

CHEMISTRY

A **European** Journal



Accepted Article

Title: Chimeric XNA - An Unconventional Design for Orthogonal Informational Systems

Authors: Tim Efthymiou, Jesse Gavette, Matthias Stoop, Francesco De Riccardis, Mathy Froeyen, Piet Herdewyn, and Ramanarayanan Krishnamurthy

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

To be cited as: *Chem. Eur. J.* 10.1002/chem.201802287

Link to VoR: <http://dx.doi.org/10.1002/chem.201802287>

Supported by
ACES

WILEY-VCH

Chimeric XNA – An Unconventional Design for Orthogonal Informational Systems

Tim Efthymiou^[a,b], Jesse Gavette^[a,b], Matthias Stoop^[a,b], Francesco De Riccardis^[a,b,c], Mathy Froeyen^[d], Piet Herdewijn^[d], Ramanarayanan Krishnamurthy*^[a,b]

Dedicated to Professor Albert Eschenmoser for his seminal work on Chemical Etiology of Nucleic Acids.

Abstract: The paradigm of homogenous-sugar-backbone of RNA and DNA has reliably guided the construction of many functional and useful Xeno Nucleic Acid (XNA) systems to date. Deviations from this monotonous and canonical design, in many cases, results in oligonucleotide systems which lack base-pairing with themselves, or with RNA or DNA. Here we show that nucleotides of two such compromised XNA systems can be combined with RNA and DNA in specific patterns to produce chimeric-backbone oligonucleotides, which in certain cases demonstrate base-pairing properties comparable to -or stronger than- canonical systems, while also altering the conventional Watson-Crick pairing behavior. The unorthodox pairing properties generated from these chimeric-sugar-backbone oligonucleotides suggest a counterintuitive approach of creating modules consisting of non-base pairing XNAs with RNA/DNA in a set pattern. This strategy has the potential to increase the diversity of unconventional nucleic acids leading to orthogonal backbone-sequence-controlled informational systems.

Introduction

RNA and DNA, which are the hub of many biological processes, are functional-informational systems with homogeneous sugar-phosphate-backbones and canonical nucleobases (Fig. 1).¹ This structural motif has been consistently emulated in the design of XNAs for biomedical applications and origins of life research.²⁻¹⁶ Needless to say, the single most important criterion in judging the utility of XNAs is their base-pairing capability, with themselves and with RNA or DNA. Therefore, it is not surprising that XNAs which lack (or have weak) base-pairing properties are rarely investigated further.²⁻¹⁵ In our recent work, within the context of oligonucleotides derived from the structural

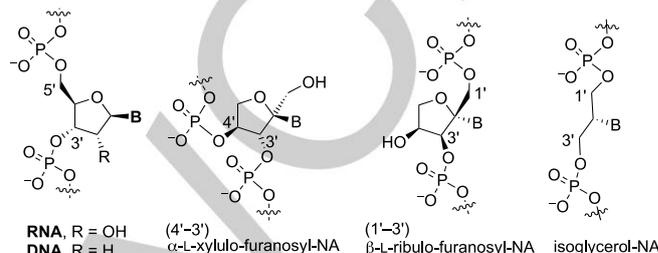


Figure 1. The constitutional and configurational representations of RNA, DNA and the XNAs investigated in this study.

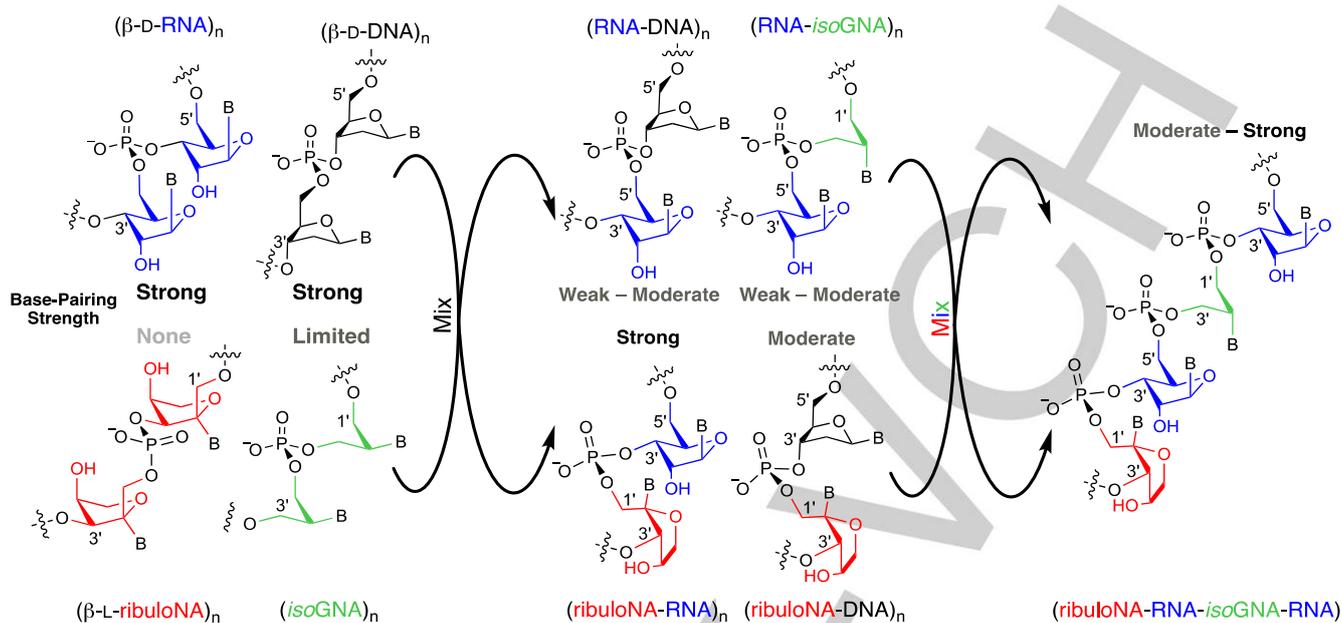
neighborhood of RNA,^{6,16} we investigated the base-pairing properties of pentulose-derived oligonucleotides (Fig. 1): β -L-ribuloNA (n , an isomer of RNA with L-ribulose in place of D-ribose) and α -L-xyluloNA (x , an isomer of RNA with L-xylulose in place of D-ribose), and found them to be devoid of base-pairing capacity.¹⁷ We observed that one or two insertions of xyluloNA mononucleotides into RNA dramatically weakened or destroyed duplex formation capabilities of the modified RNA sequences.¹⁷ A similar behavior was observed when inserting ribuloNA units into RNA for this study (Supporting Table S6, entries 12 – 19), reinforcing the view that pentulose-nucleotides compromise the base-pairing of RNA. However, when we designed oligonucleotide sequences containing strictly alternating pentuloseNA- and RNA-nucleotide units (which were synthesized in the context of understanding the difficulties and the low yields encountered when synthesizing full-pentuloseNA oligomers¹⁷), and investigated their base-pairing potential, we observed surprising base-pairing behavior that led us to reevaluate our previous assessment of the pentuloseNA-RNA chimeric systems. The results described below demonstrate that the strategic incorporation of XNAs (with no or weak hybridization properties with themselves and with RNA/DNA) into RNA (rN) and DNA (dN) in a specific pattern (Scheme 1) can create chimeric-XNA-systems with enhanced orthogonal base-pairing properties. This approach further exemplifies the concept of an “aperiodic crystal” proposed by Schroedinger.¹⁸

Results and Discussion

We began the investigation with the alternating (4',3')-xyluloNA(x n)-RNA(N) chimeric self-complementary sequence $r(^x tA)_8$ forming a duplex **1**, which showed remarkably, duplex stability slightly greater than that of the corresponding parent RNA-based rT duplex **6** (Table 1, Supporting Fig. S29a). The same behavior was observed, when DNA(A) was substituted for RNA in duplex **1**, leading to the alternating chimeric self-

- [a] Dr. T. Efthymiou, Dr. J. Gavette, Dr. M. Stoop, Professor Dr. R. Krishnamurthy
Department of Chemistry
The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037, USA
E-mail: krishna@scripps.edu
- [b] NSF/NASA Center for Chemical Evolution, Atlanta, GA 30332, USA
- [c] Professor Dr. F. De Riccardis
Department of Chemistry and Biology
University of Salerno
Via Giovanni Paolo II, 132, 84084, Fisciano, Salerno, Italy
- [d] Dr. M. Froeyen, Professor Dr. P. Herdewijn
Department of Medicinal Chemistry
Institute for Medical Research, KU Leuven, Herestraat, 49, 3000, Belgium

Supporting information for this article is given via a link at the end of the document.



Scheme 1. Gain-of-function (enhanced base-pairing) hybrid-sugar backbone chimeric oligonucleotides. Constitutional and conformational representations of self- and non-self-complementary homogeneous-backbone and chimeric oligonucleotides considered in this study, with their respective base-pairing profiles. Black = DNA (d, N); Blue = RNA (r, N); Red = 1',3'-ribuloNA (n); Green = 1',3'-isoGNA (n).

complementary sequence $d(^x tA)_8$, forming stable duplex **3** (Table 1, Supporting Fig. S29b). This suggested that the (4',3')-xyluloNA units when placed in a specific arrangement with RNA in a chimeric strand, were pairing with a complementary RNA(A) or DNA(A) unit leading to chimeric-duplex formation, even though the individual homogeneous sequences of xyluloNA($x t$) does not pair with RNA(A) or DNA (A)¹⁷. Interestingly, the reverse xylulo-purine and ribo-pyrimidine combination in oligomer $r(^x aT)_8$ (**2**) and other various alternating non-self-complementary xyluloNA($x n$)-RNA(N) chimeric sequences formed no self- or cross-pairing stable duplexes (Supporting Table S5). Stimulated by these observations we switched to the ribuloNA-RNA system and investigated a series of self-complementary chimeric sequences composed of alternating (t)-r(A) units, $r(^t A)_n=4-8$, which again formed duplexes, but were -unexpectedly- far more thermostable (T_m , 45.3 – 85.9 °C) than the parent RNA-duplexes (duplex **4** vs **6**, Fig. 2A, Table 1). Furthermore, unlike the xyluloNA-RNA $r(^x aT)_8$, the inverse ribulo(a)-ribo(T), $r(^t A)_8$ did form a stable duplex **5** and was only slightly less thermostable (ΔT_m –3.6 °C) than its parent RNA $r(AT)_8$ duplex **7**, though far less stable than chimeric duplex **4** (ΔT_m – 36.4 °C) (Table 1, Fig. 2A). This unusual behavior of the ribuloNA-RNA system vis-à-vis the xyluloNA-RNA system, led us to explore the scope of the alternating (ribuloNA-RNA) $_n$ backbone design as scaffold for an informational system by synthesizing non-self-complementary sequences (Supporting Table S6). The anti-parallel duplex **12** exhibited a slightly weaker thermostability compared to its corresponding parent RNA duplex **13** (ΔT_m –2.8 °C, Table 1, Fig. 2A), while an identical chimeric duplex designed with a forced parallel orientation was

unable to form duplexes (Table 1, Supporting Table S6, entry 24). This strict preference for an anti-parallel alignment seems to suggest an ordered conformation based on backbone-compatibility as a consequence of the (ribuloNA-RNA) $_n$ design which places a ribuloNA unit opposite to a complementary RNA residue. This result is to be considered in light of the highly-destabilizing ribuloNA/ribuloNA interactions,¹⁸ where even a single ribuloNA/ribuloNA pair insertion in an RNA duplex is detrimental to its thermostability (Supporting Table S6, entry 18). Duplex formation was confirmed in select cases by Job-plots, and by observing the thermal stability of duplexes by varying sequence length and oligonucleotide concentration (Supporting Figs. S28, S30 and S42). In addition, a pronounced dependence of duplex thermostability on ionic strength (150 mM to 1 M NaCl) was observed with chimera-duplex **12** (ΔT_m +19.1 °C, Table 1, Supporting Fig. S34), when compared with canonical RNA and DNA duplex counterparts (ΔT_m +7.3 – 10.8 °C; Table 1). This pronounced “salt effect” is likely attributed to the shorter inter-strand distances between phosphates in the chimeric duplexes due to the location of both the nucleobase and the phosphate on the same side of the ribulo-sugar unit¹⁷ (Fig. 1).

The greater duplex stability of $r(^t A)_8$ **4** versus $r(^x aT)_8$ **5** suggested a trend whereby the ribulo(Pyrimidine)-ribo(Purine) combination was stronger than the reverse ribulo(Purine)-ribo(Pyrimidine) combination. One possible explanation, based on qualitative model building, suggests a pronounced slide of the ribo-(purines) towards the center of the duplex with recessed, complementary ribulo-(pyrimidines), resulting in an inter-strand base-stacking pattern in duplex **4** – a pattern not manifested in

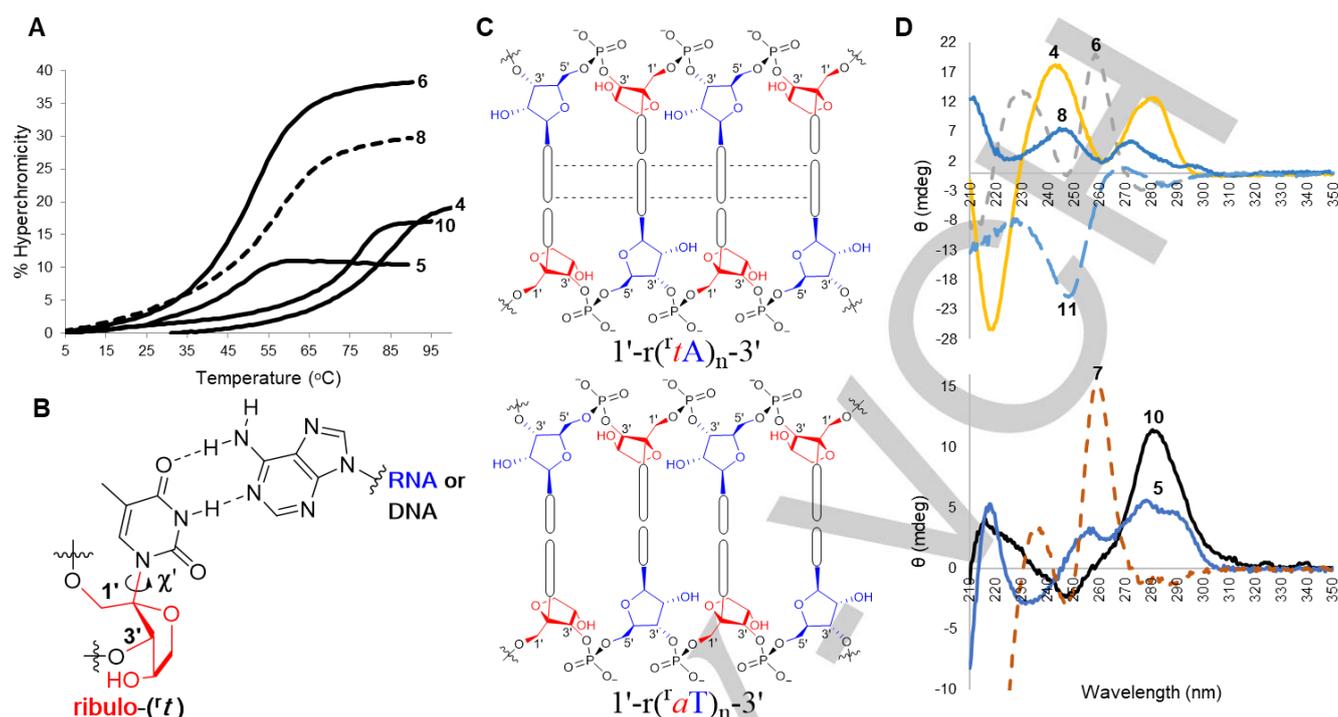


Figure 2. Base-pairing behavior of self-complementary strictly-alternating (ribuloNA-RNA)_n sequences. All panels refer to sequences in Table 1. A) UV-T_m curves of synthetic (ribuloNA-RNA)_n and RNA (rT) sequences. Buffer: 10 mM Na₂HPO₄, 100 μM EDTA, pH 7.0; dashed line = 1 M NaCl; solid line = 150 mM NaCl; Sample [strand] = 4 μM. B) The proposed ribulo-syn-nucleobase model in a Watson-Crick interaction, illustrating rotation of the χ' torsion angle. C) Proposed inter-strand purine-purine stacking interaction (dashed lines) in r(^tA)_n series (top) versus its absence in r(^tA)_n series (bottom). Red = ribuloNA ('n'); blue = RNA (r, N). D) Circular Dichroism (CD) curves of (ribulo(Pyrimidine)-RNA)_n sequences (top) and (ribulo(Purine)-RNA)_n sequences (bottom) at 0 °C in 150 mM NaCl buffer with 4 μM [strand] (11 = 1 M NaCl, 6 μM [strand]); dashed line represents canonical sequence.

duplex **5** due to the reverse purine-on-ribulo/pyrimidine-on-ribo design (Fig. 2C). Based on these models we applied the (ribuloNA-RNA)_n pattern to self-complementary guanine-cytosine duplexes; however, we observed the exact opposite behavior: the length-sensitive r(^tcG)_n series was less thermostable than r(^tgC)_n (ΔT_m -17.4 °C, Table 1, oligonucleotide duplexes **8** vs **9**). Furthermore, chimeric-duplexes r(^tcG)₆ **8** (T_m 46.5 °C) and r(^tgC)₆ **9** (T_m 63.9 °C) were not only weaker than the r(^tA)_n sequences of corresponding length, but far weaker than their parent RNA sequences (Supporting Table S6). To our knowledge, this is the first example where a guanine-cytosine Watson-Crick base-pair is weaker than an adenine-thymine base-pair within the same backbone framework in a strictly antiparallel duplex,¹⁹ and challenges the traditional three- versus two-hydrogen bond model as the primary reason¹ for stronger GC versus AT base-pairs.²⁰ In addition, modelling suggests a syn conformation of the ribulo-nucleobase (purine and pyrimidine). The presumed preference for syn-(T and A) over syn-(G and C) on the ribulo-unit, offers an alternative-design option to the known preference of syn-G (over A) and syn-purine (over pyrimidine) paradigms,²¹ and is atypical of ribo-nucleosides.¹ The syn-preference in ribulo-nucleosides can be understood based on the fact that the ribulose-C(5')-position

(which is equivalent to the C(4')-position of ribose) has no steric hindrance to offer to the syn-oriented nucleobase (Fig. 2B).

Computational modeling using Amber 16/AmberTools17 software²² was attempted to gain insight into the peculiar behavior of these alternating (ribuloNA-RNA)_n oligonucleotides. Details of modeling and simulation procedures are given in the Supporting information. Models of r(^tcG)₆ and r(^tgC)₆ self-complementary duplexes were created starting from a classical antiparallel A-RNA duplex with Watson-Crick base pairing. Figure 3 shows representative conformations of the r(^tcG)₆ and r(^tgC)₆ duplexes, as well as their superposition. Examination of structures in UCSF Chimera²³ revealed that the r(^tcG)₆ structure suffered from base pair opening at the end residues. So, for r(^tcG)₆ duplex, the fraying ends were omitted and only the remaining internal residues were considered in the duplex analysis. In Table 2 we give a summary of some structural parameters of both duplexes r(^tcG)₆ and r(^tgC)₆ compared with a reference pdb structures of DNA (1bna), RNA (1qcu) and Z-DNA (4ocb). The most striking difference, when compared with RNA and DNA duplexes, is that r(^tcG)₆ and r(^tgC)₆ have a repeating unit of 2 base pairs, as can be seen from the different base orientation (anti for RNA residues, syn for ribulo residues).

Table 1. Thermal duplex stability (UV- T_m values) of select chimeric xyluloNA-, ribuloNA- containing oligonucleotides, and comparison with RNA, DNA and (RNA-DNA) $_n$.^[a]

Duplex	Sequence	T_m (°C)
Self-Complementary <i>xyluloNA</i> -RNA and <i>xyluloNA</i> -DNA Sequences		
1	4'-r(^x tA) ₈ -3'	54.4 ^[b]
2	4'-r(^x aT) ₈ -3'	— ^[b]
3	4'-d(^x tA) ₈ -3'	49.4 ^[b]
Self-Complementary <i>ribuloNA</i> -RNA Sequences		
4	1'-r(^t tA) ₈ -3'	85.9 ^[b]
5	1'-r(^t aT) ₈ -3'	49.5 ^[b]
6	5'-r(TA) ₈ -3'	51.1 ^[b]
7	5'-r(AT) ₈ -3'	53.1 ^[b]
8	1'-r(^t cG) ₆ -3'	46.5 ^[b] , 56.9 ^[c]
9	1'-r(^t gC) ₆ -3'	63.9 ^[b]
10	1'-r(^t gC) ₈ -3'	77.3 ^[b]
11	5'-d(TA) ₈ -3'	42.2 ^[b]
Non-Self-Complementary <i>ribuloNA</i> -RNA(DNA) Duplexes		
12	1'-r(^t aA ^t aA ^t tA ^t tA ^t aT ^t tA)-3'/ 3'-r(T ^t T ^t tA ^t aA ^t tA ^t aT ^t aA ^t t)-1'	35.2 ^[b] , 54.3 ^[c]
13	5'-r(AAAATTTATATTATTA)-3'/ 3'-PO ₄ -r(TTTTAAATATAATAAT)-5'	48.9 ^[b] , 56.2 ^[c] (5 μM)
14	1'-r(^t gA ^t gA ^t tC ^t tA ^t tC ^t gC ^t tA)-3'/ 3'-r(C ^t tC ^t tA ^t gA ^t tA ^t gC ^t gA ^t t)-1'	67.5 ^[c]
15	5'-r(GAGAUCUAUAUCGCUA)-3'/ 3'-r(CUCUAGAUUAGCGAU)-5'	72.3 ^[c]
16	1'-d(^t aA ^t aA ^t tA ^t tA ^t aT ^t tA)-3'/ 3'-d(T ^t T ^t tA ^t aA ^t tA ^t aT ^t aA ^t t)-1'	34.5 ^[c]
17	5'-(A ^d A ^d A ^d A ^d U ^d T ^d U ^d A ^d U ^d A ^d U ^d T ^d A ^d T ^d U ^d A)-3'/ 3'-(^d T ^d U ^d T ^d U ^d A ^d A ^d U ^d A ^d U ^d A ^d A ^d T ^d A ^d A ^d U)-5'	24.2 ^[c]
18	1'-d(^t aA ^t aA ^t tA ^t tA ^t aT ^t tA)-3'/ 3'-r(T ^t T ^t tA ^t aA ^t tA ^t aT ^t aA ^t t)-1'	42.0 ^[c]
19	5'-r(AAAUUUUUUAUUUUUA)-3'/ 3'-r(UUUUUAAAUAUUUUUA)-5'	47.1 ^[c]
20	5'-d(AAAATTTATATTATTA)-3'/ 3'-d(TTTTAAATATAATAAT)-5'	37.0 ^a , 47.8 ^[c]

[a] [Duplex] = 2 μM (4 μM C_{tot}) unless otherwise specified. r or ^tN = RNA; d or ^dN = DNA; ^rn = ribuloNA. [b]: Phosphate buffer: 10 mM Na₂HPO₄, 100 μM EDTA, 150 mM NaCl, pH 7.0; [c]: 10 mM Na₂HPO₄, 100 μM EDTA, 1 M NaCl, pH 7.0. 4' in xyluloNA and 1' in ribuloNA are equivalent to RNA 5' and DNA 5' end.

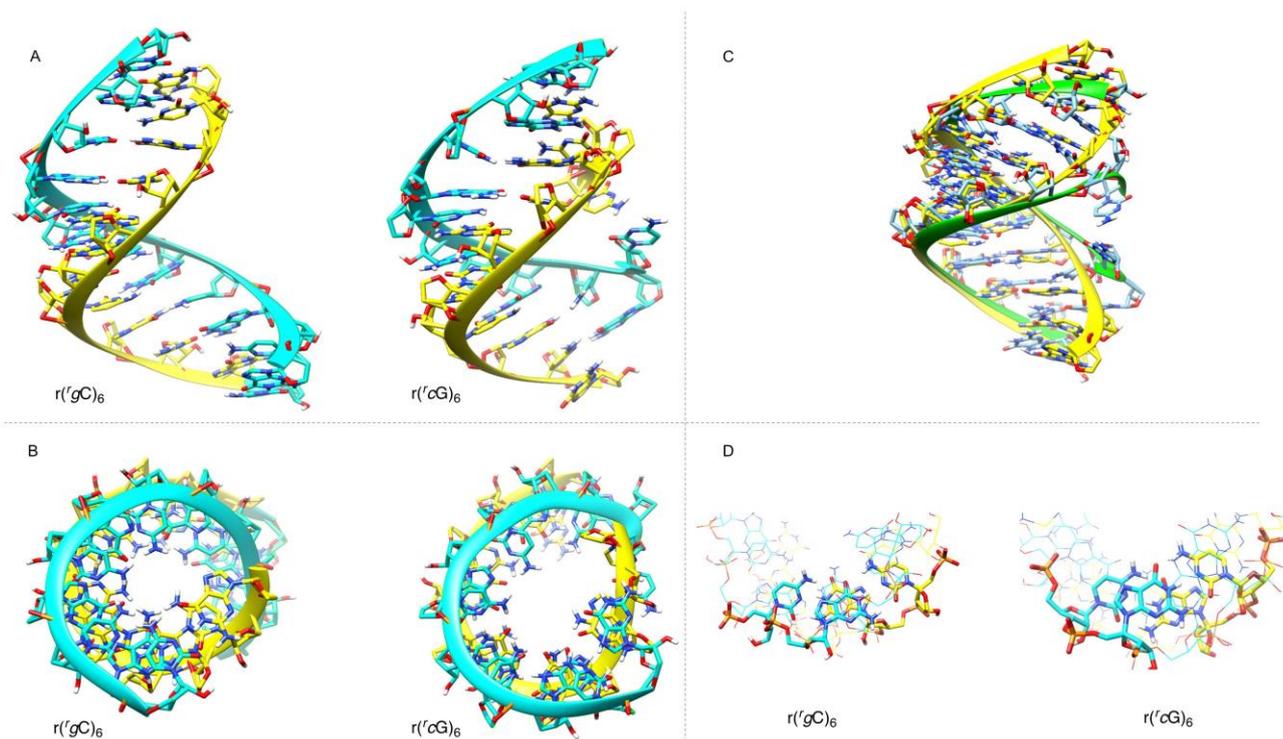


Figure 3. Molecular model of the self-complementary double helical structure of β -L-ribulo-furanosylINA-RNA chimeric nucleic acids with 'gC and 'cG 'dinucleotide repeats: left for $r('gC)_6$ and right for $r('cG)_6$. In both cases side-view (A) and top-view (B) of the helices are shown. The $r('cG)_6$ simulated structure suffers from base pair fraying at the ends. C) A superposition of both structures is shown. D) A close-up of the base stacking pattern in the $r('gC)_6$ duplex (left) and in the $r('cG)_6$ duplex (right) is represented at the bottom right side. Images were created using UCSF Chimera.²³

This difference is also reflected in some of the backbone dihedral angles (Supporting Table S25). The curves+ analysis²⁴ output (Supporting Tables S15 and S16) also shows that the simulated $r('gC)_6$ duplex is more regular than the $r('cG)_6$ duplex which is in agreement with the RMSF analysis (Fig. S65). The difference between $r('gC)_6$ and $r('cG)_6$ is that the structural parameters of $r('gC)_6$ more resembles the parameters of RNA, while $r('cG)_6$ has a lower twist value and a much higher inclination. It should be mentioned that $r('gC)_6$ gives a more stable MD trajectory than $r('cG)_6$, deduced from the RMSF plots. Analysis of the G:C Watson-crick hydrogen bonds shows that the hydrogen bonds in $r('gC)_6$ and $r('cG)_6$ are somewhat longer than in the DNA and RNA reference structures (Supporting Table S26). In contrast, internucleotide stacking (as also seen from the inclination) is higher in $r('cG)_6$ (with a periodicity of two bases) than in $r('gC)_6$ and RNA (Supporting Tables S23, S24 and Fig. 3). This periodicity suggests that the alternating chimeric system could start behaving as a "dimeric informational unit", that is, the dinucleotide repeat is representing a unit of information – and could now encode information that is both dependent on the nucleobase and the backbone unit.⁸ The "dimeric unit" containing a syn-ribuloNA–anti-riboNA seems to simulate the type of CpG or GpC dinucleotide unit that is found in a Z-DNA, which –unlike the chimeric $r('gC)_6$ and $r('cG)_6$ – has poor internucleotide stacking compared to B-DNA.²⁵ Currently we are unable to model $r('tA)_8$ and $r('aT)_8$ using methods that

were successful for describing the dichotomous behavior of chimeras $r('cG)_6$ and $r('gC)_6$. This has prevented a greater understanding –both at the molecular level and supramolecular level– as to why $r('tA)_8$ shows exceptional duplex stability. Though circular dichroism (CD) comparisons demonstrate that all self-complementary chimeric-systems appear with a red-shifted λ_{max} , simulating A-like conformations with matching ('tA, 'cG) or similar ('aT, 'gC) profiles (Fig. 2D), a more detailed structural elucidation (by NMR and/or X-ray) is required to understand the root cause(s) of this interesting behavior.

Based on these results, we naturally progressed to investigate the base-pairing behavior of the ribuloNA-riboNA sequence containing all four 'tA'gC nucleosides. The non-self-complementary fully-hybrid ('tA'gC)-containing sequences formed a thermostable duplex **14** with a T_m of 67.5 °C (Table 1, Fig. 4A), stronger than $r('aA'tT)$ chimera duplex **12**. Unlike the self-complementary (ribuloNA-RNA)_n systems, the replacement of 'a-rT and 't-rA base-pairs in the non-self-complementary chimera **12** with 'g-rC pairs is stabilizing and follows the canonical pairing trends. This contrast in pairing profiles suggests a sequence dependence,¹⁹ based on the interaction dictated by which particular base (purine vs pyrimidine) is present on the ribulose – a behavior similar to other XNA systems like TNA.²⁶

Table 2. Comparison of some structural parameters of the duplexes derived from the curves plus calculations. Values are averages over all nucleotides. Numbers in brackets are standard deviations for averages on nucleotides in both strands.

	r'(cG)6	r'(gC)6	DNA	RNA	Z-DNA
	r'c=ribulo/G=RNA	r'g=ribulo/C=RNA	1bna	1qcu	4ocb
	12mer	12mer	12mer	11mer	12mer
	(cGcGcGcGcGcG) ₂	(gCgCgCgCgCgC) ₂	(CGCGAATTCGCG) (CGCGAATTCGCG)	(GGGGGGGGGGG) (CCCCCCCCC)	(CGCGCGCGCGCG) ₂
Base pairing	Watson-Crick	Watson-Crick	Watson-Crick	Watson-Crick	Watson-Crick
Helix sense	Right	Right	Right	Right	Left
Repeating unit	2 bp (cG)	2 bp (gC)	1 bp	1 bp	2 bp
Inclination °	29(5)	18(5)	0(5)	18(2)	8(1)
Twist °	cG 40(4) Gc 14(9)	gC 38(4) Cg 23(6)	36(4)	32(1)	CG -10(1) GC -49(2)
Glycosyl angle °	c -62(10) syn G -172(8) anti	g -65(11) syn C -165(9) anti	-117(14) anti	-162(3) anti	C -153(4) anti G 63(8) syn
Sugar pucker	C3'endo	C3'endo	C2'endo	C3'endo	C C2'endo G C3'endo

Encouraged by the potential of this sequence-design platform, we translated the strictly-alternating pattern to ribuloNA-DNA (Scheme 1). Synthetic d('tA) non-self-complementary sequences formed duplex **16**, which was weaker than the corresponding ribuloNA-RNA counterpart **12**, and its full-DNA parent duplex **20** ($\Delta T_m -13.3$ °C). However, duplex **16** was more thermostable ($\Delta T_m +10.3$ °C) when compared to the corresponding chimeric-RNA-DNA duplex **17** with alternating canonical RNA inserts within DNA (rAdArUdT) (Table 1, Fig. 4A). Such contrasting behavior of chimeric RNA-DNA duplexes versus ribuloNA-DNA or ribuloNA-RNA duplexes raises the possibility of combining non-base-pairing (instead of base-pairing) nucleotides in constructing XNA chimeric systems. Thermostability is also gained ($\Delta T_m +7.5$ °C) when a ribuloNA-DNA sequence is hybridized to its corresponding complimentary ribuloNA-RNA complement (duplex **18**, Fig. 4A). These results imply that ribuloNA residues favor accommodation within, and hybridization to RNA over DNA. The cross-pairing exhibited by chimeric duplex **18** is noteworthy for its hybrid sugar-backbone composition (ribulo-, ribo- and deoxyribo-), especially in light of the fact that none of these (ribuloNA-RNA)_n or (ribuloNA-DNA)_n sequences cross-pair with complementary homogeneous RNA or DNA sequences (Supporting Table S7), a behavior characteristic of orthogonal base-pairing systems.⁶

The CD profiles of the pentuloNA-based non-self-complementary chimeric duplexes were equally unpredictable. For example, the xyluloNA-RNA r('tA)₈ and xyluloNA-DNA d('tA)₈ exhibited a left-handed ("Z-type")²⁵ CD-curve at 150mM NaCl phosphate buffer (Supporting Fig. S46a). While both ribuloNA-RNA duplexes **12** and **14** demonstrated A-form RNA-

like helices (Fig. 4B), the corresponding ribuloNA-DNA duplex **16** and its single-strands, displayed a left-handed ("Z-type")²⁷ CD-curve at room temperature in 1M NaCl phosphate buffer (Fig. 4C; Supporting Fig. S54). The conformational behavior of this chimeric (AT)-only duplex, while atypical (compared with the RNA or DNA transition to Z-form usually observed in (GC)-rich sequences under high salt concentration, reduced hydration and/or increased temperatures),²⁷⁻²⁹ further seems to support the proposed ribulo-syn-NA model. Highlighting the unconventional nature of this chimeric system, increased salinity (5 M NaClO₄) completely melted ribuloNA-DNA duplex **16** into single-strands adopting, what appears to be, an A-form conformation at room temperature (Fig. 4C; Supporting Fig. S55). The comparatively milder room temperature inducement of this unique Z-to-A transition is likely attributed to the duplex's 50% DNA constitution³⁰ and weaker thermostability, while the ribuloNA residues dictate the resultant conformation. This type of divergent behavior can only be a consequence of the chimeric composition of the sugar-backbone, identifying a potential paradigm for designing sequence-controlled systems³⁰ with tunable, context-sensitive properties.^{4,7,15} Furthermore, (ribuloNA-RNA)_n chimeras demonstrate stability to enzymatic digestion (Fig. 4D) – a trait favored for potential biomedical applications.^{4,31,32}

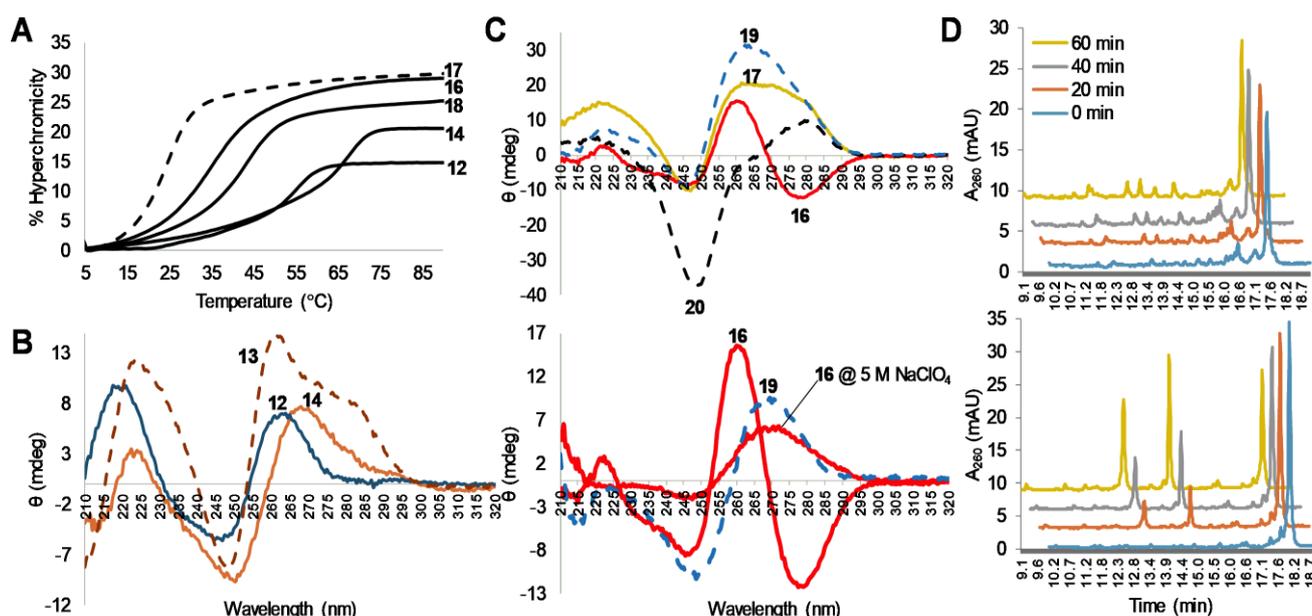


Figure 4. Base-pairing Behaviour and Enzymatic Stability of Non-self-complementary (ribuloNA-RNA(DNA))_n Duplexes. All panels refer to sequences in Table 1 and Supporting Table S6. A) UV-T_m curves of strictly-alternating ribuloNA-RNA and ribuloNA-DNA chimeric sequences; sample [strand] = 4 μM. B) CD curves of r(ⁱaAⁱt) and (ⁱtAⁱgC) chimeric duplexes at 0 °C compared to canonical duplexes; sample [duplex] = 2 – 3 μM. C) CD curves representing Z-like conformation of (ribuloNA-DNA)_n (**16**) at 0 °C compared to canonical duplexes (top) and the Z-to-A form transition with 5 M NaClO₄ (bottom). **19** (bottom) at 60 °C; sample [duplex] = 4 μM. D) AE-FPLC chromatograms of ribuloNA-RNA sequences (Top: 1'-r(ⁱgAⁱgAⁱtCⁱtAⁱtAⁱtCⁱgCⁱtA)-3'; Bottom: 5'- r(TⁱaATⁱaATⁱaTAⁱaATⁱtT)-PO₄-3'; Supporting Table S2) incubated in 0.6% Human Serum over a 60 min period (Supporting Fig. S60). Dashed lines (A – C) = canonical duplex; buffer (A – C): 10 mM Na₂HPO₄, 100 μM EDTA, 1 M NaCl, pH 7.0 (except for **16** in panel C (bottom) = 5 M NaClO₄). See Supporting Information for panel D protocol.

Expanding the repeating chimeric arrangement from (ribuloNA-RNA)_n to (ribuloNA-RNA-RNA)_n or (ribuloNA-ribuloNA-RNA)_n led to loss of hybridization (Supporting Table 6, entries 28 and 29). In addition, the increased RNA content in r(ⁱaAA)_n sequences enhanced their susceptibility to nuclease-mediated degradation (Fig. 4D). These results illustrate (a) the strict nature of the alternating ribuloNA-RNA “dimeric” arrangement unique to this system, and (b) the equal importance of both backbone- and nucleobase-complementarity for designing novel informational systems.

The “dimeric” backbone design was also applied to another related alternative informational system with limited base-pairing properties, *iso*GNA³³ – an acyclic version of ribuloNA where the C4' and C5' carbons of ribuloNA have been removed (Fig. 1). The synthetic alternating (RNA(rA)-*iso*G(ⁱt))_n sequence r((Tⁱa)₈T) **21** (Table 3) formed a hairpin structure with a thermostability of 41.2 °C, while the reverse r((Aⁱt)₈U) sequence formed a weaker hairpin with a T_m 25.9 °C (Supporting Figs. S38, S39 and Supporting Table S8), mirroring the trend in thermostability observed when comparing ribulo-(ⁱt) and ribulo-(ⁱa) inserts in an RNA backbone. This interspersed blending of *iso*GNA and RNA provides yet another example of how heterogeneous-mixed backbone systems can yield base-pairing systems, even when homogeneous-backbone *iso*GNA sequences (ⁱaⁱt)₈ themselves do not.³³ Based on the observation that both ribulo-(ⁱt) and *iso*G-(ⁱt) inserts form stable chimeric constructs with RNA, we

fashioned, as proof-of-principle, three self-complementary oligonucleotides **22** – **24** containing ribuloNA-RNA-*iso*GNA-RNA arrangements (Table 3, Scheme 2, Supporting Table S9), with an aim to extend the generality of the alternating-sugar-residue backbones. We found that these sequences **22** – **24** were equally capable of forming stable duplexes with T_m values of 41.7 – 79.5 °C, comparing favorably to the parent RNA duplexes **6** and **7**. Duplex thermostability of the full (ribuloNA-RNA)_n (**5**) self-complementary duplex is weakened in a linear fashion by – 6.2 to –7.4 °C per additional *iso*G-(ⁱt) insert, eventually forming hairpin **24** which highlights *iso*GNA's inherent flexibility, mimicking the formation of hairpin **21** (Fig. 5A/B; Supporting Fig. S45C/D). Being the acyclic version of ribuloNA, the *iso*GNA unit could adopt the syn-like conformation necessary for the ribuloNA-based chimeras to hybridize. Moreover, thermodynamic analysis of this tri-sugar-backbone system reveals that the strictly-alternating ribuloNA-RNA design has achieved “chimeric homogeneity”, which like homogeneous backbones, pays a penalty when the pattern is interrupted by destabilizing modifications. For instance, the entropic penalty incurred by duplex **22** containing three more *iso*G-(ⁱt) inserts than duplex **23** reveals a less-favorable, higher-energy duplex (Δ(ΔH) = 31.8 kcal/mol, Supporting Table S10). The CD spectra also reveal that the ribuloNA residues contribute strongly to control the helical structure, while diminishing to an overall A-like conformation with increasing *iso*GNA and RNA character (Fig. 5C).

Table 3. UV- T_m Values of Select Chimeric ribuloNA- and isoGNA-Containing RNA Oligonucleotides.^[a]

Duplex	# isoGNA:ribuloNA residues	Self-Complementary Sequences	T_m (°C)
21*	8 : 0	5'-r((A ⁱ t) ₈ A)-3'	41.2 ^[a]
22	4 : 4	5'-r((A ⁱ tArt) ₂ -(A ⁱ tA ⁱ t) ₂ -A)-3'	59.3 ^[a]
23	1 : 7	5'-r((A ⁱ t) ₂ -A ⁱ t-(A ⁱ t) ₅ -A)-3'	79.5 ^[a]
24*	7 : 1	5'-r((A ⁱ t) ₂ -A ⁱ t-(A ⁱ t) ₅ -A)-3'	41.7 ^[a]

[a] [Duplex] = 2 μ M (4 μ M C_{tot}) unless otherwise specified. r or 'N = RNA; 'r = ribuloNA; 'n = isoGNA; UV- T_m measured in phosphate buffer: 10 mM Na₂HPO₄, 100 μ M EDTA, 150 mM NaCl, pH 7.0.

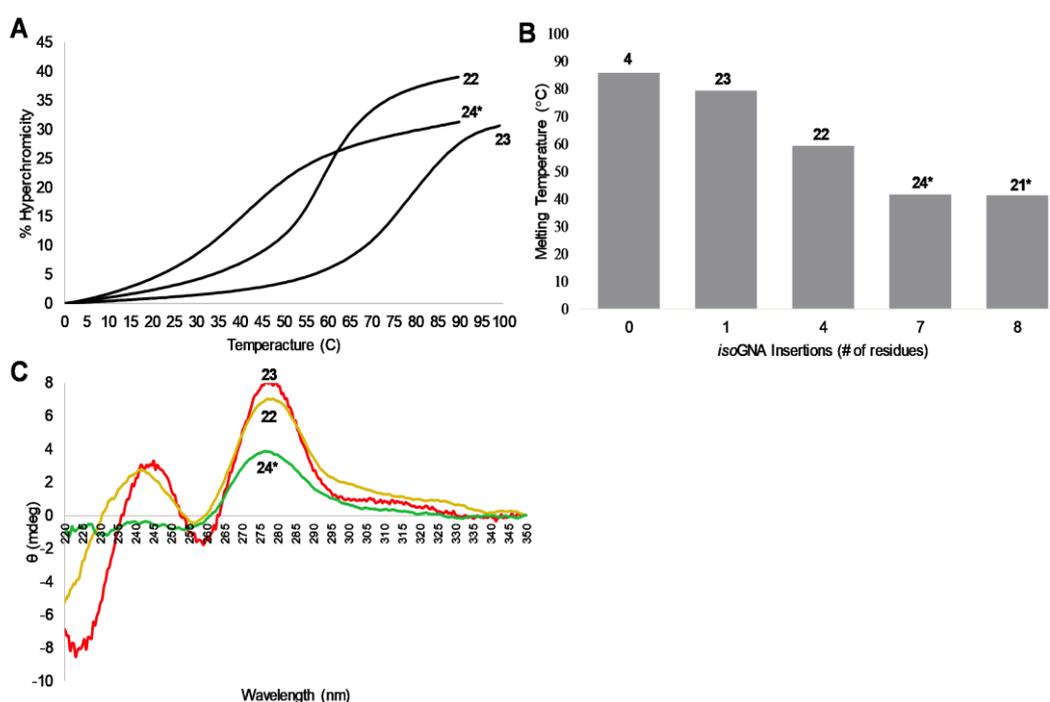


Figure 5. Comparison of Base-pairing propensity of isoGNA and tri-sugar ribuloNA-RNA-isoGNA-RNA chimeric-backbone sequences. All panels refer to sequences in Table 3; sequence 4 is from Table 1. *: represents hairpin secondary structure. A) UV- T_m curves of tri-sugar chimeric sequences. B) Thermostabilities (T_m) of tri-sugar-backbone sequences as a function of isoGNA inserts within the parent sequence 4. Sample [strand] (A/B) = 4 μ M. C) The influence of isoGNA inserts on CD behavior at 0 °C. Sample [duplex] = 2 – 2.5 μ M; buffer (A – C): 10 mM Na₂HPO₄, 100 μ M EDTA, 150 mM NaCl, pH 7.0 (23 in panel C, no salt).

Conclusions

In principle, such an “expanded chimeric-repeat unit” strategy consisting of two isomeric cyclic-sugars with differing chirality (L-ribose and D-ribose) and an acyclic (glycerol) unit, demonstrates a simple, yet viable potential of generating nucleic acid systems with diversified properties.^{30,34} For example, a functioning system based on a two-letter code^{35,36} using only adenine and thymine can, in principle, potentially be developed,

since it may be possible to mimic a four letter code (“GC” and “AU”). By using the stronger ribulo-pyrimidine–ribo-purine (‘t-rA) combination, the “GC” unit can be impersonated, while the weaker ribulo-purine–ribo-pyrimidine ‘a-rT combinations would simulate the “AU” unit.

The observation that heterogeneous-backbone duplexes containing alternating ribuloNA-RNA or ribuloNA-DNA residues are more stable than corresponding alternating RNA-DNA

duplexes is significant from two aspects: a) ribuloNA is a highly-destabilizing alternative backbone system with no self-pairing or RNA/DNA cross-pairing capability,¹⁷ and, more importantly, b) chimeric-backbone-heterogeneous RNA-DNA duplexes are weaker than the corresponding homogeneous RNA or DNA duplexes,³⁷ even though RNA and DNA cross-pair efficiently with each other. The latter aspect is also true for many of the XNAs which base-pair with complementary RNA and DNA¹⁵ with few exceptions.³⁸ Considering these facts, the results presented here suggest an unusual design – that combining oligonucleotide systems which do not pair with themselves or cross-pair with others may lead to novel sequence-specific pairing systems with more favorable functional profiles; while the opposite seems to occur upon combining self- and cross-pairing-capable systems. Governed by this trend and the relative simplicity with which it can be employed, it is possible to envision another dimension to expanding the repertoire of modified nucleic acid backbones through varied combinations employing non-functioning XNAs to fashion hybrid systems (oligomers with heterogeneity in both their backbone and nucleobase composition) with varied properties.

Experimental Section

Monomer Synthesis: Ribulofuranoside ('t and 'a) phosphoramidite monomers were synthesized as outlined in the literature¹⁸, and the synthesis of both 'c and 'g phosphoramidite intermediates are outlined in the Supporting Information (Sections 2.1 – 2.2). Microwave-assisted phosphorylation of all four ('t, 'a, 'c and 'g) phosphoramidites was carried out as described in literature.³⁹

Oligonucleotide Synthesis: All modified oligonucleotides were synthesized in-house through automated solid-support DMT-phosphoramidite chemistry as described in the literature⁴⁰ using slightly modified reagents, coupling protocols and deprotection/purification strategies (see Supporting Section 4.0).

Biophysical Analysis: Oligonucleotide samples (single- and double-stranded) were analyzed by UV-monitored thermal denaturation experiments, monitored at λ 254, 260 and 272 nm. UV curves were standardized and represented as % Hyperchromicity (%H), from which thermal melting (T_m) values were calculated by taking the first derivative ($\Delta\%H/\Delta^\circ C$) using Microsoft Excel and averaged from a minimum of two heating/cooling cycles. Oligonucleotide samples were further characterized by temperature-dependent circular dichroism spectroscopy (CD) as outlined in the Supporting Information (Section 5.0).

Nuclease Stability Assays: Modified and canonical single-stranded oligonucleotide samples were subjected to nuclease digestion through incubation in a 0.6% Human Serum buffer for 60 min. Nuclease activity was halted as outlined in the literature,⁴¹ and analyzed by anion-exchange FPLC monitoring at λ 254, 260 and 280 nm (see Supporting Information, Section 6.0, for further details).

Computational Modeling: Oligonucleotide structures were generated submitted to conformational calculations using the Amber16 molecular mechanics package.²² The ff99OL3 force field was chosen (parmbsc0 α/γ + χ OL3 to ff99).^{23, 24} Additionally, RESP atomic charges were

calculated for the ribulo nucleotides.²⁵ Gamess2016 was used to calculate electrostatic potentials at the 6-31G* level.²⁶ Those potentials were then converted into RESP atomic charges as described in Cieplak *et al.*²⁷

Acknowledgements

This work was jointly supported by NSF and the NASA Astrobiology Program under the NSF Center for Chemical Evolution, CHE-1504217. We thank and Dr. Jayasudhan Yerabolu for providing *iso*GNA intermediates, Professors Reza Ghadiri, Gerald Joyce, Nicholas Hud and our lab members for helpful discussions and feedback. T. E. thanks the NASA Astrobiology Postdoctoral Program for a fellowship. M.F. and P.H. are indebted to the Fund of Scientific Research, Flanders (contract FWO G0188) and to the KU Leuven (OT/14/128) for financial support.

Keywords: Chimeric Nucleic Acid • RNA • DNA • XNA • Oligonucleotides

- [1] W. Saenger, *Principles of Nucleic Acid Structure*. (Springer-Verlag New York, Incorporated, New York, NY, U.S.A., 1984).
- [2] P. Herdewijn, *Liebigs Ann.* **1996**, 1337-1348.
- [3] C.J. Leumann, *Biorg. Med. Chem.* **2002**, *10*, 841-854.
- [4] P. Herdewijn, P. Marliere, *Chem. Biodiv.* **2009**, *6*, 791 – 808.
- [5] P.E. Nielsen, *Chem. Biodiv.* **2010**, *7*, 786-804.
- [6] A. Eschenmoser, *Angew. Chem. Int. Ed.* **2011**, *50*, 12412-12472.
- [7] A.V. Pinheiro, Han, W.M.D., Shih, H. Yan, *Nat. Nano.* **2011**, *6*, 763-772.
- [8] M.P. Robertson, G.F. Joyce, *The Origins of the RNA World. Cold Spring Harbor Perspectives in Biology* **4**, (2012).
- [9] V.B. Pinheiro, P. Holliger, *Curr. Opin. Chem. Biol.* **2012**, *16*, 245-252.
- [10] V.B. Pinheiro, *et al. Science* **2012**, *336*, 341-344.
- [11] A.E. Engelhart, M. W. Powner, J. W. Szostak, *Nat. Chem.* **2013**, *5*, 390-394.
- [12] M. Winnacker, E. T. Kool, *Angew. Chem. Int. Ed.* **2013**, *52*, 12498-12508.
- [13] D.A. Malyshev, F.E. Romesberg, *Angew. Chem. Int. Ed.* **2015**, *54*, 11930-11944.
- [14] A.I. Taylor, *et al. Nature* **2015**, *518*, 427-430.
- [15] I. Anosova, *et al. Nucleic Acids Res.* **2016**, *44*, 1007-1021.
- [16] R. Krishnamurthy, *On the Emergence of RNA. Isr. J. Chem.* **2015**, *55*, 837-850.
- [17] M. Stoop, G. Meher, P. Karri, R. Krishnamurthy, *Chem. Eur. J.* **2013**, *19*, 15336-15345.
- [18] E. Schrödinger, 1944 *What Is Life? The Physical Aspect of the Living Cell*. Cambridge University Press, Cambridge
- [19] K. Rippe, N.B. Ramsing, R. Klement, T.M. Jovin, *J. Biomol. Struct. Dyn.* **1990**, *7*, 1199-1209.
- [20] P. Yakovchuk, E. Protozanova, M.D. Frank-Kamenetskii, *Nucleic Acids Res.* **2006**, *34*, 564-574.
- [21] N. Foloppe, B. Hartmann, L. Nilsson, A.D. MacKerell Jr, *Biophys. J.* **2002**, *82*, 1554-1569.
- [22] D. A. Case, *et al.*, AMBER 2017, University of California, San Francisco.
- [23] E.F. Pettersen, T.D. Goddard C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605-1612.
- [24] R. Lavery, M. Moakher, J.H. Maddocks, D. Petkeviciute, K. Zakrzewska, *Nucleic Acids Res.* **2009**, *37*, 5917–5929.

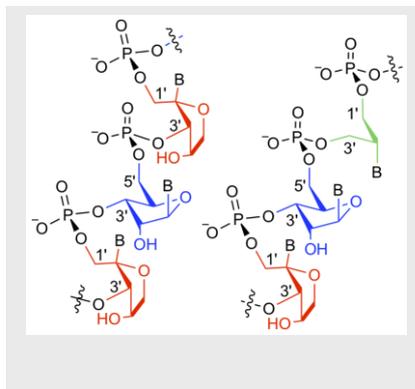
- [25] A. Rich, A. Nordheim, A.H.J. Wang, *Annu. Rev. Biochem.* **1984**, *53*, 791-846.
- [26] K. U. Schöning, *et al. Helv. Chim. Acta.* **2002**, *85*, 4111-4153.
- [27] H.H. Klump, E. Schmid, M. Wosgien, *Nucleic Acids Res.* **1993**, *21*, 2343-2348.
- [28] J. Wereszczynski, I. Andricioaei, *J. Phys. Chem. B* **2010**, *114*, 2076-2082.
- [29] K. Hall, P. Cruz, I. Tinoco Jr., T. M. Jovin, J. H. van de Sande, *Nature* **1984**, *311*, 584-586.
- [30] J.-F. Lutz, M. Ouchi, D. R. Liu, M. Sawamoto, *Science* **2013**, *341*, 1238149.
- [31] R.M. Dirks, M. Lin, E. Winfree, N.A. Pierce, *Nucleic Acids Res.* **2004**, *32*, 1392-1403.
- [32] Y. Krishnan, M. Bathe, *Trends Cell Biol.* **2012**, *22*, 624-633.
- [33] P. Karri, V. Punna, K. Kim, R. Krishnamurthy, *Angew. Chem. Int. Ed.* **2013**, *52*, 5840-5844.
- [34] M.K. Schlegel *et al. J. Am. Chem. Soc.* **2017**, *139*, 8537-8546.
- [35] J.S. Reader, G. F. Joyce, *Nature* **2002**, *420*, 841-844.
- [36] V. Pezo, F. W. Liu, M. Abramov, M. Froyen, P. Herdewijn, P. Marlière, *Angew. Chem. Int. Ed.* **2013**, *52*, 8139-8143.
- [37] J.V. Gavette, M. Stoop, N.V. Hud, R. Krishnamurthy, *Angew. Chem. Int. Ed.* **2016**, *55*, 13204-13209.
- [38] J. Wengel, *Acc. Chem. Res.* **1999**, *32*, 301-310.
- [39] G. Meher, T. Efthymiou, M. Stoop, R. Krishnamurthy, *Chem. Commun.* **2014**, *50*, 7463-7465.
- [40] M.H. Caruthers *et al. Method. Enzymol.* **1987**, *230*, 287-313.
- [41] I.S. Cho *et al. Biotechnol. Lett.* **2008**, *30*, 1901-1908.

Entry for the Table of Contents

FULL PAPER

Greater than the sum of its parts:

Incorporating weakly or non-base pairing XNA with RNA and DNA gives rise to chimeric orthogonal informational systems.



Tim Efthymiou^[a,b], Jesse Gavette^[a,b],
Matthias Stoop^[a,b], Francesco De
Riccardis^[a,b,c], Mathy Froeyen^[d], Piet
Herdewijn^[d], Ramanarayanan
Krishnamurthy*

Page No. – Page No.

**Chimeric XNA – An Unconventional
Design for Orthogonal Informational
Systems**