Molecular Therapy
Nucleic Acids
Original Article

Molecular dynamics simulations of acyclic analogs of nucleic acids for antisense inhibition

Rodrigo Galindo-Murillo,[1] Jack S. Cohen,[2] and Barak Akabayov[2]

[1]Medicinal Chemistry Department, Skaggs Pharmacy Institute, University of Utah, Salt Lake City, UT, USA; [2]Department of Chemistry, Ben-Gurion University of the Negev, Beer-Sheva, Israel

For antisense applications, oligonucleotides must be chemically modified to be resistant to endogenous nucleases. Until now, antisense oligonucleotide (ASO) analogs have been synthesized and then tested for their ability to duplex with a target nucleic acid, usually RNA. In this work, using molecular dynamics calculations simulations, we systematically tested a series of chemically modified analogs in which the 2-deoxyribose was substituted for by one or two methylene groups on each side of the phosphate backbone, producing four compounds, of which three were previously unknown. We used a 9-mer sequence of which the solution structure has been determined by NMR spectroscopy and tested the ability to form stable duplexes of these acyclic analogs to both DNA and RNA. In only one case out of eight, we unexpectedly found the formation of a stable duplex with complementary RNA. We also applied limitations on end fraying because of the terminal AT base pairs, in order to eliminate this as a factor in the comparative results. We consider this a predictive method to potentially identify target ASO analogs for synthesis and testing for antisense drug development.

INTRODUCTION

The structure of DNA that is the basis of life and reproduction evolved through millennia. The Watson-Crick (WC) double-helical structure of DNA is both chemically stable to ensure the retention of the genetic information and enzymatically attainable. That means that the enzymes required for its synthesis, DNA polymerase, and the many other enzymes responsible for the biosynthesis of its components were able to evolve concurrently. However, from a purely structural point of view, the structure of DNA is unnecessarily complex and chemically redundant and may have had relevant pre-biotic precursors.[1]

To illustrate this point, consider the double-helical structure of DNA as an assemblage of stacked bases. As shown very early in the first X-ray diffraction study of DNA by Astbury,[2] the strongest feature in the diffraction pattern of DNA corresponds to a distance of ca. 3.3 Å, on the basis of which Astbury speculated that this was the distance between the bases, and he suggested the structure of DNA was a stack of bases, which he likened to a stack of coins.[2] In order to connect these closely stacked bases together, one requires only three or four covalent bonds. It is not generally realized that in DNA there are, in fact, 10 atoms, or nine bonds, separating the bases due to the presence of the deoxyribose phosphate backbone (Figure 1).

As a result of pursuing the use of oligodeoxynucleotides as antisense inhibitors of gene expression,[3] it became necessary to synthesize chemically modified analogs to protect them against prevalent nucleases. However, most of these analogs were minimally modified, such as the phosphorothioate analog,[4] to preserve the need for duplex formation with a target mRNA. These PS analogs are widely used, and many are in US Food and Drug Administration (FDA) trials; six have been approved for human use.[5,6] Nevertheless, the concept of synthesizing and testing biomimetic analogs, yet that differ radically from the natural deoxyribose phosphate backbone, was conceived.

There have been no systematic attempts as far as we are aware of synthesizing biomimetic analogs of DNA that do not have the natural deoxyribose-phosphate backbone. Certainly, such simplified analogs may have different topological features compared with DNA, but they can be expected to have important biological properties. The simplest type of analog of this kind one can envisage consists of bases with an inter-base linkage devoid of the deoxyribose sugar (note that because the sugar moiety is removed, these analogs will be the same for RNA).

Based on the above analysis, it should be possible to synthesize analogs of DNA in which the bases are linked by one or two simple groups, such as phosphate, peptide, or methylene bonds. Some of these compounds would represent new compositions of matter that have not been synthesized before. Further, preliminary analysis by computer molecular modeling with energy minimization indicates that although three bonds between bases are insufficient to preserve stacking as a result of base distortion, four bonds are sufficient (T.-J. Syi, J. Maizel, and J.S.C., unpublished data).

Received 30 May 2020; accepted 28 November 2020;
https://doi.org/10.1016/j.omtn.2020.11.023.

Correspondence: Barak Akabayov, Chemistry Department, Ben Gurion University, Beer Sheva, Israel.
E-mail: akabayov@bgu.ac.il

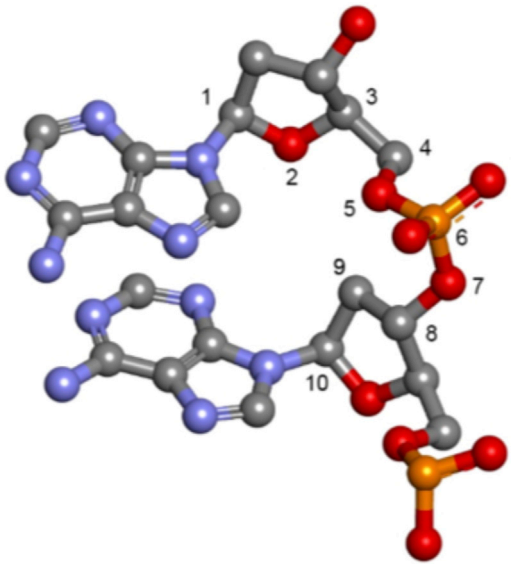


Figure 1. Showing the number of atoms/bonds between adjacent bases in DNA

We propose to synthesize several types of such analogs connected by one and two phosphates, peptide, carboxyl, methylene, and other groups, and to test them both for their duplex-forming capability and for their potential antisense activity. Ultimately, we would hope to develop these compounds as potential drugs of clinical importance.

Many analogs of DNA have previously been synthesized (Figure 2), most of which are based on minor substitutions of the canonical structure.[7,8] There have also been more extensive modifications, such as peptide nucleic acids (PNAs),[9,10] that have a peptide-like linker between the bases (Figure 3), as well as modified versions of these. Another analog that is reported to have biological properties is the morpholino analog (Figure 3).[11] There is also so-called acyclic or linear DNA[12,13] in which the deoxyribose sugar has been opened in one way or another (Figure 4). Some of these analogs particularly tend to be too flexile for duplex formation. There are also sugar substituted analogs of DNA (e.g., Declercq et al.[14]). These analogs can all be considered special cases of the general approach of replacing the deoxyribose phosphate backbone.

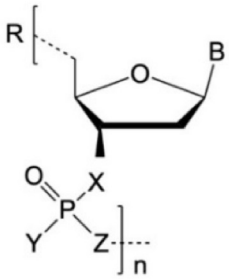
A striking example of these kinds of simplified intra-base analogs is that of glycol nucleic acid (GNA) that has been synthesized based on propylene glycol as a linker.[15,16] This has been shown to form a highly stable homo-duplex (Figure 5), although not a stable duplex with a natural DNA oligo. The synthesis and duplex-forming properties of this analog provide a proof of concept for the general approach to the synthesis of such analogs as detailed here. One major reason for developing nucleic acid analogs in the first place was to render them resistant to nuclease degradation *in vitro* and *in vivo*.[4] Although the simplified analogs presented here may not be susceptible to nucleases, those containing phosphate may be degraded by phosphatases, and this is a factor that will have to be investigated.

In a previous publication we used MD methodology to study the ability to form stable duplex structures of a phosphorothioate analog of DNA with a natural congener.[17] Given the wide range of nucleic acid analogs being tested, there needs to be some self-consistent means of comparing their effectiveness as antisense complements without actually synthesizing them. Computer-aided molecular design (CAMD) is a valuable methodology to give insight into the complexation properties of DNA analogs prior to attempting synthesis. To explore the structure and dynamics of acyclic DNA chains and the capabilities of forming stable duplex structures with both DNA and RNA receptors, we performed microsecond, multiple replicas MD simulations of the systems presented in Figure 6, which consists of linear saturated carbon chains of the following form: -CH(base)-PO_4-CH(base)-, -CH(base)-PO_4-CH_2-CH(base)-, -CH(base)-CH2-PO_4-CH(base)-, and -CH(base)-CH_2-PO_4-CH_2-CH(base)-.

RESULTS

Double-strand hybrid

To study the capability of these modified backbones to form a double strand with either DNA or RNA, we conducted molecular dynamics simulations of double-strand hybrids with both DNA and RNA. The modified backbones forming a duplex with DNA resulted in highly



Name of Analogue	R	X	Y	Z
DNA	OH	O	O	O
Imidazole	Imidazole	O	O	O
Quinone	Quinone	O	O	O
Phosphorothioate	OH	O	S	O
Methylphsphonate	OH	O	CH3	O
3'-Methylene-phosphonate	OH	CH2	O	O

Figure 2. DNA analogs based on minor substituents of the canonical structure

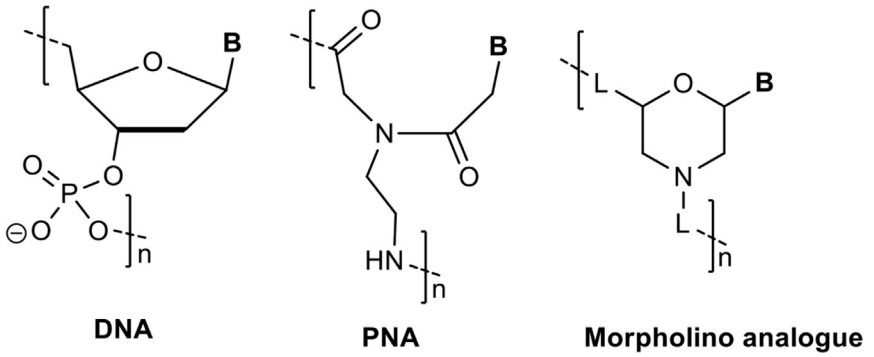


Figure 3. PNA and morpholino analog of DNA structures

distorted structures as evidenced in an initial root-mean-square analysis (Figure 7). The DNA reference structure's (PDB: 6ed9) inner base pairs present a root-mean-square deviation (RMSD) value close to 2 Å, whereas only the CH-CH system shows a small amount of structure population within that range, when forming a duplex with RNA. The CH2-CH systems also presents short-lived duplex structures, although as the simulation progresses, the pairing is lost, and globular un-ordered configurations are formed. The information of RMSD versus time for each of the three replicas shows that the duplex pairing is maintained for ~100 ns in each case until the duplex is lost (Figure S1). In the case of duplex with an RNA strand, the CH-CH system presents RMSD populations within ~1–4 Å when considering the inner base pairs, which suggest a possible stable duplex formation. The percentage of stacking for each system from base steps 2–8 is presented in Table 1 (see Materials and methods for stacking analysis criteria). As a control, the data of the reference structure is presented, which is expected to show 100% stacking base steps throughout the simulation. For the DNA duplex, on average, the CH2-CH2 system presents the lowest stacking percentage and the CH-CH the highest. In the case of duplex with RNA, the CH2-CH2, CH2-CH, and CH-CH2 present similar values across every base step, and the CH-CH systems showed stacking configurations of ~90%–98%, which suggest a stable duplex is indeed occurring.

To extract the most populated structure, we performed a clustering analysis using the aggregated three independent calculated trajectories. The most populated clusters for each double-strand system are represented in Figure 8. The number of clusters detected for each system was 95, 46, 39, and 38 for DNA and for RNA, and the detected clusters were 72, 35, 39, and 7 for CH2-CH2, CH2-CH, CH-CH2, and CH-CH, respectively. Elevated number of clusters shows a higher variability of structures, which for both the DNA and RNA duplex is the CH2-CH2 system, and the lowest cluster count corresponds to the CH-CH system. The clustering analysis was also informative regarding the level of convergence of the studied systems. When performing an analysis of the number of structures per cluster through time, we can observe the CH-CH/RNA system as being the most uniformly distributed as the simulation progresses (Figure S2). The processes on which the duplexes with DNA and RNA evolve over time and unfold for most systems is complex and difficult to characterize. It is comparable with the myriad of structural interactions observed in normal-occurring fraying events in the terminal base pairs of both DNA/RNA duplexes.[18] Visual inspection of the calculated trajectories shows long-lived fraying events that could potentially start the un-pairing of nucleobases, and thus the loss of the duplex hybridization. To test this theory, we have conducted a 3.0-μs simulation of the CH2-CH2 and CH-CH systems with DNA and RNA as the complementary chain on which a distance restraint has been added in the terminal base pairs (residues 1/18 and 9/10) for 1.5 μs. After this time, the restraint is shut off and the system is allowed to run for the remaining 1.5 μs. The results of this test are presented in Figure 9 (see also Video S1 that shows the entire sampled trajectory for the CH2-CH2(DNA/RNA) and CH-CH(DNA/RNA) systems). It is clear that the CH-CH/DNA and CH-CH/RNA systems present lower RMSD values both before and after the 1.5 μs because of its capacity to form and maintain a duplex structure. Figures 9A–9H show selected structural snapshots to provide a visual representation of the distorted conformers observed. Structure A, which represents CH2-CH2/RNA with the distance restraints, remains in a duplex-like form, although the central base pairs are already mismatched, causing an RMSD of almost 8 Å at the ~250 ns time mark. This same system deviates highly once the restraint is turned off at the 1.5-μs mark, with multiple fraying events at both ends of the structure and completely losing the WC pairing. An unpaired structure is shown at structure B with a deviation of ~13 Å. In contrast, for the CH-CH/DNA system, a stable duplex structure is presented in Figure 9C, and when the restraint is turned off, fraying increases, causing the duplex formation to lose the pairing (Figure 9D).

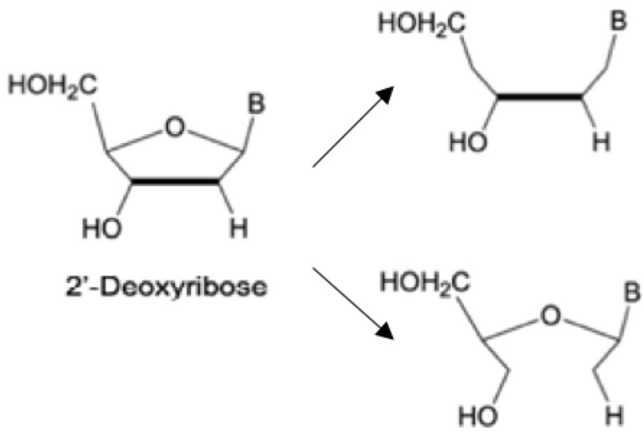


Figure 4. Acyclic analogs of DNA

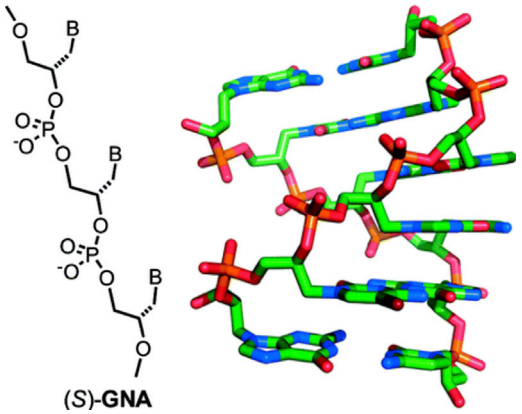


Figure 5. Homo-duplex formed by glycol nucleic acid (GNA) (from Meggers and Zhang[15])

For the CH2-CH2/RNA system, the structure at Figure 9F presents a deviation of ~8 Å at the 1.1-μs mark with multiple mispairing throughout the entire chain and when the restraint between the terminal base pairs is turned off, the double helix winds upon itself, forming globular disordered structures (structure in Figure 9F). The CH-CH/RNA system forms a stable duplex structure with the restraint with the occasional flip of the central cytosine residue at position 5 that causes the RMSD to increase to ~6 Å (structure in Figure 9G). When the restraint is turned off, there is a clear increase in fraying events at both ends of the chain, but the central seven residues remain paired for the rest of the simulation. Elevated deviation values are mainly originated by the constant flipping of the central C base, as observed in the structure in Figure 9H at the 1.9-μs mark. This result suggests that the stability of both CH-CH/DNA and CH-CH/RNA systems is not entirely dependent on fraying events, and the composition of the backbone (CH2-CH2 versus CH-CH) is responsible for the ability of the system to form a stable and long-lived duplex pairing.

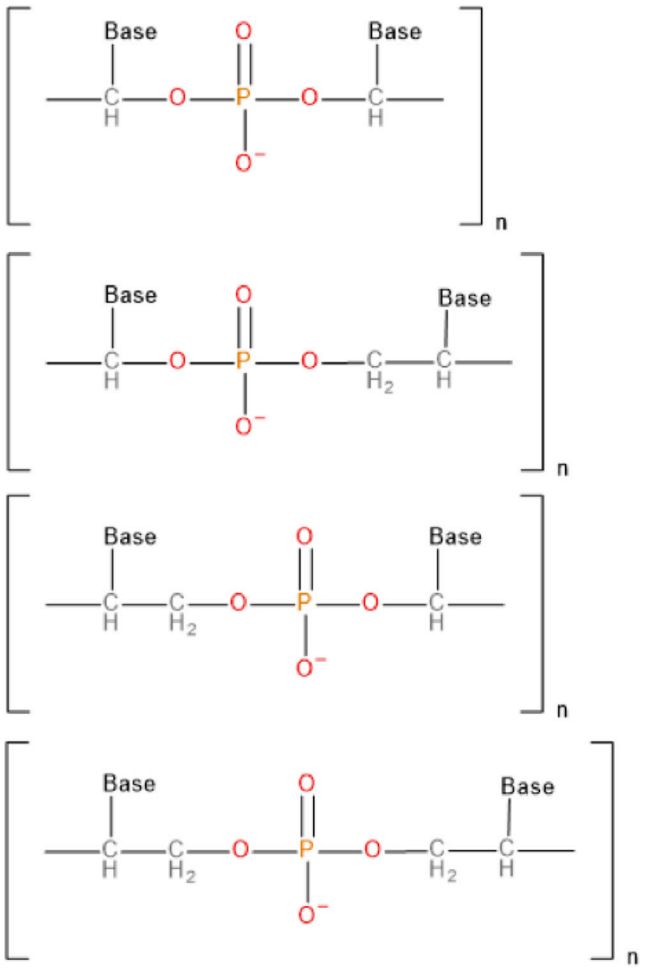


Figure 6. Backbone analogs
In the manuscript we will refer to these systems as (from top to bottom) CH-CH, CH-CH2, CH2-CH, and CH2-CH2.

Base-pair influence of the CH-CH backbone in duplex with DNA and RNA

To study the structural differences that the DNA/RNA hybridization with the CH-CH produces with a canonical RNA duplex, we calculated the sequence $(UGGGCGGGA)_2$ using the same protocol as the rest of the systems (three independent copies, each 5 μs of sampling time). This will serve as the reference structure in the same way we are using our DNA reference structure from PDB: 6ed9.

Selected base-pairing parameters are presented in Table 2. Because we have observed that both CH-CH/DNA and CH-CH/RNA systems form a stable duplex (more with RNA than DNA), we expect the base-pairing values to be close to the canonical reference values. No significant differences are observed with the shear, stretch, and stagger properties. Buckle reference for B-DNA is 0.5, and we obtained a value of −5.2 in our RNA reference simulations. The CH-CH/DNA system is close to the B-DNA value with 0.6, which is similar to the CH-CH/RNA value of −0.31, although this is the highest standard deviation error among the measured parameters (±20.2°), which suggests high value fluctuations. The propeller value of the CH-CH/DNA system is 1.34, whereas the same system with RNA is −8.93, closer to the RNA reference of −9.51. The DNA backbone is shifting the propeller value of the nucleobases to positive values, reducing stacking interactions and stability.

Another measure to study the influence of the CH-CH backbone into the formation of a duplex with DNA and RNA is the distance between the carbons bonded directly to the nucleobase and the C1′ carbon that forms the glycosidic bond with DNA and RNA. The results are presented in Table 3. The reference structures for both DNA and RNA show a value of ~10.7 Å, as expected from the literature.[19] We observe that the CH-CH/DNA duplex shows increased values throughout the sequence with highest values (~14.4 Å) in the central base pairs. The average distance considering all the base pairs is 13.1 Å. In the case of

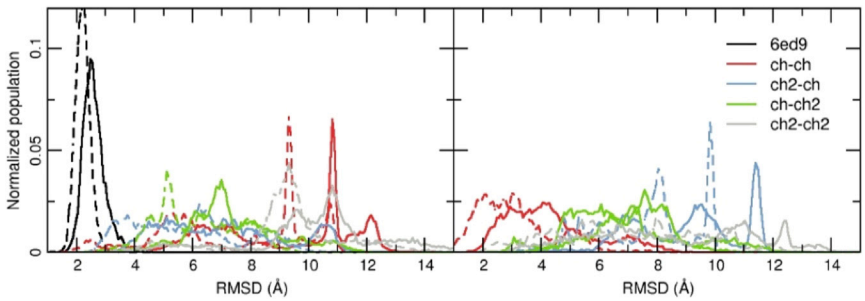


Figure 7. Normalized population of root-mean-square (RMS) values using the first frame as a reference

Solid lines are considering all residues, and dashed lines are considering only the central base pairs. Left panel is RMS values with DNA with black line showing the control structure; right panel is with RNA.

the CH-CH/RNA system, the overall average is 11.1 Å, which is considerably lower. The central base pairs are in the range of ~10.2 to 11.1 Å, which is closer to the reference values.

Referring to the previous data about nucleobase pairing, we can observe that the average reference opening value is 0.6°. Even though the opening average for CH-CH/RNA is 9.51, in contrast with 4.93 of the CH-CH/DNA, this conformation allows for the CH-C1′ carbons to come closer, which is observed in the information presented in Table 3 and can potentially help in maintaining the structural duplex stability. The increased opening value could also explain the multiple flipping events of the cytosine at position 5.

DISCUSSION

Antisense technology, as well as the need to explore the many possibilities in nucleic acid analog design, has led to many alternative analogs being tested for biological activity.[11] But it is not sensible to synthesize an oligo analog with a chemically modified backbone to be tried as an antisense agent without first testing its ability to strongly duplex with the target, usually a natural mRNA. Our present results have explored a series of acyclic nucleic acid analogs to test the capabilities of duplex formation using molecular dynamics simulations.

We would like to stress the importance of sampling time regarding MD simulations, specifically, simulations of double-stranded DNA or RNA duplexes of the size presented in this work. It is clear that without extended sampling time, the results would be biased, erroneous, and probably wrong. The importance of multiple independent copies of the same biomolecular system has been explored by previous published works and independent groups.[19,20] We present the extended simulations of acyclic backbone systems to guide the subsequent synthesis and biological experimental testing of the proposed structures.

Our results explored four different models to consider the formation of stable DNA/RNA duplexes. For the most part, the simulations

Table 1. Percentage of stacking of the backbone-modified strand between base steps calculated using the entire trajectory data

	Base step - DNA						
	2/3 (%)	3/4 (%)	4/5 (%)	5/6 (%)	6/7 (%)	7/8 (%)	Average
63d9[a]	100.00	100.00	100.00	100.00	100.00	100.00	1.00
CH2-CH2	25.40	27.00	16.50	44.80	25.90	11.90	0.25
CH2-CH	33.30	47.60	45.00	71.00	40.00	27.70	0.44
CH-CH2	37.60	65.40	51.80	15.40	61.40	14.90	0.41
CH-CH	66.90	68.60	45.10	11.90	61.80	77.30	0.55
RNA							
CH2-CH2	28.59	30.84	22.29	7.55	51.25	33.81	0.29
CH2-CH	12.06	73.50	24.69	21.53	30.76	3.57	0.28
CH-CH2	30.58	56.59	81.48	33.20	23.24	7.19	0.39
CH-CH	91.02	97.28	53.76	54.91	99.95	98.82	0.83
Single strand							
CH2-CH2	18.98	23.68	11.30	14.79	10.33	2.66	0.14
CH2-CH	30.65	13.81	5.36	7.79	17.50	8.52	0.14
CH-CH2	51.02	49.78	23.20	22.77	24.46	2.75	0.29
CH-CH	63.87	51.16	27.25	13.79	84.99	27.25	0.45

Details of stacking criteria are explained in the Materials and methods.
[a]Normal duplex.

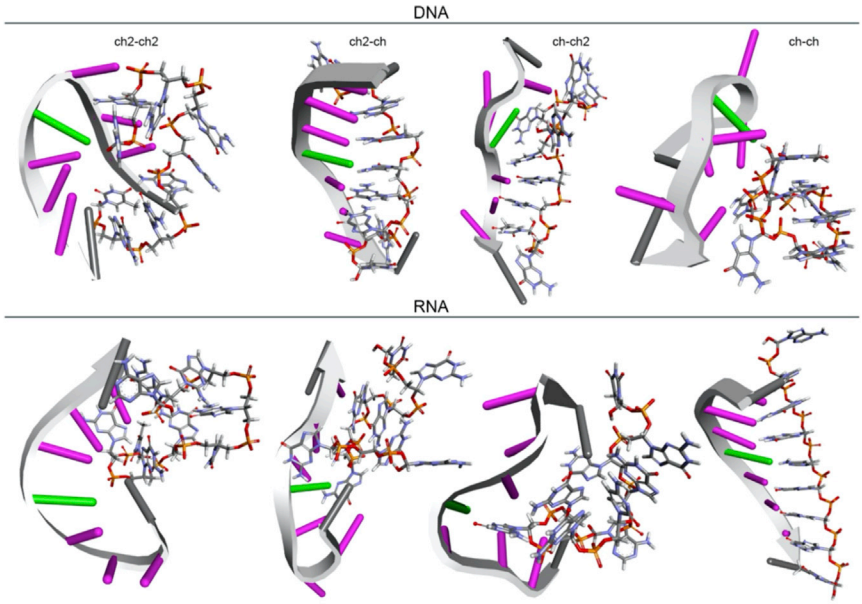


Figure 8. Representative structures from the most populated cluster for each system
Analysis performed using the entire sampled data; clustering details are described in the Materials and methods.

showed that non-stable duplexes are formed, leading to globular unordered structures that mainly formed interactions within the nucleobases and the backbone with no discernable shape. It was to our surprise that the -CH(base)-PO_4-CH(base)- system forms a stable duplex with RNA. Analysis of the structure of the CH-CH/RNA duplex suggests that the stability is mainly due to the following: (1) lack of strong fluctuations of the nucleobases, which allows the stacking interactions within neighboring bases to stabilize the duplex form of the analog/RNA complex; (2) distance between the carbon atom that is bonded to the nucleobase and the C1′ atom of the complementary chain is close in value to the canonical DNA/DNA and RNA/RNA duplexes; and (3) nucleobase inter-step distance is also close to the canonical DNA/DNA and RNA/RNA duplex value, which promotes the correct formation of WC pairing.

Our results also suggest the importance of having converged simulations, multiple copies, and extended sampling time. Considering our simulations are extended to only ~100 ns, the studied systems in this work all remain in a stable duplex configuration, which would produce a biased result suggesting that the modified backbone could work as a potential antisense analog. Most of these analogs will not work experimentally, they are just too flexible to pair to DNA/RNA. The stacking interactions between nucleobases is not enough to keep the bases oriented to form a stable duplex. We also note that the results can depend on length; however, we chose a specific 9-mer for our studies, of which the structure of the duplex has been determined by NMR spectroscopy.[21]

Analogs that have been used as antisense agents, such as PNAs[22] and morpholino analogs,[23] have not been subjected to this stringent objective comparative test. Although there have been MD calculations performed on PNA,[24] there have been none as far as we are aware on morpholino analogs. In future work we are proceeding to test and compare these analogs using our MD methods. Given the large number of analogs that have been developed and can be conceived, this is a practical approach to antisense drug design, avoiding expensive and fruitless synthetic procedures. Of course, this does not provide any information on other important properties of antisense agents, such as cell penetration,[25] non-selective effects, or *in vivo* distribution.

What we have shown here is that MD calculations (with the stringent details given) are a valuable tool to guide decisions in the design of oligo analogs that are intended to duplex with a target mRNA. CAMD should be used as a valuable tool in guiding further antisense analog design. We will now proceed with the synthesis of the only acyclic analog CH-CH that shows evidence of stable duplex formation with its RNA complement.

MATERIALS AND METHODS

Building and generation of models

As a reference, we used the 9-mer NMR structure d $(TGGGCGGGA)_2$ (PDB: 6ED9). The 10 deposited structures were used to create an average structure. The DNA variants CH-CH, CH-CH2, CH2-CH (this corresponds to GNA), and CH2-CH2 (see Figure 6) with modified backbone were constructed using DS Visualizer. Structural and geometry optimizations for each proposed backbone were performed using quantum mechanical calculations at the DFT level of theory (M06-2x/6-311G(d,p)[26]). Bond, angles, dihedrals, and improper parameters were extracted from the General AMBER Force Field (GAFF).[27] Charges were calculated using the restricted electrostatic potential (RESP) at the HF/6-31G* level of theory as required for the AMBER force field.

Molecular dynamics protocol

MD simulations were performed using the parm99 force field[28] with the bsc0[29] and the OL15 modifications.[30,31] The nonstandard backbone was described by the GAFF using the OL15 parameters for the nucleobases. Using the models described above as starting structures, the topology and coordinate files were created using the LEaP module present in AmberTools 1. Explicit water was added using the optimum point charge (OPC) water model with a truncated octahedral box using a minimum distance of 10 Å around the solute and the edge of the box.[32] Na^+ counter ions were added to reach a net charge of 0 using the Joung-Cheatham ion parameters.[33] Periodic boundary conditions were applied. A non-bonded cutoff of 8 Å was employed and the SHAKE algorithm[34] to contain hydrogen bonds.

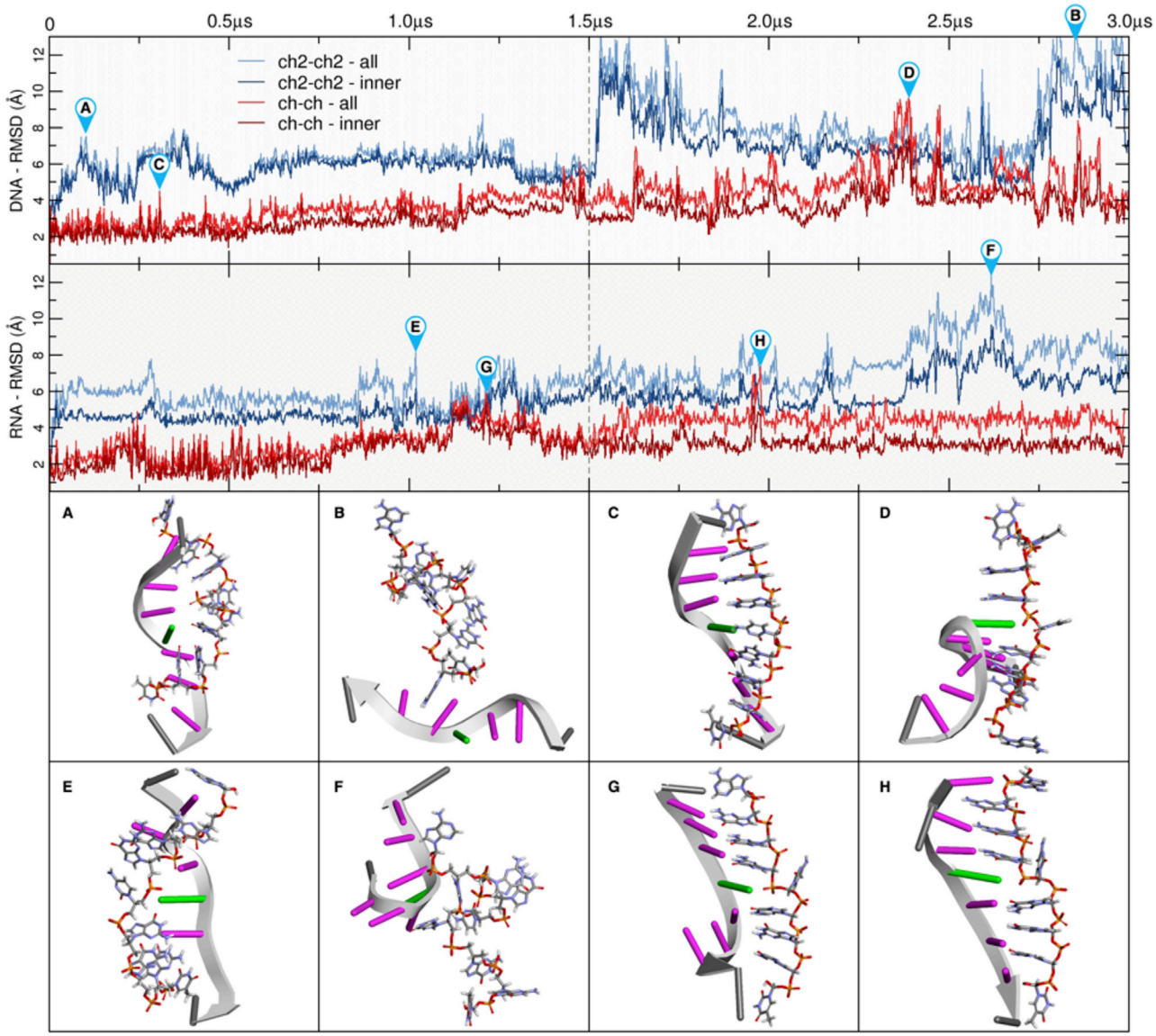


Figure 9. Simulations of the CH2-CH2 and CH-CH systems with DNA (top) and RNA (bottom) as the complementary chain

Light blue and light red represent the RMSD using the first frame as reference including all the residues; dark blue and dark red are the RMSD considering only the inner base pairs (2–8 and 11–17). A distance restraint of 5 kcal/mol·Å^2 is applied between residues 1/18 and 9/10 (terminal base pairs) for a duration of 1.5 μs. The restraint is removed after this time (gray dashed line at the 1.5-μs mark), and the system is allowed to freely sample for the remainder of 1.5 μs.

Long-range electrostatics were calculated using the particle mesh Ewald method with default parameters.[35,36] Each model was initially minimized using 500 steps of steepest descent and 500 steps of conjugated gradient using a harmonic restriction on the solute with a value of 20 kcal/mol·Å. A heating process was done using the same restriction on the solute and slowly heating for 50 ps to a final temperature of 300K. Temperature was maintained with Langevin dynamics (collision frequency = 2 ps^{-1}). After heating, the restraints on the DNA atoms were slowly reduced from 20 to 0.5 kcal/mol·Å in 5 steps, with each step lasting 50 ps. Three independent copies were minimized for each system, and unrestrained MD was then conducted at NTP conditions for at least 10 μs, with each copy using the pmemd.cuda module.[37–39] The resulting trajectories were concatenated, and all analyses and further post-processing were performed in this aggregated trajectory data. All simulations were run using AMBER18,[40] and analysis was done using CPPTRAJ.[41] Clustering was performed using every 1/10th of a frame of the three independent runs for each case using the hierarchical agglomerate algorithm[42] using an epsilon value of 5 and including all the residues (heavy atoms only). Stacking criteria were based on previous work by Hayatshahi et al.[43] In brief, for each nucleobase to be considered stacked, a minimum distance between any two heavy atoms is 4 Å; the center of

Table 2. Selected base-pair parameters for the CH-CH system

		DNA Ref		RNA Ref		CH-CH/DNA		CH-CH /RNA	
	B-DNA Ref21	**Avg.**	**SD**	**Avg.**	**SD**	**Avg.**	**SD**	**Avg.**	**SD**
Shear (Å)	0.00	−0.08	0.09	−0.10	0.12	−0.20	0.47	0.03	0.24
Stretch (Å)	−0.15	−0.06	0.00	−0.06	0.04	0.00	0.33	0.17	0.53
Stagger (Å)	0.09	0.00	0.11	−0.10	0.16	0.09	0.12	−0.07	0.73
Buckle (°)	0.50	0.79	3.67	−5.29	3.91	0.69	4.78	−0.31	20.29
Propeller (°)	−11.40	−7.64	1.29	−9.51	3.49	1.34	4.88	−8.93	7.11
Opening (°)	0.60	−0.28	0.58	−0.22	1.35	4.93	7.95	9.51	19.03

Values are calculated considering only the seven inner base pairs and the entire sampling trajectory. Avg., average; Ref, reference.

mass of two bases should be a minimum of 5 Å, and angles between the normal to the bases is between 0° and 45° or 135° and 180°. Any parameters not falling in these criteria are not counted as a stacking base-step event. Helicoidal parameters were performed using the NASTRUCT command as available in CPPTRAJ creating new reference libraries for each of the modified nucleobases; these files are available in the Supplemental Information.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.omtn.2020.11.023.

ACKNOWLEDGMENTS

J.S.C. wishes to acknowledge J.L. Syi and J. Maizel, Mathematical Biology Section, NCI, for their help with preliminary unpublished results before he left the NCI in 1990. We also thank Dror Sherf for help with preliminary searches for MD analyses of DNA. We also acknowledge the cooperation of Dr. E. Yavin, Hebrew University, Jerusalem, for advice on PNAs. R.G.-M. wishes to acknowledge the Texas Advanced Computing Center for their generous time in the Longhorn supercomputer and the Center for High Performance Computing at the University of Utah for their support. B.A. wishes to thank the Israel Science Foundation (ISF, Grant No. 1023/ 18).

Table 3. Distance (Å) between the carbon between the phosphate groups of the CH-CH system and the C1′ carbon of the glycosidic bond for DNA and RNA

	DNA Ref		RNA Ref		CH-CH/ DNA		CH-CH/ RNA	
Base Pairs	**Avg.**	**SD**	**Avg.**	**SD**	**Avg.**	**SD**	**Avg.**	**SD**
1/18	8.5	2.4	12.9	2.7	14.2	6.4	11.0	6.0
2/17	10.6	0.3	10.7	0.2	12.4	5.6	10.5	3.8
3/16	10.7	0.2	10.7	0.2	12.8	3.9	10.9	2.2
4/15	10.8	0.2	10.7	0.2	13.0	2.6	10.2	1.2
5/14	10.8	0.2	10.7	0.2	14.4	4.2	10.5	0.7
6/13	10.7	0.2	10.7	0.2	14.4	3.7	11.1	0.9
7/12	10.7	0.2	10.7	0.2	11.3	2.0	10.6	0.5
8/11	10.7	0.2	10.7	0.2	12.1	3.5	9.6	1.7
9/10	10.4	2.0	12.3	2.4	13.0	4.8	15.0	2.9

Average (Avg.) is calculated using the entire sampled trajectory and including the three independent copies.

AUTHOR CONTRIBUTIONS

Conceptualization, B.A. and J.S.C.; Formal Analysis, R.G.-M.; Investigation, R.G.-M. All authors designed research and wrote the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Dalton, L. (2005). Simpler Than DNA: bare-bones nucleic acid structures may provide insight into pre-biotic past. Chem. Eng. News *83*, 13.
2. Astbury, W., and Bell, F.O. (1938). Some recent developments in the X-ray study of proteins and related structures. Cold Spring Harb. Symp. Quant. Biol. *6*, 109–121.
3. J.S. Cohen, ed. (1977). Oligodeoxynucleoties: Antisense Inhibitors of Gene Expression (Macmillan).
4. Stein, C.A., Subasinghe, C., Shinozuka, K., and Cohen, J.S. (1988). Physicochemical properties of phosphorothioate oligodeoxynucleotides. Nucleic Acids Res. *16*, 3209–3221.
5. Stein, C.A., and Castanotto, D. (2017). FDA-Approved Oligonucleotide Therapies in 2017. Mol. Ther. *25*, 1069–1075.
6. Rüger, J., Ioannou, S., Castanotto, D., and Stein, C.A. (2020). Oligonucleotides to the (Gene) Rescue: FDA Approvals 2017-2019. Trends Pharmacol. Sci. *41*, 27–41.
7. Caruthers, M.H. (1991). Chemical Synthesis of DNA and DNA Analogs. Acc. Chem. Res. *24*, 278–284.
8. Leumann, C.J. (2002). DNA analogues: from supramolecular principles to biological properties. Bioorg. Med. Chem. *10*, 841–854.
9. Nielsen, P.E. (1995). DNA analogues with nonphosphodiester backbones. Annu. Rev. Biophys. Biomol. Struct. *24*, 167–183.
10. Gupta, A., Mishra, A., and Puri, N. (2017). Peptide nucleic acids: Advanced tools for biomedical applications. J. Biotechnol. *259*, 148–159.
11. Karkare, S., and Bhatnagar, D. (2006). Promising nucleic acid analogs and mimics: characteristic features and applications of PNA, LNA, and morpholino. Appl. Microbiol. Biotechnol. *71*, 575–586.
12. Schneider, K.C., and Benner, S.A. (1990). Oligonucleotides containing flexible nucleoside analogs. J. Am. Chem. Soc. *112*, 453–455.
13. Augustyns, K., Rozenski, J., Van Aerschot, A., Janssen, G., and Herdewijn, P. (1993). Synthesis of 2,4-dideoxy-.beta.-D-erythro-hexopyranosyl nucleosides. J. Org. Chem. *58*, 2977–2982.

14. Declercq, R., Van Aerschot, A., Read, R.J., Herdewijn, P., and Van Meervelt, L. (2002). Crystal structure of double helical hexitol nucleic acids. J. Am. Chem. Soc. *124*, 928–933.

15. Meggers, E., and Zhang, L. (2010). Synthesis and properties of the simplified nucleic acid glycol nucleic acid. Acc. Chem. Res. *43*, 1092–1102.

16. Johnson, A.T., Schlegel, M.K., Meggers, E., Essen, L.O., and Wiest, O. (2011). On the structure and dynamics of duplex GNA. J. Org. Chem. *76*, 7964–7974.

17. Jaroszewski, J.W., Syi, J.L., Maizel, J., and Cohen, J.S. (1992). Towards rational design of antisense DNA: molecular modelling of phosphorothioate DNA analogues. Anticancer Drug Des. *7*, 253–262.

18. Zgarbová, M., Otyepka, M., Šponer, J., Lankaš, F., and Jurečka, P. (2014). Base Pair Fraying in Molecular Dynamics Simulations of DNA and RNA. J. Chem. Theory Comput. *10*, 3177–3189.

19. Knapp, B., Ospina, L., and Deane, C.M. (2018). Avoiding False Positive Conclusions in Molecular Simulation: The Importance of Replicas. J. Chem. Theory Comput. *14*, 6127–6138.

20. Galindo-Murillo, R., Roe, D.R., and Cheatham, T.E.I., 3rd (2015). Convergence and reproducibility in molecular dynamics simulations of the DNA duplex d(GCACGAACGAACGAACGC). Biochim. Biophys. Acta *1850*, 1041–1058.

21. Olson, W.K., Bansal, M., Burley, S.K., Dickerson, R.E., Gerstein, M., Harvey, S.C., Heinemann, U., Lu, X.J., Neidle, S., Shakked, Z., et al. (2001). A standard reference frame for the description of nucleic acid base-pair geometry. J. Mol. Biol. *313*, 229–237.

22. Nielsen, P.E. (1999). Applications of peptide nucleic acids. Curr. Opin. Biotechnol. *10*, 71–75.

23. Warren, T.K., Shurtleff, A.C., and Bavari, S. (2012). Advanced morpholino oligomers: a novel approach to antiviral therapy. Antiviral Res. *94*, 80–88.

24. Soliva, R., Sherer, E., Luque, F.J., Laughton, C.A., and Orozco, M. (2000). Molecular Dynamics Simulations of PNA·DNA and PNA·RNA Duplexes in Aqueous Solution. J. Am. Chem. Soc. *122*, 5997–6008.

25. Miller, K.J., and Das, S.K. (1998). Antisense oligonucleotides: strategies for delivery. Pharm. Sci. Technol. Today *1*, 377–386.

26. Zhao, Y., and Truhlar, D.G. (2007). The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of 4 M06-Class Functionals and 12 Other Functions. Theor. Chem. Acc. *120*, 215–241.

27. Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A., and Case, D.A. (2004). Development and testing of a general amber force field. J. Comput. Chem. *25*, 1157–1174.

28. Cheatham, T.E., 3rd, Cieplak, P., and Kollman, P.A. (1999). A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. J. Biomol. Struct. Dyn. *16*, 845–862.

29. Pérez, A., Marchán, I., Svozil, D., Sponer, J., Cheatham, T.E., 3rd, Laughton, C.A., and Orozco, M. (2007). Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. Biophys. J. *92*, 3817–3829.

30. Zgarbová, M., Šponer, J., Otyepka, M., Cheatham, T.E., 3rd, Galindo-Murillo, R., and Jurečka, P. (2015). Refinement of the Sugar-Phosphate Backbone Torsion Beta for AMBER Force Fields Improves the Description of Z- and B-DNA. J. Chem. Theory Comput. *11*, 5723–5736.

31. Galindo-Murillo, R., Robertson, J.C., Zgarbová, M., Šponer, J., Otyepka, M., Jurečka, P., and Cheatham, T.E., 3rd (2016). Assessing the Current State of Amber Force Field Modifications for DNA. J. Chem. Theory Comput. *12*, 4114–4127.

32. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of Simple Potential Functions for Simulating Liquid Water. J. Chem. Phys. *79*, 926.

33. Joung, I.S., and Cheatham, T.E., 3rd (2008). Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. J. Phys. Chem. B *112*, 9020–9041.

34. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H.J.C. (1977). Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. J. Comput. Phys. *23*, 327–341.

35. Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., and Pedersen, L.G. (1995). A Smooth Particle Mesh Ewald Method. J. Chem. Phys. *103*, 8577.

36. Cheatham, T.E., 3rd, Miller, J.L., Spector, T.I., Cieplak, P., and Kollman, P.A. (1998). Mol. Model. Nucl. Acids.

37. Götz, A.W., Williamson, M.J., Xu, D., Poole, D., Le Grand, S., and Walker, R.C. (2012). Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. J. Chem. Theory Comput. *8*, 1542–1555.

38. Le Grand, S., Götz, A.W., and Walker, R.C. (2013). SPFP: Speed without Compromise—A Mixed Precision Model for GPU Accelerated Molecular Dynamics Simulations. Comput. Phys. Commun. *184*, 374–380.

39. Salomon-Ferrer, R., Götz, A.W., Poole, D., Le Grand, S., and Walker, R.C. (2013). Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. J. Chem. Theory Comput. *9*, 3878–3888.

40. Case, D.A., Darden, T.A., Cheatham, T.E., 3rd, Simmerling, C.L., Wang, J., Duke, R.E., Luo, R., Walker, R.C., Zhang, W., Merz, K.M., et al. (2014). AMBER, 14.

41. Roe, D.R., and Cheatham, T.E., 3rd (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. J. Chem. Theory Comput. *9*, 3084–3095.

42. Shao, J., Tanner, S.W., Thompson, N., and Cheatham, T.E. (2007). Clustering Molecular Dynamics Trajectories: 1. Characterizing the Performance of Different Clustering Algorithms. J. Chem. Theory Comput. *3*, 2312–2334.

43. Hayatshahi, H.S., Henriksen, N.M., and Cheatham, T.E., 3rd (2018). Consensus Conformations of Dinucleoside Monophosphates Described with Well-Converged Molecular Dynamics Simulations. J. Chem. Theory Comput. *14*, 1456–1470.