PROSIT: A Pseudo-Rotational Online Service and Interactive Tool Applied to Conformational
Survey of Nucleosides and Nucleotides

Guangyu Sun, Johannes H. Voigt, Victor E. Marquez* and Marc C. Nicklaus*

* E-mail: mn1@helix.nih.gov Tel: (301)-846-5903 Fax: (301)-846-6033

Laboratory of Medicinal Chemistry, NCI-Frederick, NIH

376 Boyles St., Frederick, MD 21702

Supplementary Material:

Ab initio calculation on 2'-O,4'-C-methylene bridged nucleic acids

The synthetic 2'-O,4'-C-methylene bridged nucleic acids have a very large puckering amplitude $\nu_{\max}$ ¹. We performed quantum chemical calculations to confirm that the large puckering amplitude results from a significantly strained structural motif rather than from crystal packing forces.

Starting from the crystal structure in EBOZAW¹, we first obtained the fully optimized structure for compound **1**. Density functional theory (DFT) at the B3LYP/6-31G* level^{2,3} was employed in the gas phase calculations. All calculations were performed using the Gaussian 03 program package⁴. We calculated the dihedral angles ν_0 and ν_4 corresponding to a fixed P value and performed geometry optimizations at $\nu_{\max}$ from 42° to 69° with increments of 3°. Since only two out of the five endocyclic dihedral angles in a furanose ring are independent, we chose to fix ν_0 and ν_2 in these partial geometry optimizations.

The calculated bond lengths for compound **1** are in good agreement with experimental values, as shown in Table S1. The optimized structure of **1** has a pseudorotational phase angle P of 20.63° and a puckering amplitude $\nu_{\max}$ of 56.80°. These values are very close to those found in the crystal structure (16.67° and 57.59°, respectively).

Table S1: The experimental and calculated bond lengths (in Å) in compound **1**.^a

bonds	crystal	calculation	bonds	crystal	calculation
C1'-C2'	1.513	1.539	C1'-N1	1.498	1.479
C1'-O4'	1.418	1.430	N1-C2	1.380	1.423
C2'-C3'	1.524	1.536	C2-O2	1.246	1.233
C2'-O2'	1.436	1.440	C2-C3	1.428	1.446
C3'-C4'	1.535	1.551	C3-C4	1.340	1.366
C3'-O3'	1.401	1.406	C4-C5	1.398	1.423
C4'-O4'	1.458	1.460	C5-C6	1.345	1.365
C4'-C5'	1.498	1.510	C6-N1	1.368	1.367
C5'-O5'	1.421	1.420			
C4'-C6'	1.507	1.535			
C6'-O2'	1.453	1.456			

^a Experimental values measured from the X-ray single crystal structure of EBOZAW in CSD.

The structures obtained from partial geometry optimizations have the expected values of puckering amplitude $v_{\max}$, while the phase angle P changes slightly from the optimized value 20.63° to 21.63° at $v_{\max}=42^\circ$ and 19.58° at $v_{\max}=69^\circ$. The slight changes in P result from the slight changes in the v_1 , v_3 and v_4 , which were fully relaxed. Large relative energies are predicted when the puckering amplitude is varied from 42° to 69° while the phase angle is fixed at the calculated optimal value of 20.63° . Figure S1 shows the potential energy surface of compound **1**. When the phase angle is fixed to 20.63° and the puckering amplitude is decreased from 56.80° to 42° , the energy increased by 8.56 kcal/mol. When the puckering amplitude is increased to 69° , the energy increased by 6.62 kcal/mol. The large energies required in the gas phase to either reduce or enlarge $v_{\max}$ indicate that the large $v_{\max}$ observed in the single crystal structure is intrinsic to the molecular structure itself and not due to crystal packing forces.

Figure S1. Potential energy surface of compound **1** calculated by DFT at the B3LYP/6-31G* level of theory with the puckering amplitude $v_{\max}$ varying from 42° to 69° .

High-resolution structures of DNA and RNA

The search of high-resolution crystal structures of DNA and RNA duplexes with ten or more residues and 1.55 Å or higher resolution in the Nucleic Acid Database (NDB)⁵ returned 72 entries. The entries are listed in Table S2.

Table S2: The crystal structures of DNA and RNA duplexes with ten or more residues and 1.55 Å or higher resolution in NDB.

AD0003	AD0026	BD0013	BD0036	BDL084	PD0341
AD0004	AD0027	BD0014	BD0038	BDLB84	PR0037
AD0007	AD0028	BD0018	BD0041	BDLB85	UD0019
AD0012	AD0031	BD0021	BD0044	DD0058	UD0026
AD0013	AH0007	BD0023	BD0054	DR0005	UDJ032
AD0014	AH0012	BD0024	BD0060	PD0068	UR0006
AD0015	AR0020	BD0025	BDD001	PD0101	UR0007
AD0016	AR0021	BD0026	BDJ008	PD0111	UR0020
AD0018	BD0006	BD0030	BDJ019	PD0121	UR0026
AD0019	BD0007	BD0033	BDJ025	PD0253	URL064
AD0020	BD0008	BD0034	BDJ031	PD0254	URX075
AD0025	BD0012	BD0035	BDJB44	PD0312	ZD0006

References:

- 1 Obika, S.; Hari, Y.; Sekiguchi, M.; Imanishi, T. A 2',4'-bridged nucleic acid containing 2-pyridone as a nucleobase: Efficient recognition of a C·G interruption by tripler formation with a pyrimidine motif. *Angew. Chem.-Int. Edit.* **2001**, *40*, 2079-2081.
- 2 Becke, A. D. Density-Functional Thermochemistry 3. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648-5652.
- 3 Lee, C. T.; Yang, W. T.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. *Phys. Rev. B* **1988**, *37*, 785-789.
- 4 Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J., J. A.; Vreven, T.; Kudin, K. N. et al. *Gaussian 03*; Gaussian, Inc., Pittsburgh PA 2003.
- 5 Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S. H.; Srinivasan, A. R.; Schneider, B. The Nucleic-Acid Database - a Comprehensive Relational Database of 3-Dimensional Structures of Nucleic-Acids. *Biophys. J.* **1992**, *63*, 751-759.