C3′ Stereocentre of Furanose Determines the Helicity of Double-Stranded Nucleic Acids: Towards Etiology of Nucleic Acids

Anuj Kumar,[1,2] Amol Tagad,[1] and G. Naresh Patwari[1,*]

[1]Department of Chemistry, Indian Institute of Technology Bombay, Powai, Mumbai 400076 India; [2]Prithwi Chand Vigyan College, Jai Prakash Vishwavidyalaya, Chapra-Saran 841301 Bihar India.

ABSTRACT Ribose containing double-stranded nucleic acids exhibit helical structure, whereas sugar modified (xeno) nucleic acids may exhibit different structural features. The structural landscape of four stereo variants of furanosal nucleic acids and their C2′ deoxy counterparts, explored with molecular dynamics simulations, suggest that the configuration at the C3′ position plays a pivotal role in determining the helicity. The C3′ stereocentre acts as toggle-switch for the helix to ladder structural transformation by changing the nature of intra-strand interactions resulting in the optimal helices for ribose containing double-stranded nucleic acids. Interestingly, lack of chirality at the C2′ position results in better quality helices than inversion of stereochemistry relative to ribose. The etiology of furanosal-RNA over other furanoses can be hypothesized based on the helical structure, which can effectively be exploited by the biological machinery.

SIGNIFICANCE The double helix structure of furanosal RNA is governed by the configuration at the C3′ position. Furanose sugars such as xylose and lyxose where in the configuration at the C3′ position is inverted relative to the ribose do not form double helix structure, instead result in ladder-like structure. The configuration at the C3′ position acts as a toggle switch for the helix to ladder structural transition. Among four furanose sugars viz., ribose, arabinose, xylose, and lyxose, the double-stranded nucleic acids incorporating ribose form helices with best aspect ratio between major and minor grooves.

INTRODUCTION

The chemical etiology of nucleic acids has been subject of great curiosity (1), and several synthetic approaches have been employed to address this issue by varying one or more 'structural elements' of the canonical nucleic acids viz., the nucleobases, the sugar and the phosphate linkage (2, 3). Xeno nucleic acids (XNA) are sugar modified nucleic acids (4). which include chemical modification of the ribose sugar with several variants that may or may not include furanose sugar. An interesting question raised by Eschenmoser in this context is (5), "why did nature choose furanosyl-RNA and not pyranosyl-RNA?" and concluded based on the experimental evidence that pyranosyl-RNA binds more strongly and selectively relative to its natural counterpart the furanosyl-RNA (5). However, the primary question remains unanswered. Subsequently, many variants of XNAs with non-furanosyl moieties such as cyclohexene (6), glycol (7), hexitol (8), peptides (9), threose (10), and others have been reported, which demonstrate their ability to pair can be modulated by appropriately modifying the 'sugar' moiety even while the nucleobases are conserved to form Watson-Crick pairs. A large number of XNAs have been explored in various biological technologies such as antisense oligonucleotides, siRNAs, and aptamers (11, 12), prominent among them being C2′ modified analogs. On the other hand, arabino nucleic acids, which have inverted C2′ stereocentre are known to equilibrate between single-stranded hairpin structure and canonical B-form (13).

The ability of canonical nucleic acids to elongate and structurally transform, albeit transitorily, have been captured during the transcription process (14). However, a standalone ladder-like structure of canonical nucleic acids has not been reported, whilst structural transformation of helical B-DNA into ladder-like structure can be explored by quenching of dispersion energy component of the potential function (15) or by force stretching (16). Interestingly, substitution of ribose with xylose results in a ladder-like structure for the double-stranded furanosyl-xylo nucleic acid (xyloNA) (17), and its deoxy analog (DxyloNA) (18). The structural landscape of xyloNA and DxyloNA, explored with the aid of molecular dynamics simulations, reveals their propensity to form ladder-like structures along with left-handed helix (19–22). Interestingly, the thermodynamic stability of DxyloNA was found to be higher than DNA (23), which is in contrast to the facile unzipping of xyloNA relative to RNA (24). While there has been a flurry of activity to explore the structure and dynamics of both native and deoxy versions of xylo nucleic acids, using both experimental and computational tools, investigations on nucleic acid constructs with the C2′ epimer of xylose, the lyxose, remain largely unexplored (25). An interesting question that arises at this stage is, "how does the configuration at the C2′ and C3′ stereocentres influence the structural landscape of double-stranded nucleic acids?" More specifically, "is the ladder-like structure of double-stranded xyloNA (and DxyloNA) is associated with the inversion of C3′ stereocentre, relative to RNA (and DNA)?" To this end, molecular dynamics simulations on double-stranded nucleic acids consisting of four variants of furanose and two of C2′ deoxy counterparts, with the sequence 5′-[GCAU(T)]$_7$G-3′, were carried out to understand the role of sugar stereochemistry, and the results are discussed herein.

METHODOLOGY

The structures of the double-stranded nucleic acids of RNA, arabinoNA, xyloNA, and lyxoNA and their corresponding deoxy counterparts DNA and DxyloNA were explored using MD simulations. The initial structures of double-stranded B-RNA and B-DNA with sequence 5′-**(**GCAU)$_7$G-3′ along with its complementary strand was generated using the make-NA server (28). The initial structure of the double-stranded xyloNA structure was obtained from Ramaswamy *et al*. (20) The structures of arabino- and lyxose- nucleic acids were generated by inverting the stereochemistry at the C2' position of furanose sugar in the ribo- and xylo- nucleic acids, respectively. Further, the initial structure of deoxy-xylose form was generated by replacing the OH group with H in the xyloNA. All-atom molecular dynamics simulations were carried out for 29-mer double-stranded nucleic acids with Amber-14SB forcefield (26) using single precision GROMACS 2018.1 (29, 30). The double-stranded nucleic acid was oriented along the principal axis of a rectangular box of dimensions 6x6x14 nm. The simulation box contained about 16000 TIP3P (31) water molecules along with 56 sodium cations in each case to achieve the electro-neutrality of the system. The potential energy of the system was minimized using the steepest descent algorithm (32) with a tolerance force of 100 kJ mol^{-1} to remove any bad contacts and steric clashes. After energy minimization, the system was equilibrated in the NVT ensemble by employing a velocity-rescale thermostat (33) with a relaxation time of 0.1 ps and a reference temperature of 298 K for 100 ps. This was followed by NPT equilibration for 200 ps using Berendsen barostat (34) with a relaxation time of 2 ps and reference pressure of 1 bar. For the production runs, the temperature of the system was maintained at 298 K using a velocity rescaling thermostat with a relaxation time of 0.1 ps and the pressure of the system was maintained at 1 bar using the Parrinello-Rahman barostat (35) with a relaxation time of 2 ps. The equations of motion during the simulations were integrated using the leapfrog algorithm (36) with a time step of 2 fs. The length of MD simulations was 100 ns. The coordinates were saved every 10 ps during the course of simulation. For each of the six double-stranded nucleic acids, five independent trajectories of 100 ns were calculated to evaluate the and statistical validity of the structures obtained.

RESULTS and DISCUSSION

The structural evolution of all the six sets of double-stranded nucleic acids was followed through the course of the simulations and the converged structures at the end of 100 ns simulation are depicted in Figure 1. This figure illustrates that the helical structure for RNA, arabinoNA, and DNA is retained at the end of the simulation. On the other hand, for the xyloNA, lyxoNA, and DxyloNA an apparent loss of helicity can be observed. *Prima-facie*, it appears that the inversion of configuration of the furanose ring at C3′ position in

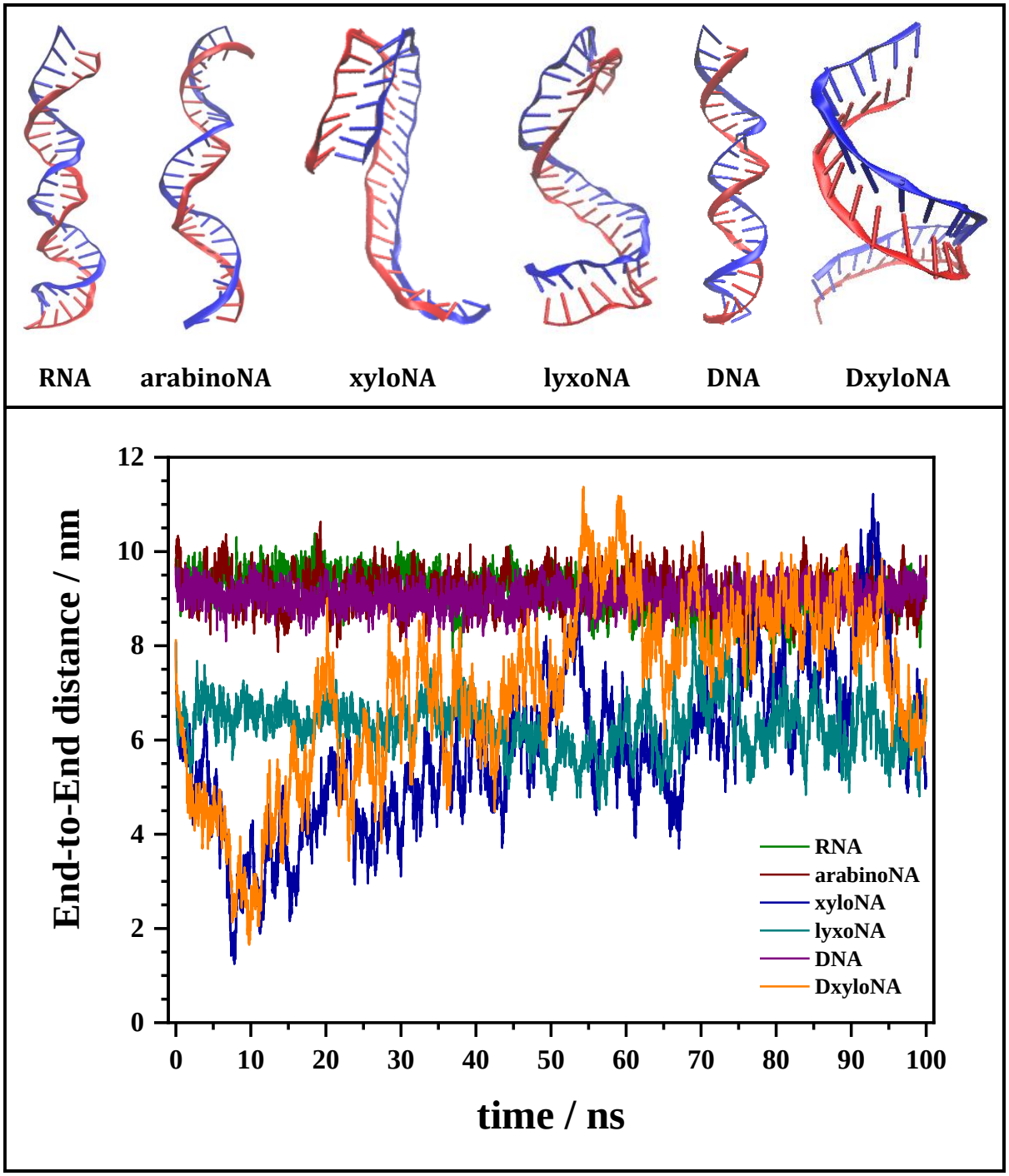

FIGURE 1 Snapshots of structures of double-stranded nucleic acids obtained after 100 ns of simulation. The structures of RNA, arabinoNA, and DNA appear helical while the structures of xyloNA, lyxoNA and DxyloNA show non-helical structures. The bottom panel shows that plot of end-to-end distance (see text for details), which indicates that lyxoNA, xyloNA, and DxyloNA (in that order) are flexible in comparison to corresponding C3′ inverted counterparts arabinoNA, RNA and DNA, respectively.

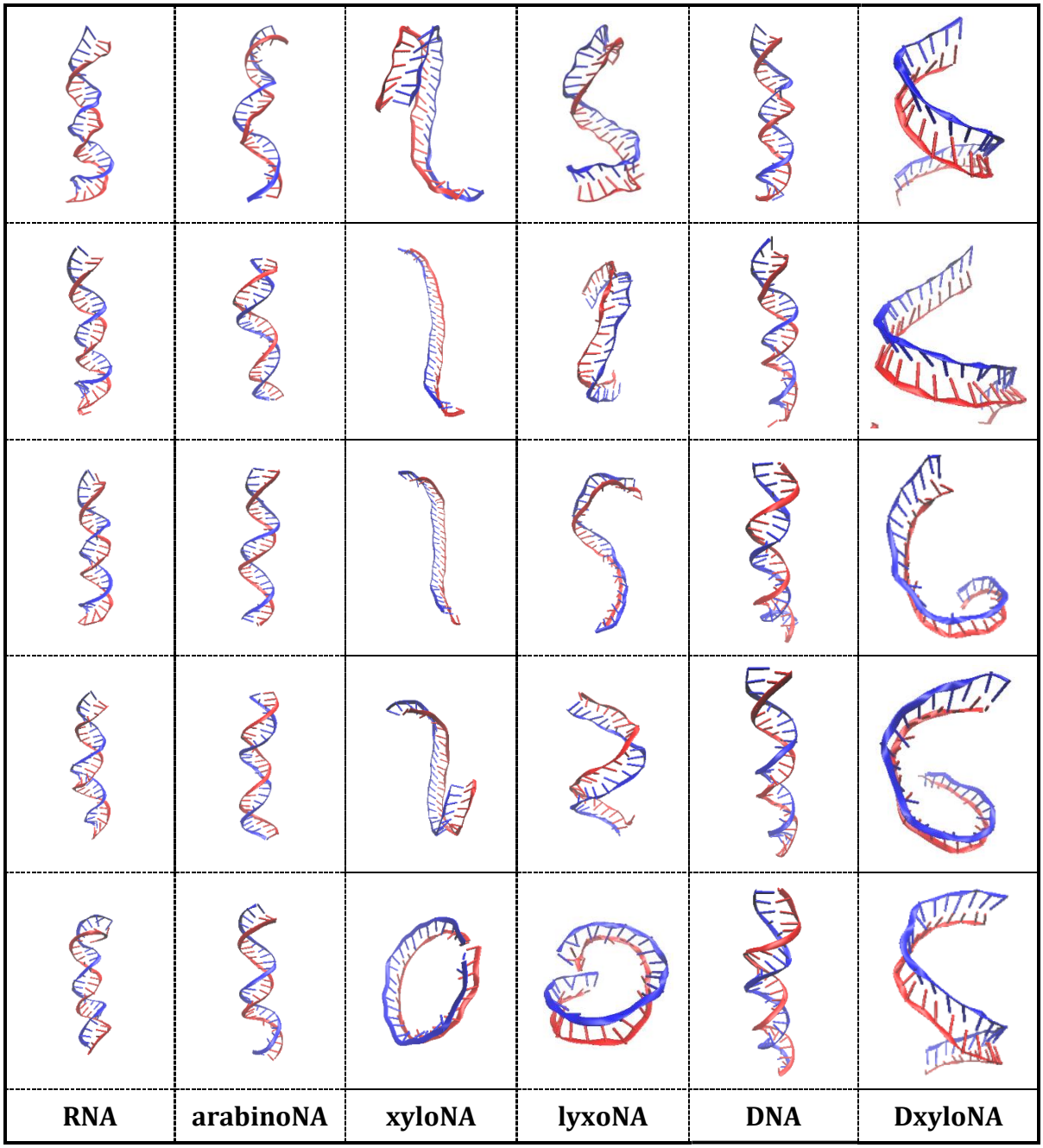

FIGURE 2 Collage of snapshots of structures of double-stranded nucleic acids obtained after five different trajectories of 100 ns each. This figure indicates that RNA, arabinoNA, and DNA maintain their structural integrity as a helix, while xyloNA, lyxoNA, and DxyloNA display a myriad of structures suggesting their flexible nature.

the xyloNA, lyxoNA, and DxyloNA vis-à-vis RNA, arabinoNA, and DNA respectively, results in loss of helicity. Fig. 1 also depicts a plot of end-to-end distance (the distance between the oxygen atoms of the furanose ring at two ends of the same strand) over the period of simulation. Interestingly, for the RNA, arabinoNA, and DNA the end-to-end distance is conserved (around 9 nm), while the corresponding plots for the xyloNA, lyxoNA, and Dxy-loNA show marked variation over the period of simulation.

In an effort to statistical validity these results, four additional (a total of five) independent trajectories of 100 ns for each of the six variants of double-stranded nucleic acids were carried out and Fig. 2 shows the collection of snapshots of structures obtained at the end of 100 ns of simulation time. Fig. 2 depicts that RNA, arabinoNA, and DNA maintain their structural integrity over each of the trajectory and converge to a right-handed double-helical structure. Analysis of the end-to-end distances (Table S1; see the Supporting Material) suggests that the DNA forms the most rigid double-helical structure followed by arabinoNA and RNA. On the other hand, xyloNA, lyxoNA, and DxyloNA are flexible and exhibit a diverse set of structures, which cannot be classified as right-handed helices.

Further, various average structural parameters such as intra-strand stacking distance, inter-strand N---N hydrogen bond distance, and inter-strand P---P distances were carried out, and Figure 3 depicts the observed trends. The average stacking distance, which is the separation of center-of-masses of two adjacent intra-strand nucleo-bases, is lowest for the DNA, followed by RNA, arabinoNA, lyxoNA, xyloNA and DxyloNA. The average intra-strand stacking distance is around 0.5 nm for RNA, arabinoNA, and DNA, while the corresponding distances are larger in the case of xyloNA, lyxoNA and DxyloNA, which is the result of the weakening of π-stacking interaction between adjacent nucleobases which leads to loss of helicity in the latter case (37). In the case of xyloNA the inversion of stereochemistry at the C3′ position relative to ribose leads to the formation of favorable Lp⋯π interactions between the oxygen atom of the xylose and the adjacent nucleobase due to the disposition of the sugar moiety relative to the nucleobases (21). Careful structural analysis suggests the formation of Lp⋯π interactions even in the case of lyxoNA and DxyloNA (Figure S1; see the SUPPORTING MATERIAL). Therefore, it can be inferred that stereochemistry at the C3′ position is a switch that toggles between π-stacking and Lp⋯π interactions. The inter-strand interactions in all the six double-stranded nucleic acids were analyzed by examining the average inter-strand distances between the phosphorus atoms and the average N–H⋯N hydrogen-bonded distances which are also shown in Figure 3. In comparison with dsRNA, the inter-strand distances between the phosphorus atoms are shorter for the lyxoNA and longer in arabinoNA and xyloNA in that order. However, the average N–H⋯N hydrogen-bonded distances show a considerably different trend, wherein the average N–H⋯N hydrogen-bonded distances for the RNA arabinoNA and xyloNA are almost identical, while the corresponding distance for the lyxoNA are much longer, indicating a weaker inter-strand hydrogen bonding in this case. Interestingly, hydrogen-bonded distances show opposite trends in xyloNA (DxyloNA) and lyxoNA, which suggest the buckling of the double-stranded structure in opposite directions. The helicity or lack of it in the six variants of double-stranded nucleic acids, analyzed using 3DNA (38, 39), suggests that RNA, arabinoNA, and DNA form right-handed helical structures, while lyxoNA has multiple left-handed helical regions and associated structural heterogeneity (Table S2; see the SUPPORTING MATERIAL). On the other hand, xyloNA and DxyloNA form left-handed ladder-like structures (W-form) in some instances. The most important consequence of the inversion of configuration at the C3′ position is the loss of helicity resulting in the loss of both the major and the minor grooves. The

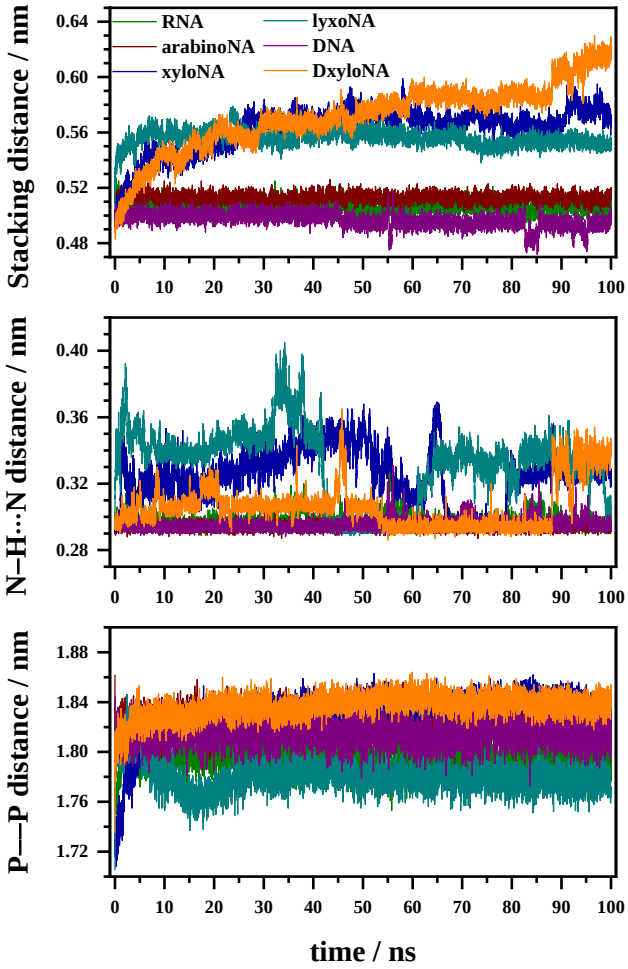

FIGURE 3 Collage of snapshots of structures of double-stranded nucleic acids obtained after five different trajectories of 100 ns each. This figure indicates that RNA, arabinoNA, and DNA maintain their structural integrity as a helix, while xyloNA, lyxoNA, and DxyloNA display a myriad of structures suggesting their flexible nature

helical parameters suggest that average turn-per-residue is largest for the DNA (31.5°) followed by RNA (30.7°) and arabinoNA (27.4°), with almost similar end-to-end distance (about 9 nm), which indicates untwisting of arabinoNA relative to DNA and RNA. This untwisting of arabinoNA double helix results in a marginal decrease in minor groove width and a substantial increase in the major groove width, wherein, the ratio of major-to-minor groove widths for the arabinoNA (2.00) is disproportionately large in comparison with RNA (1.15) and DNA (1.45).

The structural landscape of the furanosyl double-stranded nucleic acids is primarily governed by the configuration of the C3′ position. Inversion at C3′ position relative to ribose results in loss of helicity as in the case of xyloNA and lyxoNA. These structural changes can be associated with sugar pucker, which is influenced to a larger extent by the inversion of configuration at the C3′ position. On the other hand, inversion at C2′ position (arabinoNA) leads to ineffective helix in comparison to RNA, However, the loss of chirality at C2′ position in DNA appears to be a better proposition than inversion, which is

intriguing. The present set of simulations bring out that incorporation of ribose results in better defined helices among the four possible furanose variants with robust structural integrity, which might have led to the natural choice of ribose over the other three furanoses. Additionally, the aspect ratio of the major-to-minor groove widths could have played some role. The formation of the helix and its associated structural features are utilized by a variety of biological machinery such as helicases, topoisomerases, groove binders, and several others, which could not have been possible with xylose and lyxose derived nucleic acids and to a much lesser degree of control with arabinoNA. Returning to the question raised by Eschenmoser (5), "why did nature choose furanosyl-RNA and not pyranosyl-RNA?" is still a matter of debate, however, among the furanoses, ribose appears to be the best natural choice.

CONCLUSIONS

Summarizing, all-atom molecular dynamics simulations on six double-stranded nucleic acids viz., RNA, arabinoNA, xyloNA, lyxoNA, DNA, and DxyloNA suggest that the configuration of C3′ stereocenter determines the ability of double-stranded nucleic acids to form right-handed helices. The inversion of the C3′ stereocenter acts as a toggle switch for the helix to ladder structural transformation. On the other hand, the inversion of the C2′ stereocenter has a less pronounced effect, resulting in the modification of groove widths. A comparison of structural features suggests that ribose containing double-stranded nucleic acids are structurally robust and form better defined helices, which might have been the nature's choice

SUPPORTING MATERIAL

Supporting material can be found online at

AUTHOR CONTRIBUTIONS

The problem was formulated by G.N.P. Most of the calculations and the analysis was carried out by A.K. with some help from A.T. in running calculations on DNA and DxyloNA. Results were jointly interpreted by all the authors and the manuscript was written by A.K. and G.N.P.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the SpaceTime-2 supercomputing facility at IIT Bombay for the computing time. The authors wish to thank Dr. Amutha Ramaswamy for making the initial structure of double-stranded xyloNA available. The authors also wish to thank Prof. R. B. Sunoj for his comments and suggestions.

REFERENCES

1. Eschenmoser, A. 1999. Chemical etiology of nucleic acid structure. *Science (80-.).* 284:2118–2124.

2. Appella, D.H. 2009. Non-natural nucleic acids for synthetic biology. *Curr. Opin. Chem. Biol.*

3. Anosova, I., E.A. Kowal, M.R. Dunn, J.C. Chaput, W.D. Van Horn, and M. Egli. 2016. The structural diversity of artificial genetic polymers. *Nucleic Acids Res.* 44:1007–1021.

4. Chaput, J.C., and P. Herdewijn. 2019. What Is XNA? *Angew. Chemie - Int. Ed.* 58:11570–11572.

5. Bolli, M., R. Micura, and A. Eschenmoser. 1997. Pyranosyi-RNA: Chiroselective self-assembly of base sequences by ligative oligomerization of tetranucleotide-2',3'-cyclophosphates (with a commentary concerning the origin of biomolecular homochirality). *Chem. Biol.* 4:309–320.

6. Robeyns, K., P. Herdewijn, and L. Van Meervelt. 2008. Structure of the fully modified left-handed cyclohexene nucleic acid sequence GTGTACAC. *J. Am. Chem. Soc.* 130:1979–1984.

7. Zhang, L., A. Peritz, and E. Meggers. 2005. A simple glycol nucleic acid. *J. Am. Chem. Soc.* 127:4174–4175.

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

9. Nielsen, P.E. 1999. Peptide nucleic acid. A molecule with two identities. *Acc. Chem. Res.* 32:624–630.

10. Herdewijn, P. 2001. TNA as a potential alternative to natural nucleic acids. *Angew. Chemie - Int. Ed.* 40:2249–2251.

11. Morihiro, K., Y. Kasahara, and S. Obika. 2017. Biological applications of xeno nucleic acids. *Mol. Biosyst.* 13:235–245.

12. Joyce, G.F. 2012. Toward an alternative biology. *Science (80-.).* 336:307–308.

13. Martín-Pintado, N., M. Yahyaee-Anzahaee, R. Campos-Olivas, A.M. Noronha, C.J. Wilds, M.J. Damha, and C. González. 2012. The solution structure of double helical arabino nucleic acids (ANA and 2′F-ANA): Effect of arabinoses in duplex-hairpin interconversion. *Nucleic Acids Res.* 40:9329–9339.

14. Liu, B., and T.A. Steitz. 2017. Structural insights into NusG regulating transcription elongation. *Nucleic Acids Res.* 45:968–974.

15. Černý, J., M. Kabeláč, and P. Hobza. 2008. Double-helical → ladder structural transition in the B-DNA is induced by a loss of dispersion energy. *J. Am. Chem. Soc.*

130:16055–16059.

16. Konrad, M.W., and J.I. Bolonick. 1996. Molecular dynamics simulation of DNA stretching is consistent with the tension observed for extension and strand separation and predicts a novel ladder structure. *J. Am. Chem. Soc.* 118:10989–10994.

17. Maiti, M., M. Maiti, C. Knies, S. Dumbre, E. Lescrinier, H. Rosemeyer, A. Ceulemans, and P. Herdewijn. 2015. Xylonucleic acid: Synthesis, structure, and orthogonal pairing properties. *Nucleic Acids Res.* 43:7189–7200.

18. Maiti, M., V. Siegmund, M. Abramov, E. Lescrinier, H. Rosemeyer, M. Froeyen, A. Ramaswamy, A. Ceulemans, A. Marx, and P. Herdewijn. 2012. Solution structure and conformational dynamics of deoxyxylonucleic acids (dXNA): An orthogonal nucleic acid candidate. *Chem. - A Eur. J.* 18:869–879.

19. Ramaswamy, A., M. Froeyen, P. Herdewijn, and A. Ceulemans. 2010. Helical structure of xylose-DNA. *J. Am. Chem. Soc.* 132:587–595.

20. Ramaswamy, A., D. Smyrnova, M. Froeyen, M. Maiti, P. Herdewijn, and A. Ceulemans. 2017. Molecular Dynamics of Double Stranded Xylo-Nucleic Acid. *J. Chem. Theory Comput.* 13:5028–5038.

21. Kumar, A., and G.N. Patwari. 2019. Probing the role of dispersion energy on structural transformation of double-stranded xylo- and ribo-nucleic acids. *Phys. Chem. Chem. Phys.* 21:3842–3848.

22. Sharpe, D.J., K. Röder, and D.J. Wales. Energy landscapes and dynamics of xylo-nucleic acids arXiv : 1906 . 00941v1 [physics . bio-ph] 3 Jun 2019. 1–27.

23. Schöppe, A., H.J. Hinz, H. Rosemeyer, and F. Seela. 1996. Xylose-DNA: Comparison of the thermodynamic stability of oligo(2'-deoxyxylonucleotide) and oligo(2'-deoxyribonucleotide) duplexes. *Eur. J. Biochem.* 239:33–41.

24. Ghosh, S., and R. Chakrabarti. 2016. Spontaneous Unzipping of Xylonucleic Acid Assisted by a Single-Walled Carbon Nanotube: A Computational Study. *J. Phys. Chem. B*. 120:3642–3652.

25. Danielsen, M.B., C. Lou, J. Lisowiec-Wachnicka, A. Pasternak, P.T. Jørgensen, and J. Wengel. 2020. Gapmer Antisense Oligonucleotides Containing 2′,3′-Dideoxy-2′-fluoro-3′-C-hydroxymethyl-β-d-lyxofuranosyl Nucleotides Display Site-Specific RNase H Cleavage and Induce Gene Silencing. *Chem. - A Eur. J.* 26:1368–1379.

26. Maier, J.A., C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, and C. Simmerling. 2015. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* 11:3696–3713.

27. Abraham, M.J., T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, and E. Lindah. 2015. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 1–2:19–25.

28. Macke, T.J., and D.A. Case. 1997. Modeling Unusual Nucleic Acid Structures. In: ACS Symposium Series. . pp. 379–393.

29. Abraham, M.J., T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, and E. Lindah. 2015. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 1–2:19–25.

30. Van Der Spoel, D., E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, and H.J.C. Berendsen. 2005. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* 26:1701–1718.

31. Mark, P., and L. Nilsson. 2001. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A*. 105:9954–9960.

32. Petrova, S.S., and A.D. Solov'ev. 1997. The Origin of the Method of Steepest Descent. *Hist. Math.* 24:361–375.

33. Bussi, G., D. Donadio, and M. Parrinello. 2007. Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126.

34. Berendsen, H.J.C.C., J.P.M.M. Postma, W.F. Van Gunsteren, A. DiNola, and J.R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.

35. Parrinello, M., and A. Rahman. 1981. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* 52:7182–7190.

36. Feller, S.E., Y. Zhang, R.W. Pastor, and B.R. Brooks. 1995. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* 103:4613–4621.

37. Cooper, V.R., T. Thonhauser, A. Puzder, E. Schröder, B.I. Lundqvist, and D.C. Langreth. 2008. Stacking interactions and the Twist of DNA. *J. Am. Chem. Soc.* 130:1304–1308.

38. Lu, X.J., and W.K. Olson. 2008. 3DNA: A versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protoc.* 3:1213–1227.

39. Lu, X.J., and W.K. Olson. 2003. 3DNA: A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* 31:5108–5121.