ACS central science

http://pubs.acs.org/journal/acscii

Research Article

Observation of the Unbiased Conformers of Putative DNA-Scaffold Ribosugars

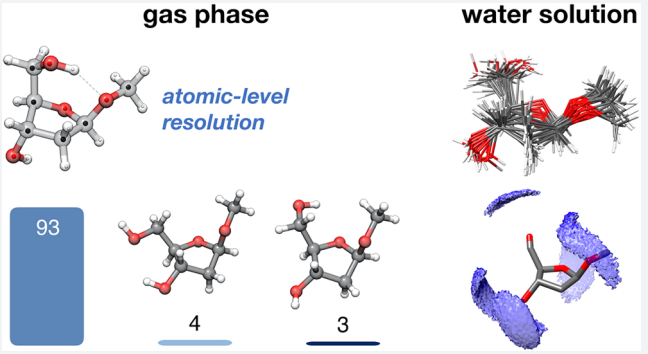
Camilla Calabrese, Iciar Uriarte, Aran Insausti, Montserrat Vallejo-López, Francisco J. Basterretxea, Stephen A. Cochrane, Benjamin G. Davis,* Francisco Corzana,* and Emilio J. Cocinero*

Cite This: *ACS Cent. Sci.* 2020, 6, 293−303

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The constitution, configuration, and flexibility of the core sugars of DNA molecules alter their function in diverse roles. Conformational itineraries of the ribofuranosides (*f*s) have long been known to finely determine rates of processing, yet we also know that, strikingly, semifunctional DNAs containing pyranosides (*p*s) or other configurations can be created, suggesting sufficient but incompletely understood plasticity. The multiple conformers involved in such processes are necessarily influenced by context and environment: solvent, hosts, ligands. Notably, however, to date the *un*biased, "naked" conformers have not been experimentally determined. Here, the inherent conformational biases of DNA scaffold deoxyribosides in unsolvated and solvated forms have now been defined using gas-phase microwave and solution-phase NMR spectroscopies coupled with computational analyses and exploitation of critical differences between natural-abundance isotopologues. Serial determination of precise, individual spectra for conformers of these 25 isotopologues in alpha (α-D) and beta (β-D); pyrano (*p*) and furano (*f*) methyl 2-deoxy-D-ribosides gave not only unprecedented atomic-level resolution structures of associated conformers but also their quantitative populations. Together these experiments revealed that typical ^{2}E and ^{3}E conformations of the sugar found in complex DNA structures are not inherently populated. Moreover, while both OH-5′ and OH-3′ are constrained by intramolecular hydrogen bonding in the unnatural α*f* scaffold, OH-3′ is "born free" in the "naked" lowest lying energy conformer of natural β*f*. Consequently, upon solvation, unnatural α*f* is strikingly less perturbable (retaining $^{2}T_{1}$ conformation *in vacuo* and water) than natural β*f*. Unnatural α*p* and β*p* ribosides also display low conformational perturbability. These first experimental data on inherent, unbiased conformers therefore suggest that it is the background of conformational flexibility of β*f* that may have led to its emergence out of multiple possibilities as the sugar scaffold for "life's code" and suggest a mechanism by which the resulting freedom of OH-3′ (and hence accessibility as a nucleophile) in β*f* may drive preferential processing and complex structure formation, such as replicative propagation of DNA from 5′-to-3′.



INTRODUCTION

Structural variability and flexibility of ribonucleic acids are not only apparent and immense in scope but also intimately linked to both the existence[1−3] and emergence[4,5] of biological function. Moreover, ever-expanding interest in the design and use of both natural and unnatural ribonucleotides in diagnostic and therapeutic applications continues to highlight a key role for an understanding of the fundamentals that generate associated structural populations.[6−8] For example, while chemical modifications at phosphate[9] or nucleobase[10] can usefully increase *in vivo* stability (reduced *reactivity*), it is also the correct manipulation of the conformations of the core sugar scaffold that has proven key to optimal functional *activity*. In DNA polymerases, an essential factor that prevents improper inclusion and extension of nucleotides appears to be governed by the preferred conformations of the furanose moiety of each incoming nucleotide during both incorporation and extension.[11,12] Interestingly, as the pioneering work of Eschenmoser[13] highlighted, there is also no necessity for ribosidic or even furanosidic structures, and alternative polynucleotides can be constructed based on, for example, L-*threo*-furanosides[14] or even configurationally varied pyranosides.[13] While their functions are typically moderated (e.g., reduced base-pairing strengths), such altered-sugar polynucleotides can still adopt relevant duplex structures via typical (e.g., Watson−Crick) patterns and can even be processed by appropriate variant enzymes, albeit at reduced rates.

Received: December 14, 2019
Published: February 18, 2020

https://dx.doi.org/10.1021/acscentsci.9b01277
ACS Cent. Sci. 2020, 6, 293−303

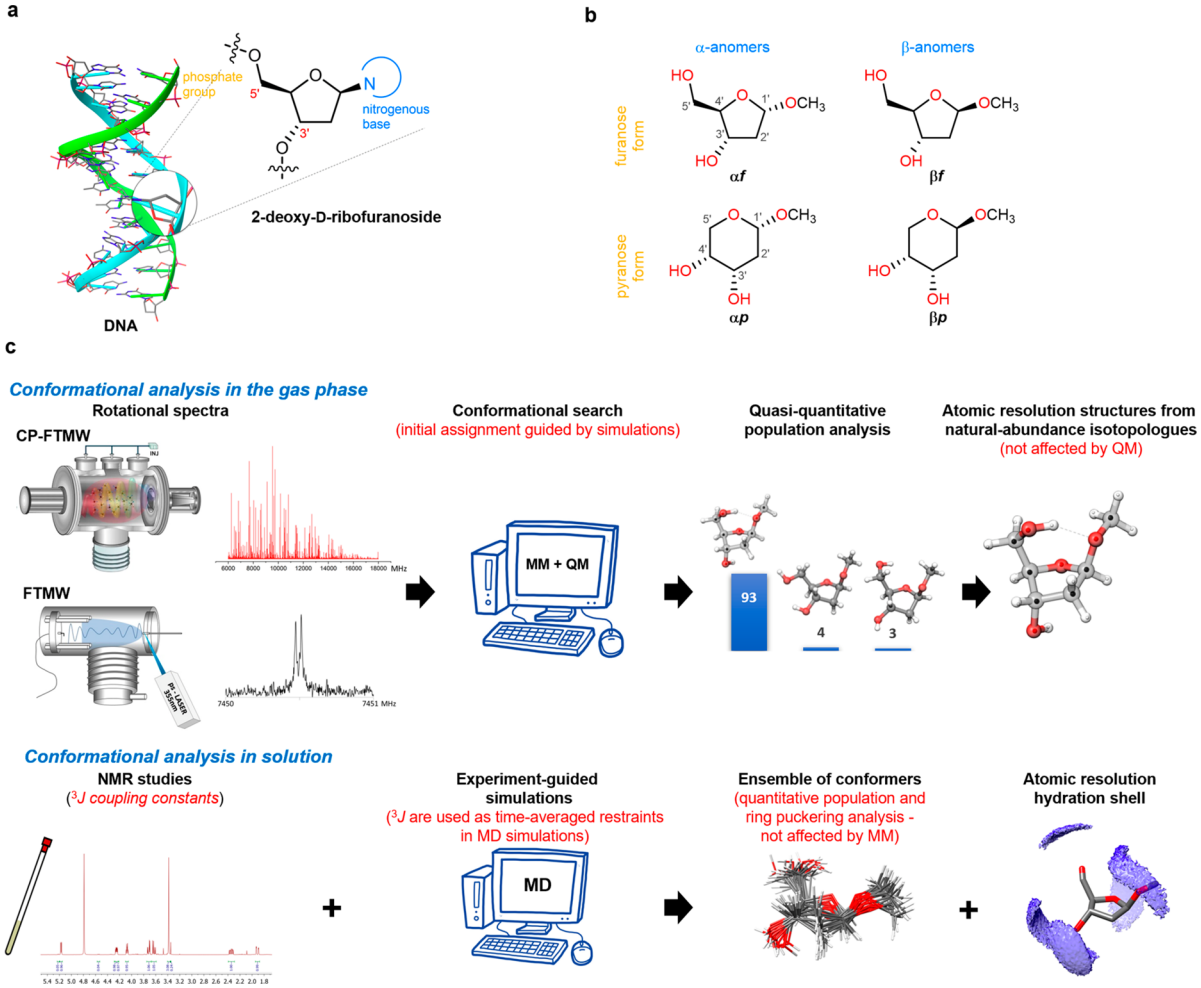


Figure 1. Coherent conformational analyses and comparison of DNA-scaffold ribosugars. (a) Schematic representation of the structure of dominant B-form DNA. (b) Compounds studied in this work: two 2-deoxyribofuranose forms (compounds *αf* and *βf*) and two corresponding pyranose forms (derivatives *αp* and *βp*). (c) Overview of the precise, comparative protocol used in this work to determine D-ribosides in the gas phase (upper panel) and in aqueous solution (lower panel).

Together, such examples illustrate that both the fuller understanding of natural nucleotide function and the development of useful unnatural nucleotides require a deep insight into associated conformational preferences and, in particular, those of the native sugar moiety, 2-deoxy-D-riboside. However, no furanoside structures to date have been determined under fully isolated (and hence artefact free) conditions. Crystallographic data can be biased by both solvent effects and crystal-packing interactions. Spectroscopy in the solution phase is similarly dominated by environmental contacts.[15] Indeed, extensive analyses performed on nucleosides and nucleotides, as well as in DNA, suggest that the sugar cores of such structures often exist in a dynamic equilibria between multiple conformations in solution.[16]

Here we present a strategy for the complete structural analysis of the core sugar scaffold of DNA, 2-deoxy-D-riboside (Figure 1a,b), that exploits custom-made, high-resolution microwave spectrometers combined with complementary vaporization and sampling techniques in the gas phase. Importantly, while associated quantum chemistry methods have been integral to prior analyses/interpretations of most gas phase structures, these microwave methods do not require quantum chemical computations during structural determination. Instead, natural abundance isotopologues act as key structural references to allow atomic-level resolution. The resulting structures are therefore critically independent of the quality or processes of any associated quantum mechanical models. Comparison of these first unbiased conformational analyses with those affected by solution was then achieved by combining NMR experiments with experiment-guided molecular dynamics (MD) simulations (Figure 1c). This allows precise study of the inherent intramolecular interactions responsible for the structural diversity found in the core of DNA without interference from surrounding molecules in condensed media (e.g., water as solvent).

RESULTS

Design of a System for Generating and Analyzing Gaseous, Isolated DNA Scaffolds. Current understanding of DNA structure is based in significant part on X-ray crystallography data,[17−20] providing information concerning global helical structure and the geometry of local features, such as base-pair stacking patterns and backbone conformation driven by that structure. However, such data can not only be biased by both solvent effects and crystal-packing interactions; it also provides only static structures. Conformational analyses of DNA fragments in solution can overcome some of these problems.[21] However, the lack of sequence variety (dominated

https://dx.doi.org/10.1021/acscentsci.9b01277
ACS Cent. Sci. 2020, 6, 293−303

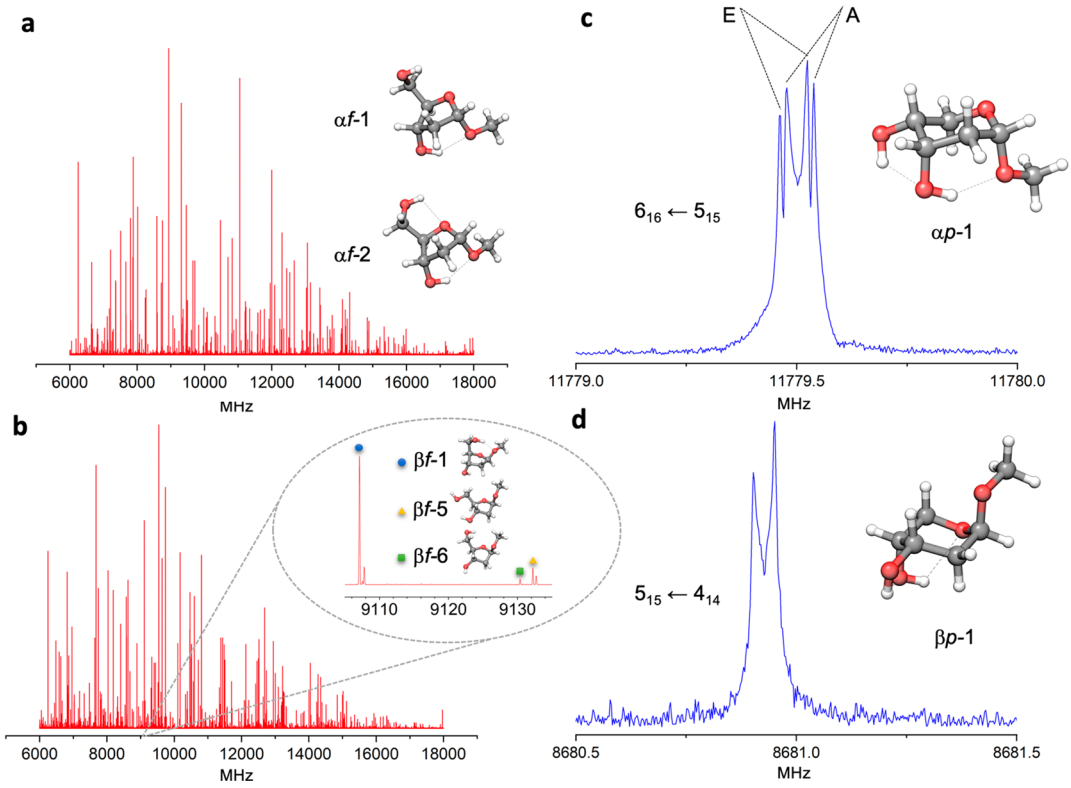


Figure 2. High-resolution rotational spectra of *αf*, *βf*, *αp*, and *βp*. (a and b) Overview of the CP-FTMW rotational spectra of *α* and *β* deoxyriboside anomers *αf*, *βf* in the 6−18 GHz region, respectively. In b, an expanded view highlights rotational transitions of each of the *β*-conformers. (c and d) Typical rotational transitions of *αp* and *βp* were observed using FTMW spectroscopy equipped with a UV ultrafast laser vaporization system. Rotational transitions showed hyperfine splitting due to internal rotation of the methyl group. Both transitions (c and d) are additionally split by the instrumental Doppler effect.

in general by A-tracts and the Drew−Dickerson dodecamer) combined with limitations in refinement protocols (sometimes biased by the quality of calculations used to interpret experimentally derived restraints[22]) has not yet delivered a data set of structures with consistent information on the intrinsic mechanics of DNA in solution.[23] For instance, in some structures, significant deviations are seen from similar NMR data when refined with different backbone restraints,[24,25] and in most studies conformational analyses rely on highly variable two-state models.[26,27] Consequently, to date, no precise (atomic resolution) structures of DNA nor the biases of its scaffold have yet been reported.

In previous pioneering studies, conformational analyses of smaller carbohydrates fragments were accomplished using gas phase vibrational laser spectroscopy.[28−31] However, these studies typically demand the use of molecules tagged with a chromophore that can be a potential cause of undesirable artefacts. While these can be replaced by external probes (e.g., toluene[32,33] or peptides[34]), this alternative methodology does not fully remove the need for chromophore and can induce additional unwanted environmental interference. Moreover, the method is critically dependent on quantum mechanical (QM) models to relate determined spectra to derived structures.

In this context, gas phase rotational spectroscopy emerges as a unique method that can avoid such artifacts. In particular, observation of multiple rotational spectra from isotopologues (e.g., where ^{13}C replaces ^{12}C) allows direct structural determination without dependence on QM models or chromophores. In addition, it possesses a superior inherent resolution due to unsurpassed frequency resolution (∼kHz) and unrivalled chemical (conformers, tautomers, isotopologues, even enantiomers) discrimination.[35,36] However, the ability to measure isotopic species directly requires sufficient sensitivity (or enriched samples), which in turn demands sufficient partial pressure of the analyte in the gas phase. For analytes, such as carbohydrates, that have low volatility and are thermolabile, this has proven challenging using traditional heating methods[37] or even IR nanosecond-pulsed laser vaporization techniques.[38] Recently, we have shown that UV (355 nm) picosecond-pulsed (∼40 ps) laser vaporization can allow observations of certain reducing sugars,[39−42] but their mutarotation[39,40] and/or lack of sensitivity[39,41] prevented either observation of biologically relevant constitutional forms or structures with atomic resolution; consequently, observation of core DNA scaffold 2-deoxy-D-ribosides (Figure 1b) was not previously possible.

We reasoned that the required critical increase in sensitivity might be gained in several ways. This proved successful through the custom design (Supplementary Figures 1 and 2) of microwave apparatus that included a Fourier-transform microwave-spectrometer (FTMW)[43] coupled to a UV ultrafast laser vaporization system[39] and a set-up with chirped-pulses (CP-FTMW) (Figure 1c).[44] Specifically, the latter system allowed: (a) simultaneous three-nozzle gas injection; (b) longer acquisition times (allowing even spectral registration for 1 week continuously averaging ∼40 million of rotational spectra); and (c) more efficient radiation pulse sequences (30 microwave pulses per molecular pulse). These features together provided an increase of an order of magnitude (or

https://dx.doi.org/10.1021/acscentsci.9b01277
ACS Cent. Sci. 2020, 6, 293−303

Table 1. Experimental Rotational Constants of the Observed Conformers of *αf*, *βf*, *αp*, and *βp*

	αf-1	*αf*-2	*βf*-1	*βf*-5	*βf*-6	*αp*-1	*βp*-1
A/MHz[a]	2056.39188(60)[d]	1882.30839(76)	1890.65747(49)	1999.87647(81)	1481.77004(72)	2153.8853(66)	2358.67704(36)
B/MHz	1019.67311(21)	1024.81006(26)	1145.05589(21)	991.81629(38)	1256.10078(60)	1058.89820(11)	1001.97250(15)
C/MHz	860.52781(23)	804.74073(31)	910.72056(22)	771.18931(31)	784.89436(42)	960.74320(13)	831.19774(12)
N[b]	95	99	140	52	86	55	90
σ[c]/kHz	8.2	9.3	10.5	6.7	7.4	2.3	3.5

[a]Rotational constants (*A*, *B*, *C*). [b]Number of rotational transitions (*N*). [c]Root-mean-square (rms) deviation (*σ*) of the fit. [d]Standard errors in units of the last digit.

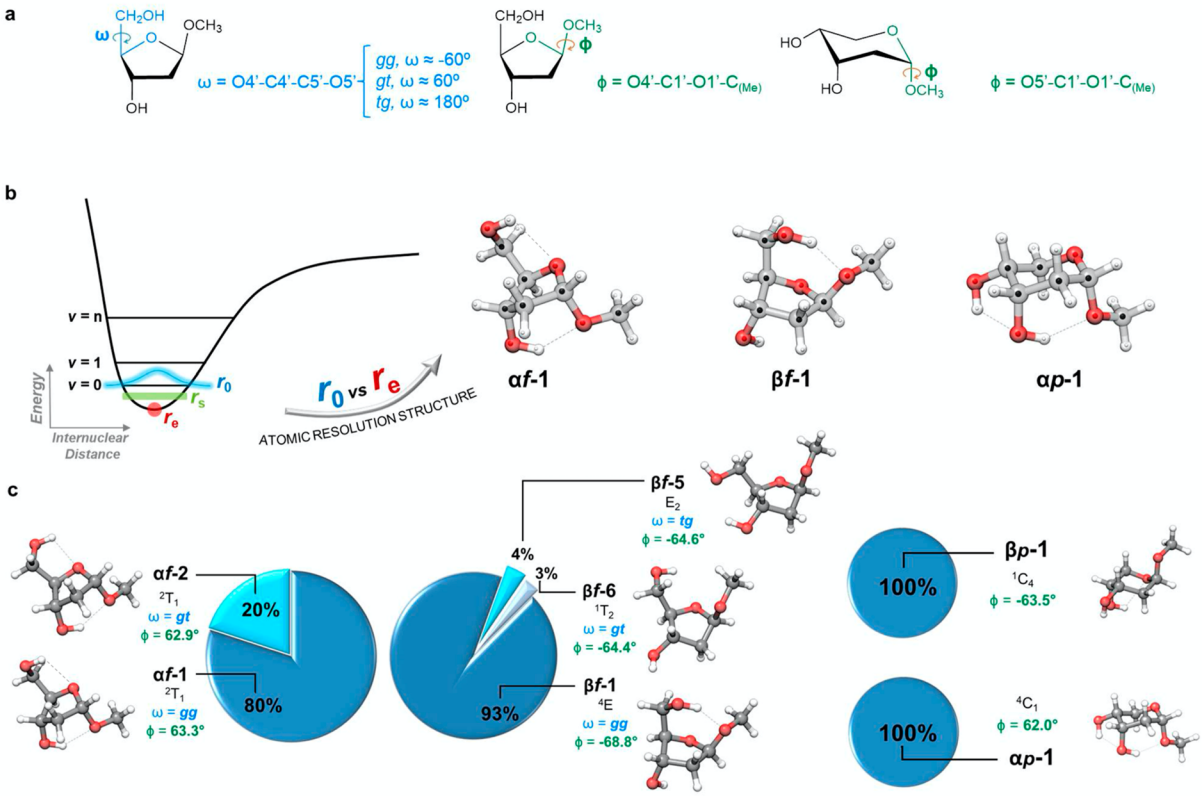


Figure 3. Observed conformers of *αf*, *βf*, *αp*, and *βp* in the gas phase. (a) Definition of relevant torsional angles. (b) Conceptualization of gas phase molecular structure-determination methods. Experimental molecular structures could be determined for *αf*-1, *βf*-1, and *αp*-1 with atomic resolution due to the observation of isotopologues in natural abundance. (c) Conformers detected for furanosides and pyranosides, together with their population and key geometrical parameters.

greater) sensitivity, which in turn allowed lower consumption of the sample (down to 0.7−2.0 g of synthetic samples). In this way, even each of the multiple isotopologues present in samples at only natural abundance ($^{13}C^{12}C_5H_{12}O_4 \approx 1.1\%$) could be detected and characterized. These advances allowed us to determine, for the first time, "nonbiased" atomic resolution structures of core DNA scaffold 2-deoxy-D-ribosides.

Experimental Determination of Unbiased Structures of DNA Scaffold Sugars. First, core putative scaffolds *αf*, *βf*, *αp*, and *βp* (Figure 1b) were readily synthesized by complementary methods on the multigram scales required. Briefly (see Supplementary Methods for further details), in one method 2-deoxy-D-ribose was subjected to classical Fischer glycosylation conditions (1% HCl in methanol); the equilibrium mixture after 18 h (57% *βp*, 15% *αp*, 11% *βf*, 17% *αf*) yielded *αp* by column chromatography and *βp* by further crystallization. To generate *αf* and *βf* on scale, we preferentially employed brief (15 min) treatment, giving an ∼1:1 mixture of essentially only furanosides; subsequent peracetylation/deacetylation allowed clean isolation via column chromatography (33% *βf*, 36% *αf*).

Next, gas-phase DNA scaffold generation was tested under various modes (Figures 1c and 2). Of these, CP-FTMW spectroscopy allowed collection of rotational spectra of *αf* and *βf*. (Figure 2, panels a and b, respectively). An FTMW spectrometer coupled to UV ultrafast laser vaporization proved most successful for *αp* and *βp*. Together, these experiments allowed long acquisition times from gram-scale samples thereby yielding rotational spectra of the "nonbiased" structures of *αf*, *βf*, *αp*, and *βp* with an unprecedented signal-to-noise ratio (S/N for most intense observed transitions: >500/1 and >1800/1 for *αf* and *βf*, respectively). In turn, such S/N orders allowed determination of structures with atomic resolution. In this way, despite heavily congested spectra—*αf* (Figure 2a) and *βf* (Figure 2b) showed ∼600 and ∼1800 rotational lines, respectively—sets of rotational transitions were identified belonging to independent structures.

https://dx.doi.org/10.1021/acscentsci.9b01277
ACS Cent. Sci. 2020, 6, 293−303

Each transition set was fitted to a semirigid rotor Hamiltonian based on Watson's symmetric reduction and I^r representation to obtain *A*, *B*, and *C* rotational constants and critical centrifugal distortion constants.[45,46] Conformational assignments were guided and supported by (but not dependent on, *vide infra*) theoretical calculations, using a two-step MM then QM (DFT and *ab initio*) strategy[39,42] (see Supplementary Methods). Rotational constants can be correlated with moments of inertia, which are a fingerprint of 3D molecular structure (Table 1). Direct comparison of simulated and experimental rotational constants (Table 1 and Supplementary Tables 1, 3, 5, and 6) therefore allowed an unequivocal assignment.

Two *αf* conformers (named *αf*-1 and *αf*-2, numbered by calculated energy ranking, Figure 2a) and three different *βf* conformers (*βf*-1, *βf*-5, and *βf*-6, Figure 2b) were identified, whereas for *αp* and *βp* only single (dominant) conformers (Figure 2c,d) were detected (Table 1 and Supplementary Tables 1, 3, 5, and 6). Other low energy conformers as *βf*-2 or *βf*-3 were not detected, likely due to relaxation to the most stable *βf*-1 through interconversion pathways (see Supplementary Figures 3 and 4). Notably, in both *αp* and *βp*, even hyperfine-splitting due to internal rotation of the methyl group could be observed and analyzed (Figure 2c,d). The resulting experimental spectroscopic parameters and experimental frequencies provide an extensive "structural map" of all possible 2-deoxy-riboside sugar scaffolds (Supporting Information).

Notably, such was the sensitivity that it allowed detection of ^{13}C isotopes in natural abundance (~1.1%). Strikingly, we were able to observe the rotational spectra for all (18 additional $^{13}C^{12}C_5H_{12}O_4$ species) of the monosubstituted ^{13}C isotopologues distributed within this small abundance for the most populated conformers *αf*-1, *βf*-1, and *αp*-1. As noted, correlation of moments of inertia with rotational constants renders them a key source for structural information. Moments of inertia depend critically on distribution of atomic masses, and so molecular geometry, and are affected by vibrational energy even in the ground state, since molecules are not rigid systems. Except for simplified cases, the explicit correction of the moments of inertia for vibrational contributions has not been possible because of related experimental difficulties. Therefore, different procedures and/or evaluations of molecular structures were tested, all exploiting isotopologue spectra (Figure 3b). *Equilibrium structures* (r_e, Figure 3b) represent the hypothetical vibration-less state that corresponds to the minimum of the potential energy surface. This structure, essentially inaccessible to molecules, is the one obtained by typical computational methods. *Effective structures* (r_0, Figure 3b)[47] reproduce rotational constants in the ground state and were derived here starting from the calculated geometry and through iterative least-squares adjustments of experimental rotational constants of each deoxyriboside (21 = 3 for parent +6 × 3 for isotopologues). *Substitution structures* (r_S, Figure 3b),[48,49] obtained from the analysis of the changes of the moments of inertia resulting from a single isotopic substitution, allow determination of the atomic coordinates of the substituted atom. This "r_S method" has the advantage that it provides the position of the substituted atom free from other assumptions about molecular structure and so does not depend on computational calculations. Such r_S structures are generally assumed to be intermediate between r_0 and r_e.[45] For these reasons, such "real" experimental geometries (r_S and r_0) are not directly comparable with those obtained by typical computational methods (r_e). In this way, the accuracy of rotational spectroscopy yielded riboside structures with unprecedented resolution (Figure 3b and Supplementary Tables 8−13).

Analysis of Gas-Phase Rotational Transitions of Core DNA Scaffolds Reveals Conformers and their Inherent Populations. Rotational spectroscopy allowed not only determination of unbiased, experimentally accurate structures but also an estimation of the population ratios between these different observed conformers by analyzing the relative intensities of associated rotational transitions.[50] *αf* adopts preferably two conformers in the gas phase, *αf*-1 ($^2T_1$ *gg*, Figure 3) and *αf*-2 ($^2T_1$ *gt*), in a ratio *αf*-1/*αf*-2 of 80(20):20, whereas for *βf*, three different conformers were identified (Figure 3c): *βf*-1 (^4E *gg*), *βf*-5 (E_2 *tg*), and *βf*-6 ($^1T_2$ *gt*) in a 93(10):4(1):3 population ratio. Only one conformer was detected for either *αp* ($^4C_1$) and *βp* ($^1C_4$).

For *αf*, both observed conformers (*αf*-1 and *αf*-2) share the same ring shape $^2T_1$, stabilized by two intramolecular hydrogen bonds: O3′−H3′···O1′ and O5′−H5′···O4′ across both *α* and *β* faces. On the other side, *βf*-1 (*gg*) adopts a partly distorted ^4E conformation. This puckering is stabilized by a strong hydrogen bond O5′−H5′···O1′, which is viable only when the C5′-hydroxymethyl group adopts a *gg* arrangement. Notably, a "switched" loss of this stabilizing interaction in *βf*-5 (to *tg*) and *βf*-6 (to *gt*) causes them to adopt somewhat distorted puckered ring E_2 and $^1T_2$ conformations, respectively, similar to those found in methyl *β*-D-ribofuranoside.[41]

The pyranosides *αp* and *βp* were much more rigid in the gas phase, exhibiting single dominant conformers (Figure 3c). *αp* adopts a $^4C_1$ chair, while *βp* adopts the inverted $^1C_4$. In both, reinforcing hydrogen bond networks on the *α* face, akin to reinforcing hydrogen bonding networks seen previously using IR spectroscopy in hexopyranosides,[51] further drive this rigidity (for *αp*-1: O4′−H4′···O3′−H3′···O-1′ and for *βp*-1: O3′−H3′···O4′−H4′···O5′) beyond the known greater barriers to conformational interconversion found in pyranosides.

In all scaffolds, the orientation of the methyl group was in accordance with the *exo*-anomeric effect,[34,52] with values of ϕ close to either 60° or −60° for the *α*- (*αf*, *αp*) and *β*- (*βp*, *βf*) anomers, respectively. This serves to place the methyl group in a site of essentially low influence with regard to conformation. Strikingly, and in clear contrast to *α*-anomer *αf* and both pyranosides *αp* and *βp*, only in *βf* was the hydroxyl O3′-H3′ not engaged in hydrogen bonding—the lowest lying conformer of natural scaffold *βf* therefore uniquely frees OH-3′, and hence the *α*-face remains consequentially more accessible (and potentially reactive, *vide infra*).

Comparison of Inherent Conformers of DNA scaffolds with Solvated Structures Reveals Selectively Driven Conformational Change for 2-deoxy-*β*-D-ribofuranoside *βf*. DNA functions in both aqueous and partly hydrated (e.g., enzyme active site) environments. To study the effect of the hydration state on the structure of these core scaffolds derivatives, *αf*, *αp*, *βp*, and *βf* were subjected to conformational analysis in aqueous solution by combining NMR spectroscopy measurements ($^3J_{H,H}$ coupling constants as constraints) with molecular dynamics (MD) simulations. Calculations were performed in AMBER18,[53] implemented with the GLYCAM06 force field.[54] Notably, extended (5 *μ*s) MD simulations *alone* performed on derivatives *αf* and *βf* grossly failed to reproduce experimental data (Supplementary

https://dx.doi.org/10.1021/acscentsci.9b01277
ACS Cent. Sci. 2020, 6, 293−303

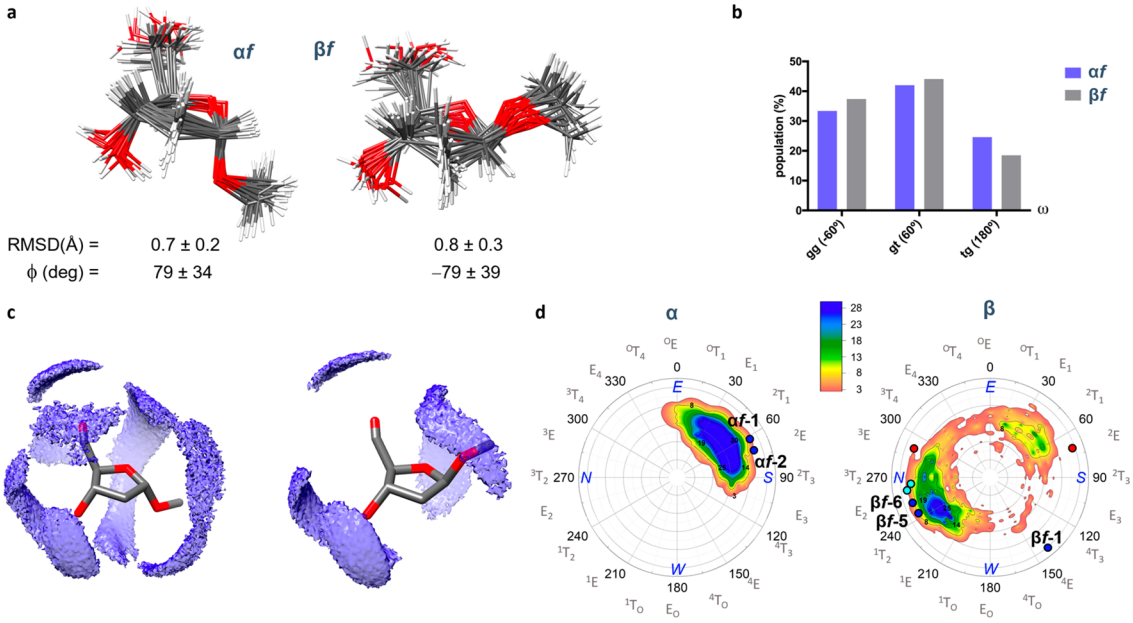


Figure 4. Conformational analysis of *αf* and *βf* in aqueous solution. (a) Structural ensembles derived from 0.2 *μs* MD-tar simulations, together with the root-mean-square deviation (RMSD) values for the heavy atoms in both anomers and the average value of ϕ torsional angle through the entire MD-tar trajectory. (b) Distribution of torsional angle ω for *αf* and *βf* derived from MD-tar simulations in aqueous solution. (c) Water oxygen density for the first hydration shell derived from experiment-guided MD simulations for *αf* and *βf*. The average structure of the carbohydrate is represented. (d) Cremer−Pople diagrams for *αf* and *βf* derived from experiment-guided MD simulations. N, S, E, and W stand for "North", "South", "East", and "West" forms. The dark blue circles represent the conformations found in the gas phase. The red circles denote the conformation found in DNA. Finally, the light blue circles show the conformations found in the gas phase for methyl *β*-D-ribofuranoside.[41] The contour coloring indicates the population (in arbitrary units) of the different conformers obtained from experiment-guided MD simulations.

Table 15), predicting a rather flexible five-membered ring with almost free pseudorotation (Supplementary Figures 5 and 6). However, MD simulations (0.2 *μs*) using experimental $^{3}J_{H,H}$ constants as time-averaged restraints (MD-tar simulations,[55] exploiting additional restraining potential terms added to the force field) successfully predicted ensembles of low-energy conformers that quantitatively reproduced NMR data (see Figure 4 and also Supplementary Tables 14 and 15 for $^{3}J_{H,H}$ constants).

These simulations revealed that in the furanosides *αf* and *βf* the C-5′ hydroxymethyl group is notably flexible in water (Figure 4b) and preferentially adopts *gg* and *gt* staggered rotamers for both. Of note, no significant intramolecular hydrogen bonds were detected in solution, with a population <10% in all cases. Analysis of the first hydration shell predicts that *αf* can accommodate more water molecules than *βf* (19 versus 15, Figure 4c). Together these solvent interactions drive scaffold-specific conformer selection from corresponding "naked" cores. Thus, while unnatural scaffold *αf* populates $^{2}T_{1}$ and ^{2}E conformations (Figure 4d) in both gas and solution phase ("unselected"), the natural scaffold *βf* is driven from ^{4}E in the gas phase *βf*-1 to E_{2} and $^{3}T_{2}$ by the loss of intramolecular bond O5′−H5′···O1′ enabled by the observed flexibility of the C-5′ hydroxymethyl group, in agreement with conformers derived by fitting the experimental homo- and heteronuclear ^{3}J coupling constants to a two-state model[56,57] (^{2}E and $^{3}T_{2}$ for *αf* and *βf*, respectively).

For the pyranosides *αp* and *βp*, characteristic $^{3}J_{H2a,H3}$ values confirmed strong propensity to remain in chair conformations[58] observed in the gas phase, yet with inverted conformation ($^{4}C_{1} \rightarrow {}^{1}C_{4}$ and $^{1}C_{4} \rightarrow {}^{4}C_{1}$, respectively). This result is also in line with previous studies.[59] In contrast to the furanosides, unrestrained MD simulations gave good agreement with experiment (Supplementary Table 16); extended analyses of total trajectory time (Supplementary Figures 7 and 8) suggested that *αp* is driven in solution to adopt only a 82(5)% partial $^{1}C_{4}$-to-$^{4}C_{1}$ equilibrium position. As expected, no significant intramolecular hydrogen bonds (population of <10%) were detected in solution through the MD simulations for either of the anomers. On the other hand, full solvation analysis performed by MD revealed that the conformational inversions are driven by an ability to properly accommodate the water in its first hydration shell, which more than offsets *endo*-anomeric preferences. Moreover, in the case of the *α*-anomer, the $^{1}C_{4}$ conformer features a larger dipole moment relative to the $^{4}C_{1}$ chair, which could also explain why this anomer preferentially adopts the $^{1}C_{4}$ chair in aqueous solution.

DISCUSSION

There are considered to be three broad conformational determinants in DNA residues: C4′−C5′ rotation (i.e., *g*/*t*); base position/"flip" (i.e., *syn*/*anti*); and ring pucker (i.e., "North"/"South") pseudorotation (Figure 4d). Although to some extent linked, sometimes even mutually "geared"[60] and accessible via typically relatively low barriers, it is the latter from these that appears to dominate function. In general terms, therefore many biomolecular interactions with such "DNA sugar scaffolds" are found to favor function with either North or South sugar ring conformers.

A North-vs-South (N-vs-S) delineation is an observation made well beyond the more obvious and static structures, such as in for example A- or B-form duplex DNA. Thus, through the use of structural biology, spectroscopy, and elegant probe molecules (e.g., conformationally restricted variants[26,61,62]), a clear role for the conformation of deoxynucleosides in

https://dx.doi.org/10.1021/acscentsci.9b01277
ACS Cent. Sci. 2020, 6, 293−303

determining their biological function has emerged in recent decades.[63] For example, clear "N-vs-S" preferences have been seen in the function of base deaminases,[64] base C-methyltransferases,[60] kinases,[65] and even receptors[66] and transporters.[67] This can have profound effects; DNA-polymerizing enzymes, such as reverse transcriptase[68] or polymerases,[69] can have striking apparent selectivity; AZT, for example, is bound in essentially an exclusively North form by HIV reverse transcriptase, and this conformational mimicry appears critical to its potency.[68]

As a result, the delineation in this work of the inherent preferences of the available ribosides *αf*, *αp*, *βp*, *βf* and their behavior upon hydration allows consideration of the fundamental biases of such scaffolds. Notably, it is the natural scaffold *βf* that displays a unique "tipping point"/"knife-edge" conformational behavior that allows the greatest malleability by environment. Here, we observe population of kinetically trapped Southern (*βf*-1) and Northern conformers (*βf*-5, *βf*-6) as well as the "tipping" by water of *βf*-1 into "Northern" conformers. Associated assumptions have often been made around the central role of "sugar pucker". For example, the itinerary of the equilibrium between North and South has always been assumed to proceed "via the East".[62] Yet, here we see that *βf* sits in a semi-Western (indeed, SW tending to W) ^{4}E *βf*-1 pucker as its baseline conformer. Any engagement by "naked" *βf*-1 with the environment (e.g., added solvent or biomolecule) therefore starts from this point.

The partial population of clear Northern conformers *βf*-5 and *βf*-6 even when "naked" also suggests that *βf* is conformationally "primed". The unnatural deoxyriboside scaffold variants *αp*, *βp*, and even *αf* do not show this behavior. The switched loss of strength of the O5′−H5′··· aglycone hydrogen bond seen here (either through hydration, protein binding, or covalent capping, e.g., as O5′-P phosphoester) thus precludes the occurrence of this pucker leading to the adoption, in turn, of Northern (i.e., E_2 or $^1T_2$ structures upon hydration) highly similar to those of *βf*-5 and *βf*-6. In this way, *βf* is therefore uniquely suitable for a switchable (e.g., South-to-North) transition, unlike all of the other scaffolds studied here—it is "ready to switch". Notably, the C2′ hydroxylated scaffold of *βf* (the RNA scaffold equivalent of *βf*) is already tipped to North ($^3T_2$), further highlighting the uniqueness of the *βf* 2-deoxyriboside scaffold.

Notably, in *βf* the OH3′ hydroxyl group is not engaged in a hydrogen bond in the lowest lying energy conformer (*βf*-1, as well as in fact *βf*-6) found in the gas phase; *βf*-5 has interaction with OH5′ but contributes only 4% to the global population. Together, these combined factors ("tipping point" conformation and exposed OH3′) allow us to speculate further on the functional benefit of the *βf* scaffold over others. N ↔ S conformational transition allows strong modulation of the O5′-to-O3′ distance; a switchable system with OH3′ exposed at modulated distances for the reaction would allow a mechanism for ready positional alteration and hence the semiexpulsion of ligand that is thought to be critical for reducing product inhibition (and hence allowing more efficient turnover) in processive enzymes such as nucleotidyl polymerases and reverse transcriptases.[70] In other words, the ready positional manipulation of a "free OH3′" uniquely in the *βf* scaffold is potentially beneficial in many biocatalytic (and hence functional) scenarios. More generally, it might even explain the observed direction of such polymerizations of DNA in nature in growing from 5′-to-3′ (via such a modulated, free OH3′). This may not be restricted to polymerases given the wide N-vs-S preferences noted above. In several nucleoside/tide kinases, which operate at the O5′-site of ribosides, a critical and evolutionarily conserved role has been identified for OH3′: hydrogen bonding to homologous Tyr-Glu diad motifs is required in several to ensure catalysis.[71]

It is also tempting to speculate that, given the context of DNA polymerases as archetypal models for elegant but sometimes nonaligned "induced fit" hypotheses (for an excellent comparative discussion see ref 72), this may be part of broader conformational, "prechemistry" mechanisms—potential triggers for induced fit.

Regardless, it is clear that the unique flexibility that we observe for "naked" *βf* for the first time is striking and likely an important determinant in its utilization/selection by nature as a scaffold from many stereo- and constitutional sugar isomers. Although such conformational "fitness" has been the subject of insightful prior speculation,[73] our work provides the first direct experiment of such a flexible trigger based on inherent conformational tendencies in the *βf* scaffold.

Finally, given the low barriers for DNA conformational interchange and the discrepancies that we observed in the work presented here exploiting existing QM and MM methods, the structures we present here (nonreliant on QM) should prove valuable in benchmarking future quantum mechanical models. In this way, our combined approach opens up the exploration of DNA's mechanistic tendencies to even larger systems in the future.

METHODS

Synthesis. The 2-deoxyribosides *αf*, *βf*, *αp*, and *βp* (Figure 1) were synthesized by modification of a previously published literature procedure.[74] Briefly, for access to *αp* and *βp*, 2-deoxy-D-ribose was dissolved in 1% HCl in methanol and stirred for 18 h. After the reaction workup, pure *αp* was obtained through purification by column chromatography. The remaining 2-deoxyribosides *βp*, *αf*, and *βf* coeluted as an inseparable mixture, but recrystallization of this mixture in diethyl ether afforded pure *βp*. For access to *αf* and *βf*, 2-deoxy-D-ribose was dissolved in 0.1% HCl in methanol and stirred for 15 min. After the reaction workup, the mixture of *αf* and *βf* was acetylated using acetic anhydride in pyridine and pure *αf*-2OAc and *βf*-2OAc was obtained through purification by column chromatography. Deacetylation using K_2CO_3 in methanol yielded pure *αf* and *βf*. Full experimental details can be found in the Supporting Information.

Pulse Fourier Transform Microwave Spectrometers. To obtain the rotational spectra of the furanose species, we used a chirped-pulse Fourier transform microwave spectrometer (CP-FTMW) built at the Spectroscopy Group at the University of the Basque Country (UPV/EHU)[44] and following the design of Pate and co-workers (Supplementary Figure 1).[75] A short broadband pulse (1 μs, 12 GHz, frequency range 6−18 GHz) is generated by an Arbitrary Waveform Generator (AWG) and amplified by a Traveling Wave Tube Amplifier (TWTA). The pulse is broadcast into a high vacuum chamber (down to 10^{-6} mbar) through a Q-par horn antenna where it interacts with the molecular supersonic jet expansion. The time-domain signal of the molecular emission is collected by another horn antenna and delivered to a digital 20 GHz oscilloscope and Fourier-transformed to obtain the rotational spectrum in the frequency domain. The spectral resolution of this apparatus is ∼10 kHz.

https://dx.doi.org/10.1021/acscentsci.9b01277
ACS Cent. Sci. 2020, 6, 293−303

Rotational spectra of pyranose conformers were recorded in a 4−18 GHz Fourier transform microwave spectrometer (FTMW) based on the Balle-Flygare design,[76] constructed at the UPV/EHU and described elsewhere (Supplementary Figure 2).[43] Appropriate excitation pulses create optimum $\pi/2$ polarization conditions on the sample, which expands in a supersonic jet coaxially within a Fabry-Pérot microwave resonator. The resulting transient spontaneous emission from the expanding molecular ensemble is amplified and down-converted to the radio frequency region, where it is digitized, and Fourier transformed to yield the frequency-domain spectrum. All transitions recorded in this spectrometer appear split into two Doppler components, because of the coaxial arrangement of the molecular jet and the resonator. The higher spectral resolution of this apparatus (better than 5 kHz) allowed us to resolve the hyperfine line splittings due to internal rotation of methyl tops, where present.

Samples in the CP-FTMW spectrometer were vaporized by a conventional heating method by wrapping a heating wire to a customized nozzle at 145 °C. The vacuum chamber hosts three pulsed solenoid valves running simultaneously in order to increase the signal-to-noise ratio. Gaseous samples were mixed with He carrier gas at 6 bar, and the resulting flow was injected into the vacuum chamber through a small orifice (1 mm diameter), producing the gas expansion and subsequent cooling (rotational temperature near 5 K). The duration of the molecular pulse was 0.5−1.2 ms. Two nozzles were enough for αf, whereas three nozzles were used for βf, at 30 chirps per molecular pulse. To obtain a good signal-to-noise ratio, a total of 20 and 40 Mcycles were collected for the αf and βf spectra, respectively. In the FTMW spectrometer equipped with UV ultrafast laser vaporization,[39] it consists of a mixture of the chemical compound and a commercial binder, mechanically pressed to get cylinder-type solid rods that are inserted in the apparatus and rotated by a stepper motor. A picosecond pulse from the third harmonic (355 nm) of a Nd:YAG laser hits the rod in the presence of Ne carrier gas (~6 bar), delivering the molecules intact into the gas phase. This technique prevents sample decomposition, that often happens with biomolecules upon thermal heating.

Computational Details. *Conformational Search.* The conformational search used a molecular mechanics approach with MMFFs, OPLS, and AMBER force fields in a 20 kJ/mol energy window and was followed by geometry optimizations with quantum chemistry calculations [*ab initio* MP2 and DFT (B3LYP-D3)] with 6-311++G(d,p) basis functions. Gaussian 16 software[77] was used in all cases.

Unrestrained MD Simulations. Simulations were performed with the AMBER18 package,[53] implemented with a GLYCAM06j[54] force field. Each molecule was immersed in a water box with a 10 Å buffer of TIP3P water molecules. A two-stage geometry optimization approach was performed. The first stage minimizes only the positions of solvent molecules, and the second stage is an unrestrained minimization of all the atoms in the simulation cell. The systems were then gently heated by incrementing the temperature from 0 to 300 K under a constant pressure of 1 atm and periodic boundary conditions. Harmonic restraints of 30 kcal mol^{-1} were applied to the solute, and the Andersen temperature coupling scheme was used to control and equalize the temperature. The time step was kept at 1 fs during the heating stages, allowing potential inhomogeneities to self-adjust. Long-range electrostatic effects are modeled using the particle-mesh-Ewald method.[78] An 8 Å cutoff was applied to Lennard−Jones interactions. Each system was equilibrated for 2 ns with a 2 fs time step at a constant volume and temperature of 300 K. Production trajectories were then run for an additional 5 μs under the same simulation conditions.

Experiment-Guided MD Simulations. The setup of the MD simulations with time-averaged restraints was identical, in terms of force fields, water model, and algorithms to that described above for the unrestrained MD simulations. The experimental $^3J_{H,H}$ coupling constants were imposed as a time-averaged restraint, applying a linear averaging. The equilibrium ^3J range was set to $^3J_{exp} - 0.2$ Hz $\leq {}^3J_{exp} \leq {}^3J_{exp} + 0.2$ Å. Trajectories were run at 300 K, with a decay constant of 20 ns and a time step of 1 fs. The force constants rk_2 and rk_3 used in each case were 0.2 kcal mol^{-1} Å^{-2}. The overall simulation length was 0.2 μs. The theoretical ^3J coupling constants were deduced from the dihedral values through the corresponding Altona equation and Sweet J software.[58,79]

Analysis of the First Hydration Shell from the MD Simulations. The water density properties were derived from the production trajectories using a cubic grid consisting of 150 × 150 × 150 bins with 0.5-Å spacing, through the "*grid*" command available in the *cpptraj* module of AMBER18. The surfaces represented in Figures 4 and Supplementary Figure 7 correspond to oxygen water density for the first hydration shell.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscentsci.9b01277.

Supplementary Figure 1: Scheme of the broadband chirped-pulse Fourier Transform microwave spectrometer. Supplementary Figure 2: Scheme of the molecular beam pulsed-jet Fourier Transform microwave spectrometer. Supplementary Figures 3−4: Interconversion paths between the lowest lying conformers. Supplementary Figure 5: Structural ensembles derived from unrestrained MD simulations in aqueous solution for αf and βf. Supplementary Figure 6: Distribution of torsional angle ω for αf and βf derived from unrestrained MD simulations in aqueous solution. Supplementary Figure 7: Conformational analysis of αp and βp in aqueous solution. Supplementary Figure 8: Evolution of the main conformers of αf (b) and βf (b) along experiment-guided MD simulations in aqueous solution. Supplementary Figures 9−20: Copies of ^{1}H and ^{13}C NMR spectra. Supplementary Tables 1−7: Experimental spectroscopic parameters and comparison with theoretical predictions of DNA Scaffolds. Supplementary Tables 8−10: Comparison between experimental (r_0) and calculated (r_e) structural parameters of DNA Scaffolds. Supplementary Tables 11−13: Experimental (r_0 and r_S) and calculated (r_e) Cartesian coordinates of DNA Scaffolds. Supplementary Table 14: Karplus equations used in the MD simulations. Supplementary Tables 15−16: Comparison between the experimental and theoretical coupling constants derived from MD simulations (PDF)

https://dx.doi.org/10.1021/acscentsci.9b01277
ACS Cent. Sci. 2020, 6, 293−303

AUTHOR INFORMATION

Corresponding Authors

Benjamin G. Davis − *Department of Chemistry, Chemistry Research Laboratory, University of Oxford, Oxford OX1 3TA, United Kingdom; The Rosalind Franklin Institute, Oxfordshire OX11 0FA, United Kingdom;* orcid.org/0000-0002-5056-407X; Email: ben.davis@chem.ox.ac.uk, ben.davis@rfi.ac.uk

Francisco Corzana − *Departamento de Química, Centro de Investigación en Síntesis Química, Universidad de La Rioja, 26006 Logroño, Spain;* orcid.org/0000-0001-5597-8127; Email: francisco.corzana@unirioja.es

Emilio J. Cocinero − *Departamento de Química Física, Facultad de Ciencia y Tecnología, Universidad del País Vasco (UPV/EHU), 48080 Bilbao, Spain; Instituto Biofisika (CSIC, UPV/EHU), 48080 Bilbao, Spain;* orcid.org/0000-0001-7632-3728; Email: emiliojose.cocinero@ehu.es

Authors

Camilla Calabrese − *Departamento de Química Física, Facultad de Ciencia y Tecnología, Universidad del País Vasco (UPV/EHU), 48080 Bilbao, Spain; Instituto Biofisika (CSIC, UPV/EHU), 48080 Bilbao, Spain;* orcid.org/0000-0003-4299-2098

Iciar Uriarte − *Departamento de Química Física, Facultad de Ciencia y Tecnología, Universidad del País Vasco (UPV/EHU), 48080 Bilbao, Spain; Instituto Biofisika (CSIC, UPV/EHU), 48080 Bilbao, Spain*

Aran Insausti − *Departamento de Química Física, Facultad de Ciencia y Tecnología, Universidad del País Vasco (UPV/EHU), 48080 Bilbao, Spain; Instituto Biofisika (CSIC, UPV/EHU), 48080 Bilbao, Spain*

Montserrat Vallejo-López − *Departamento de Química Física, Facultad de Ciencia y Tecnología, Universidad del País Vasco (UPV/EHU), 48080 Bilbao, Spain*

Francisco J. Basterretxea − *Departamento de Química Física, Facultad de Ciencia y Tecnología, Universidad del País Vasco (UPV/EHU), 48080 Bilbao, Spain*

Stephen A. Cochrane − *Department of Chemistry, Chemistry Research Laboratory, University of Oxford, Oxford OX1 3TA, United Kingdom;* orcid.org/0000-0002-6239-6915

Complete contact information is available at: https://pubs.acs.org/10.1021/acscentsci.9b01277

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank MINECO (Projects CTQ2017-89150 R, CTQ2015 68148-C2-2P), Ministerio de Ciencia, Innovación y Universidades (RTI2018-099592-B-C21), Basque Government (IT1162-19, PIBA 2018/11), the UPV/EHU (PPG17/10 and GIU18/207), CSIC (PIC2018, LINKA20249) Fundación BBVA and the Wellcome Trust (110270/A/15/Z) for financial support. C.C. and A.I. thank MINECO and Basque Government for Juan de la Cierva and predoctoral contracts, respectively. Laser and NMR resources of the UPV/EHU together with computational resources of BERONIA (Universidad de La Rioja) and UPV/EHU (SGIker and I2Basque) were used in this work.

REFERENCES

(1) Pabo, C. O.; Sauer, R. T. Protein-DNA recognition. *Annu. Rev. Biochem.* **1984**, *53*, 293−321.
(2) Pabo, C. O.; Sauer, R. T. Transcription Factors: Structural Families and Principles of DNA Recognition. *Annu. Rev. Biochem.* **1992**, *61*, 1053−1095.
(3) Rohs, R.; et al. Origins of Specificity in Protein-DNA Recognition. *Annu. Rev. Biochem.* **2010**, *79*, 233−269.
(4) Higgs, P. G.; Lehman, N. The RNA World: Molecular cooperation at the origins of life. *Nat. Rev. Genet.* **2015**, *16*, 7−17.
(5) Duzdevich, D.; Redding, S.; Greene, E. C. DNA Dynamics and Single-Molecule Biology. *Chem. Rev.* **2014**, *114*, 3072−3086.
(6) Barrangou, R.; Horvath, P. A decade of discovery: CRISPR functions and applications. *Nat. Microbiol.* **2017**, *2*, 17092.
(7) Rinaldi, C.; Wood, M. J. A. Antisense oligonucleotides: The next frontier for treatment of neurological disorders. *Nat. Rev. Neurol.* **2018**, *14*, 9−22.
(8) Wittrup, A.; Lieberman, J. Knocking down disease: A progress report on siRNA therapeutics. *Nat. Rev. Genet.* **2015**, *16*, 543−552.
(9) Herdewijn, P.; Marlière, P. Redesigning the leaving group in nucleic acid polymerization. *FEBS Lett.* **2012**, *586*, 2049−2056.
(10) Campbell, M. A.; Wengel, J. Locked vs. unlocked nucleic acids (LNA vs. UNA): Contrasting structures work towards common therapeutic goals. *Chem. Soc. Rev.* **2011**, *40*, 5680−5689.
(11) Nakamura, T.; Zhao, Y.; Yamagata, Y.; Hua, Y. J.; Yang, W. Watching DNA polymerase η make a phosphodiester bond. *Nature* **2012**, *487*, 196−201.
(12) Arangundy-Franklin, S.; et al. A synthetic genetic polymer with an uncharged backbone chemistry based on alkyl phosphonate nucleic acids. *Nat. Chem.* **2019**, *11*, 533−542.
(13) Eschenmoser, A. Chemical Etiology of Nucleic Acid Structure. *Science* **1999**, *284*, 2118−2124.
(14) Schöning, K.-U.; et al. Chemical Etiology of Nucleic Acid Structure: The α-Threofuranosyl-(3'→2') Oligonucleotide System. *Science* **2000**, *290*, 1347−1351.
(15) Gaweda, K.; Plazinski, W. The systematic influence of solvent on the conformational features of furanosides. *Org. Biomol. Chem.* **2019**, *17*, 2479−2485.
(16) Cheung, T., Ramesh, V. Biomolecular NMR Spectroscopy and Structure Determination of DNA. *Biomolecular and Bioanalytical Techniques* **2019**, 421..
(17) Watson, J. D.; Crick, F. H. C. Molecular Structures of Nucleic Acids. *Nature* **1953**, *171*, 737−738.
(18) Wang, A. H.; Fujii, S.; van Boom, J. H.; Rich, A. Molecular structure of the octamer d(G-G-C-C-G-G-C-C): modified A-DNA. *Proc. Natl. Acad. Sci. U. S. A.* **1982**, *79*, 3968−3972.
(19) Wang, A. H.-J.; Hakoshima, T.; van der Marel, G.; van Boom, J. H.; Rich, A. AT base pairs are less stable than GC base pairs in Z-DNA: The crystal structure of d(m5CGTAm5CG). *Cell* **1984**, *37*, 321−331.
(20) Larsen, T. A.; Kopka, M. L.; Dickerson, R. E. Crystal structure analysis of the B-DNA dodecamer CGTGAATTCACG. *Biochemistry* **1991**, *30*, 4443−4449.
(21) Shapiro, Y. E. NMR spectroscopy on domain dynamics in biomacromolecules. *Prog. Biophys. Mol. Biol.* **2013**, *112*, 58−117.
(22) Pérez, A.; Luque, F. J.; Orozco, M. Frontiers in Molecular Dynamics Simulations of DNA. *Acc. Chem. Res.* **2012**, *45*, 196−205.
(23) Imeddourene, A. B.; Xu, X.; Zargarian, L.; Oguey, C.; Foloppe, N.; Mauffret, O.; Hartmann, B. The intrinsic mechanics of B-DNA in solution characterized by NMR. *Nucleic Acids Res.* **2016**, *44*, 3432−3447.
(24) Schwieters, C. D.; Clore, G. M. A physical picture of atomic motions within the Dickerson DNA dodecamer in solution derived from joint ensemble refinement against NMR and large-angle X-ray scattering data. *Biochemistry* **2007**, *46*, 1152−1166.
(25) Tjandra, N.; Tate, S. I.; Ono, A.; Kainosho, M.; Bax, A. The NMR structure of a DNA dodecamer in an aqueous dilute liquid crystalline phase. *J. Am. Chem. Soc.* **2000**, *122*, 6190−6200.

(26) Taha, H. A.; Richards, M. R.; Lowary, T. L. Conformational analysis of furanoside-containing mono- and oligosaccharides. *Chem. Rev.* **2013**, *113*, 1851−1876.
(27) Olson, W. K.; Sussman, J. L. How Flexible Is the Furanose Ring? 1. A Comparison of Experimental and Theoretical Studies. *J. Am. Chem. Soc.* **1982**, *104*, 270−278.
(28) Talbot, F. O.; Simons, J. P. Sugars in the gas phase: the spectroscopy and structure of jetcooled phenyl beta-D-glucopyranoside. *Phys. Chem. Chem. Phys.* **2002**, *4*, 3562−3565.
(29) Jockusch, R. A.; Talbot, F. O.; Simons, J. P. Sugars in the gas phase. *Phys. Chem. Chem. Phys.* **2003**, *5*, 1502−1507.
(30) Barry, C. S.; et al. 'Naked' and Hydrated Conformers of the Conserved Core Pentasaccharide of N-linked Glycoproteins and Its Building Blocks. *J. Am. Chem. Soc.* **2013**, *135*, 16895−16903.
(31) Stanca-Kaposta, E. C.; et al. Solvent Interactions and Conformational Choice in a Core N-Glycan Segment: Gas Phase Conformation of the Central, Branching Trimannose Unit and its Singly Hydrated Complex. *J. Am. Chem. Soc.* **2008**, *130*, 10691−10696.
(32) Su, Z.; Cocinero, E. J.; Stanca-Kaposta, E. C.; Davis, B. G.; Simons, J. P. Carbohydrate-aromatic interactions: A computational and IR spectroscopic investigation of the complex, methyl α-l-fucopyranoside · toluene, isolated in the gas phase. *Chem. Phys. Lett.* **2009**, *471*, 17−21.
(33) Screen, J.; et al. IR-Spectral Signatures of Aromatic−Sugar Complexes: Probing Carbohydrate−Protein Interactions. *Angew. Chem., Int. Ed.* **2007**, *46*, 3644−3648.
(34) Cocinero, E. J.; Carcabal, P.; Vaden, T. D.; Simons, J. P.; Davis, B. G. Sensing the anomeric effect in a solvent-free environment. *Nature* **2011**, *469*, 76−80.
(35) Pérez, C.; et al. Structures of cage, prism, and book isomers of water hexamer from broadband rotational spectroscopy. *Science* **2012**, *336*, 897−901.
(36) Patterson, D.; Schnell, M.; Doyle, J. M. Enantiomer-specific detection of chiral molecules via microwave spectroscopy. *Nature* **2013**, *497*, 475−477.
(37) Motiyenko, R. A.; Alekseev, E. A.; Dyubko, S. F.; Lovas, F. J. Microwave spectrum and structure of furfural. *J. Mol. Spectrosc.* **2006**, *240*, 93−101.
(38) Fritzsche, J. Rotationsspektroskopische Untersuchungen von Kohlehydraten im LASER-Plasma − Ablation von D-(−)-Ribose. Diplomarbeit, Gottfried-Wilhelm-Leibniz Univ. Hann., 2006.
(39) Cocinero, E. J.; et al. Ribose found in the gas phase. *Angew. Chem., Int. Ed.* **2012**, *51*, 3119−3124.
(40) Peña, I.; et al. Six pyranoside forms of free 2-deoxy-D-ribose. *Angew. Chem., Int. Ed.* **2013**, *52*, 11840−11845.
(41) Écija, P.; et al. Furanosic forms of sugars: Conformational equilibrium of methyl β-D-ribofuranoside. *Chem. Commun.* **2016**, *52*, 6241−6244.
(42) Calabrese, C.; et al. Conformational Behavior of D-Lyxose in Gas and Solution Phases by Rotational and NMR Spectroscopies. *J. Phys. Chem. Lett.* **2019**, *10*, 3339−3345.
(43) Cocinero, E. J.; et al. Conformational equilibria in vanillin and ethylvanillin. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12486−12493.
(44) Uriarte, I.; Perez, C.; Caballero-Mancebo, E.; Basterretxea, F. J.; Lesarri, A.; Fernandez, J. A.; Cocinero, E. J. Structural Studies of Nicotinoids: Cotinine versus Nicotine. *Chem. - Eur. J.* **2017**, *23*, 7238−7244.
(45) Gordy, W.; Cook, R. L. *Microwave Molecular Spectra*; Interscience: New York, 1984.
(46) Watson, J. K. D. *Vibrational Spectra and Structure*; Elsevier, 1977.
(47) Rudolph, H. D. Contribution to the systematics of r0-derived molecular structure determinations from rotational parameters. *Struct. Chem.* **1991**, *2*, 581−588.
(48) Kraitchman, J. Determination of Molecular Structure from Microwave Spectroscopic Data. *Am. J. Phys.* **1953**, *21*, 17−24.
(49) Costain, C. C. Determination of molecular structures from ground state rotational constants. *J. Chem. Phys.* **1958**, *29*, 864−874.
(50) Fraser, G. T.; Suenram, R. D.; Lugez, C. L. Rotational spectra of seven conformational isomers of 1-hexene. *J. Phys. Chem. A* **2000**, *104*, 1141−1146.
(51) Cocinero, E. J.; et al. Conformational choice and selectivity in singly and multiply hydrated monosaccharides in the gas phase. *Chem. - Eur. J.* **2008**, *14*, 8947−8955.
(52) Wiberg, K. B.; Bailey, W. F.; Lambert, K. M.; Stempel, Z. D. The Anomeric Effect: It's Complicated. *J. Org. Chem.* **2018**, *83*, 5242−5255.
(53) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; et al. *Amber 2018*; University of California: San Francisco, 2018.
(54) Kirschner, K. N.; et al. GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.* **2008**, *29*, 622−655.
(55) Hendrickx, P. M. S.; Corzana, F.; Depraetere, S.; Tourwe, D. A.; Augustyns, K.; Martins, J. C. The use of time-averaged 3JHH Restrained molecular dynamics (tar-MD) simulations for the conformational analysis of five-membered ring systems: Methodology and applications. *J. Comput. Chem.* **2009**, *31*, 561−572.
(56) van Lieshout, H. C.; Raap, J.; van Boom, J. H.; Haasnoot, C. A. G. Conformations of Methyl 2'-Deoxy-α-D-ribofuranoside and Methyl 2'-Deoxy-β-D-ribofuranoside. A Proton Magnetic Resonance Spectroscopy and Molecular Mechanics Study. *J. Am. Chem. Soc.* **1988**, *110*, 2736−2743.
(57) Church, T. J.; Carmichael, I.; Serianni, A. S. 13C-1H and 13C-13C spin-coupling constants in methyl β-D-ribofuranoside and methyl 2-deoxy-β-D-erythro-pentofuranoside: Correlations with molecular structure and conformation. *J. Am. Chem. Soc.* **1997**, *119*, 8946−8964.
(58) Donders, L. A.; De Leeuw, F. A. A. M.; Altona, C. Relationship between proton—proton NMR coupling constants and substituent electronegativities. IV—An extended karplus equation accounting for interactions between substituents and its application to coupling constant data calculated by the Extended Hückel. *Magn. Reson. Chem.* **1989**, *27*, 556−563.
(59) Franks, F.; Lillford, P. J.; Robinson, G. Isomeric equilibria of monosaccharides in solution. Influence of solvent and temperature. *J. Chem. Soc., Faraday Trans. 1* **1989**, *85*, 2417−2426.
(60) Wang, P.; et al. Use of Oligodeoxyribonucleotides with Conformationally Constrained Abasic Sugar Targets To Probe the Mechanism of Base Flipping by HhaI DNA (Cytosine C5)-methyltransferase. *J. Am. Chem. Soc.* **2000**, *122*, 12422−12434.
(61) Marquez, V. E.; et al. Nucleosides with a twist. Can fixed forms of sugar ring pucker influence biological activity in nucleosides and oligonucleotides? *J. Med. Chem.* **1996**, *39*, 3739−3747.
(62) Mathé, C.; Périgaud, C. Recent Approaches in the Synthesis of Conformationally Restricted Nucleoside Analogues. *Eur. J. Org. Chem.* **2008**, *2008*, 1489−1505.
(63) Birnbaum, G. I.; Shugar, D. *Topics in Nucleic Acid Structure*; Macmillan Press, 1997.
(64) Marquez, V. E.; et al. Contrasting behavior of conformationally locked carbocyclic nucleosides of adenosine and cytidine as substrates for deaminases. *Nucleosides, Nucleotides Nucleic Acids* **2009**, *28*, 614−632.
(65) Sjuvarsson, E.; Marquez, V. E.; Eriksson, S. Selective Phosphorylation of South and North-Cytidine and Adenosine Methanocarba-Nucleosides by Human Nucleoside and Nucleotide Kinases Correlates with Their Growth Inhibitory Effects on Cultured Cells. *Nucleosides, Nucleotides Nucleic Acids* **2015**, *34*, 544−564.
(66) Jacobson, K. A.; et al. Methanocarba Analogues of Purine Nucleosides as Potent and Selective Adenosine Receptor Agonists. *J. Med. Chem.* **2000**, *43*, 2196−2203.
(67) Damaraju, V. L.; et al. Influence of Sugar Ring Conformation on the Transportability of Nucleosides by Human Nucleoside Transporters. *ChemBioChem* **2011**, *12*, 2774−2778.
(68) Marquez, V. E.; et al. HIV-1 reverse transcriptase can discriminate between two conformationally locked carbocyclic AZT triphosphate analogues. *J. Am. Chem. Soc.* **1998**, *120*, 2780−2789.

https://dx.doi.org/10.1021/acscentsci.9b01277
ACS Cent. Sci. 2020, 6, 293−303

(69) Eoff, R. L.; et al. Selective modulation of DNA polymerase activity by fixed-conformation nucleoside analogues. *Angew. Chem., Int. Ed.* **2010**, *49*, 7481−7485.
(70) Majumdar, C.; Abbotts, J.; Broder, S.; Wilson, S. H. Studies on the mechanism of human immunodeficiency virus reverse transcriptase. Steady-state kinetics, processivity, and polynucleotide inhibition. *J. Biol. Chem.* **1988**, *263*, 15657−15665.
(71) Johansson, K.; et al. Structural basis for substrate specificities of cellular deoxyribonucleoside kinases. *Nat. Struct. Biol.* **2001**, *8*, 616−620.
(72) Mulholland, A. J.; Roitberg, A. E.; Tuñón, I. Enzyme dynamics and catalysis in the mechanism of DNA polymerase. *Theor. Chem. Acc.* **2012**, *131*, 1286.
(73) Eschenmoser, A. Chemistry of potentially prebiological natural products. *Origins Life Evol. Biospheres* **1994**, *24*, 389−423.
(74) Deriaz, R. E.; Overend, W. G.; Stacey, M.; Wiggins, L. F. Deoxy-sugars. Part VI. The constitution of β-methyl-2-deoxy-L-ribopyranoside and of $\alpha\beta$-methyl-2-deoxy-L-ribofuranoside. *J. Chem. Soc.* **1949**, *0*, 2836−2841.
(75) Brown, G. G.; et al. A broadband Fourier transform microwave spectrometer based on chirped pulse excitation. *Rev. Sci. Instrum.* **2008**, *79*, 053103.
(76) Balle, T. J.; Flygare, W. H. Fabry-Perot cavity pulsed Fourier transform microwave spectrometer with a pulsed nozzle particle source. *Rev. Sci. Instrum.* **1981**, *52*, 33−45.
(77) Frisch, M. J. et al. *Gaussian 16*; Gaussian, Inc.: Wallingford, CT, 2016.
(78) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(79) Balacco, G. A desktop calculator for the Karplus equation. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 885−887.

https://dx.doi.org/10.1021/acscentsci.9b01277
ACS Cent. Sci. 2020, 6, 293−303