE = 0.34ppm.

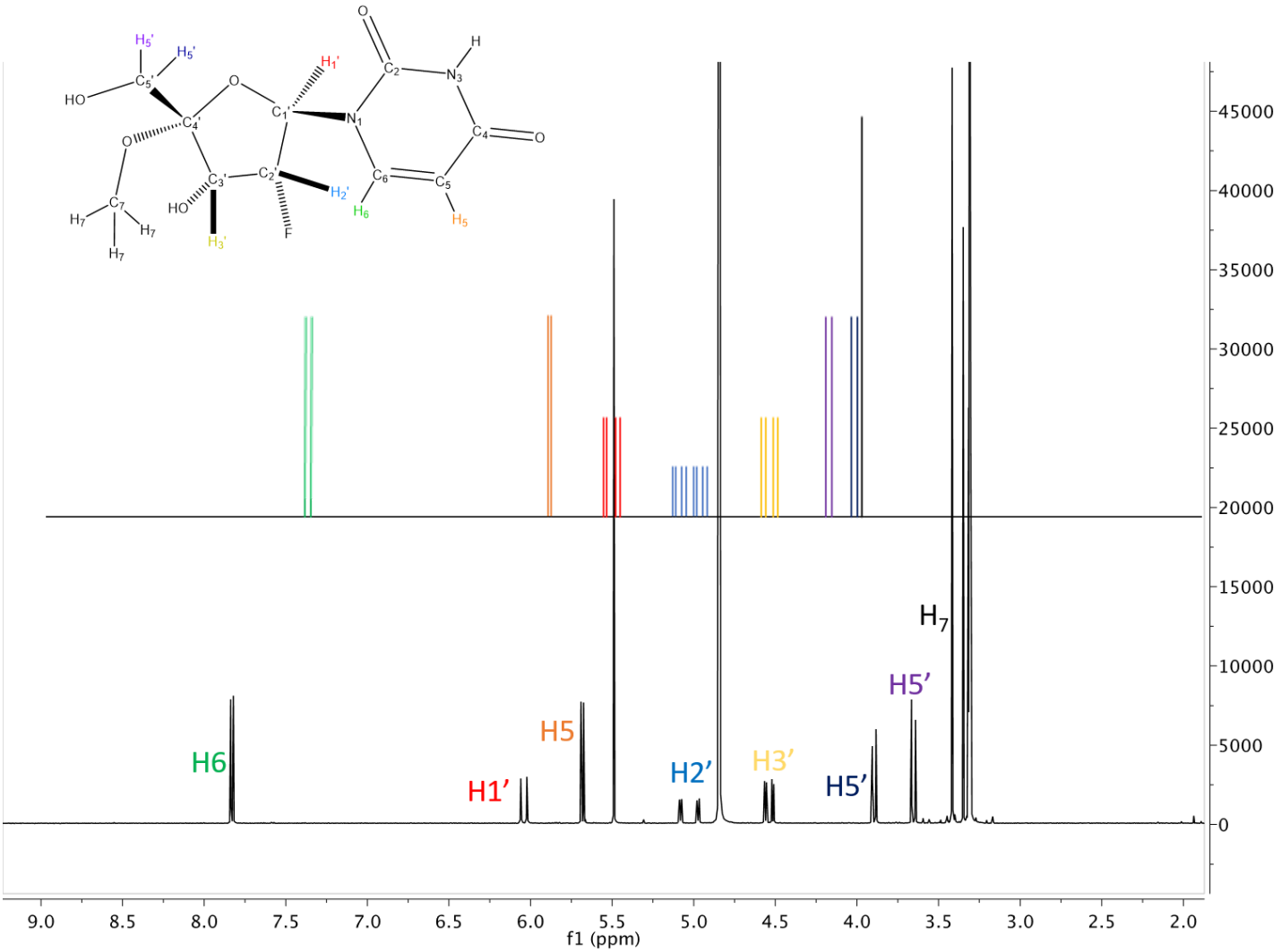


Figure S45. Superposition of the predicted NMR spectrum for the linear combination of 6 conformations (top) and the experimental NMR for **5.** RMSE = 0.39ppm.

Table S10. % population for each conformation used in assembling the linear combinations for predicting the ^{1}H NMR.

Conformation	Pseudorotational Angle Value (°)	%Population - 2 Conformation Prediction	%Population – 6 Conformation Prediction
1	18	-	7.7
2	58.5	84	75.1
3	100	-	8.3
4	168.5	16	7
5	233	-	0.6
6	305	-	1.3
TMS	0	-	-

Table S11. Absolute proton shifts (ppm) obtained for each conformation used for predicting the ^{1}NMR, as well as TMS.

Proton	Conf. 1 (ppm)	Conf. 2 (ppm)	Conf. 3 (ppm)	Conf. 4 (ppm)	Conf. 5 (ppm)	Conf. 6 (ppm)
H3’	25.705	26.818	27.351	27.201	25.837	26.215
H5’	27.480	27.526	27.613	26.946	27.372	27.169
H5’	27.355	27.242	27.870	27.731	27.488	26.887
H7	27.728	27.808	27.784	27.886	27.889	28.308
H6	24.359	24.309	24.06	23.044	23.138	22.783
H5	25.754	25.721	25.661	25.700	25.637	25.794
H1’	25.778	26.256	24.742	24.685	24.559	24.014
H2’	25.790	26.384	26.479	26.190	26.350	26.401
TMS	31.545	31.545	31.545	31.545	31.545	31.545

Table S12. Relative proton shifts (ppm) obtained for each conformation used for predicting the ^{1}NMR using TMS as a reference.

Proton	Conf. 1 (ppm)	Conf. 2 (ppm)	Conf. 3 (ppm)	Conf. 4 (ppm)	Conf. 5 (ppm)	Conf. 6 (ppm)
H3'	5.840	4.727	4.194	4.344	5.708	5.330
H5'	4.065	4.019	3.932	4.599	4.173	4.376
H5'	4.190	4.303	3.675	3.814	4.057	4.658
H7	3.817	3.373	3.761	3.659	3.656	3.237
H6	7.186	7.236	7.485	8.501	8.407	8.762
H5	5.791	5.824	5.884	5.845	5.908	5.751
H1'	5.767	5.289	6.803	6.860	6.986	7.531
H2'	5.755	5.161	5.066	5.355	5.195	5.144
TMS	0.000	0.000	0.000	0.000	0.000	0.000

XVI. References

1. Case, D.; Darden, T.; Cheatham III, T.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Walker, R.; Zhang, W.; Merz, K., *University of California: San Francisco* **2012**.
2. Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G., *Phys. Rev. B* **1998,** *58* (11).
3. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E., *Proteins* **2010,** *78* (8), 1950-1958.
4. Gordon, M. S.; Schmidt, M. W., Advances in electronic structure theory: GAMESS a decade later. In *Theory and Applications of Computational Chemistry: the first forty years*, Dykstra, C. E.; Frenking, G.; Kim, K. S.; Scuseria, G. E., Eds. Elsevier: 2005; pp 1167-1189.
5. Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., *J. Comp. Chem.* **1993,** *14* (11), 1347-1363.
6. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., *J. Chem. Phys.* **1983,** *79* (2), 926-935.
7. Adelman, S.; Doll, J., *J. Chem. Phys.* **1976,** *64* (6), 2375-2388.
8. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J., *J. Comp. Phys.* **1977,** *23* (3), 327-341.
9. Darden, T.; York, D.; Pedersen, L., *J. Chem. Phys.* **1993,** *98* (12), 10089-10092.
10. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., *J. Chem. Phys.* **1995,** *103* (19), 8577-8593.
11. Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A., *J. Comp. Chem.* **1992,** *13* (8), 1011-1021.
12. Grossfield, A., WHAM: the weighted histogram analysis method, version 2.0.2., **2008.**

13. Neese, F., *Wiley Interdisc. Rev. Comp. Mol. Sci.* **2012,** *2* (1), 73-78.
14. Weinhold, F., *J. Comp. Chem.* **2012,** *33* (30), 2363-2379.
15. Glendening, E.; Badenhoop, J.; Reed, A.; Carpenter, J.; Bohmann, J.; Morales, C.; Landis, C.; Weinhold, F., *Theoretical Chemistry Institute, University of Wisconsin, Madison, WI* **2013**.
16. Zhao, Y.; Truhlar, D. G., *Theor. Chem. Acc.* **2008,** *120* (1), 215-241.
17. Weigend, F.; Ahlrichs, R., *Phys. Chem. Chem. Phys.* **2005,** *7* (18), 3297-3305.
18. Varetto, U., Molekel 5.4. 0.8. Swiss National Supercomputing Centre, Manno, Switzerland, **2009**.
19. Adamo, C.; Barone, V., *Journal Chem Phys.* **1999,** *110* (13), 6158-6170.
20. Jensen, F., *J. Chem. Theor. Comput.* **2006,** *2* (5), 1360-1369.
21. Malek-Adamian, E.; Guenther, D. C.; Matsuda, S.; Martinez-Montero, S.; Zlatev, I.; Harp, J.; Burai Patrascu, M.; Foster, D.; Fakhoury, J.; Perkins, L.; Manoharan, R. M.; Taneja, N.; Bisbe, A.; Charisse, K.; Maier, M.; Kallanthottathil, R. G.; Egli, M.; Manoharan, M.; Damha, M. J., *under revision*.
22. Sun, G.; Voigt, J. H.; Filippov, I. V.; Marquez, V. E.; Nicklaus, M. C., *J. Chem. Inf. Comput. Sci.* **2004,** *35* (48), 1752-1762.
23. Klamt, A.; Schüürmann, G., *J. Chem. Soc. Perkin Trans 2* **1993,** (5), 799-805.
24. Islam, S. M.; Richards, M. R.; Taha, H. A.; Byrns, S. C.; Lowary, T. L.; Roy, P.-N., *J. Chem. Theor. Comput.* **2011,** *7* (9), 2989-3000.
25. Koole, L. H.; Buck, H. M.; Nyilas, A.; Chattopadhyaya, J., *Can. J. Chem.* **1987,** *65* (9), 2089-2094.
26. Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J., *J. Comp. Chem.* **2006,** *27* (10), 1101-1111.
27. Stewart, J. J., *J. Comp. Chem.* **1989,** *10* (2), 209-220.
28. McNamara, J. P.; Muslim, A.-M.; Abdel-Aal, H.; Wang, H.; Mohr, M.; Hillier, I. H.; Bryce, R. A., *Chem. Phys. Lett.* **2004,** *394* (4), 429-436.
29. Stewart, J. J. P., *J. Mol. Model.* **2007,** *13* (12), 1173-1213.
30. Barnett, C. B.; Naidoo, K. J., *J. Phys. Chem. B* **2010,** *114* (51), 17142-17154.
31. De Leeuw, F. A. A. M.; Altona, C., *J. Comp. Chem.* **1983,** *4* (3), 428-437.
32. Houseknecht, J. B.; Lowary, T. L.; Hadad, C. M., *J. Phys. Chem. A* **2003,** *107* (30), 5763-5777.
33. Evdokimov, A.; Gilboa, A. J.; Koetzle, T. F.; Klooster, W. T.; Schultz, A. J.; Mason, S. A.; Albinati, A.; Frolow, F., *Acta Cryst. Sec. B* **2001,** *57* (2), 213-220.
34. Sun, G.; Voigt, J. H.; Filippov, I. V.; Marquez, V. E.; Nicklaus, M. C., *J. Chem. Inf. Comput. Sci.* **2004,** *35* (48), 1752-1762.
35. Altona, C. T.; Sundaralingam, M., *J. Am. Chem. Soc.* **1972,** *94* (23), 8205-8212.
36. Petrová, M.; Páv, O.; Buděšínský, M.; Zborníková, E.; Novák, P.; Rosenbergová, Š.; Pačes, O.; Liboska, R.; Dvořáková, I.; Šimák, O.; Rosenberg, I., *Org. Lett.* **2015,** *17* (14), 3426-3429.
37. Islam, S. M.; Roy, P.-N., *J. Chem. Theor. Comput.* **2012,** *8* (7), 2412-2423.
38. Houseknecht, J. B.; Lowary, T. L.; Hadad, C. M., *J. Phys. Chem. A* **2003,** *107* (3), 372-378.
39. Evdokimov, A. G.; Martin, J. M. L.; Kalb, A. J., *J. Phys. Chem. A* **2000,** *104* (22), 5291-5297.
40. Evdokimov, A.; Gilboa, A. J.; Koetzle, T. F.; Klooster, W. T.; Martin, J. M. L., *J. Phys. Chem. A* **1999,** *103* (744).

41. Yamada, K.; Wahba, A. S.; Bernatchez, J. A.; Ilina, T.; Martínez-Montero, S.; Habibian, M.; Deleavey, G. F.; Götte, M.; Parniak, M. A.; Damha, M. J., *ACS Chem. Biol.* **2015,** *10* (9), 2024-2033.
42. Martínez-Montero, S.; Deleavey, G. F.; Dierker-Viik, A.; Lindovska, P.; Ilina, T.; Portella, G.; Orozco, M.; Parniak, M. A.; González, C.; Damha, M. J., *J. Org. Chem.* **2015,** *80* (6), 3083-3091.
43. Morita, K.; Takagi, M.; Hasegawa, C.; Kaneko, M.; Tsutsumi, S.; Sone, J.; Ishikawa, T.; Imanishi, T.; Koizumi, M., *Bioorg. Med. Chem.* **2003,** *11* (10), 2211-2226.
44. Lipnick, R. L.; Fissekis, J. D., *Biochim. Biophys. Acta Nucl. Acids Prot. Synth.* **1980,** *608* (1), 96-102.
45. Watts, J. K.; Sadalapure, K.; Choubdar, N.; Pinto, B. M.; Damha, M. J., *J. Org. Chem.* **2006,** *71* (3), 921-925.