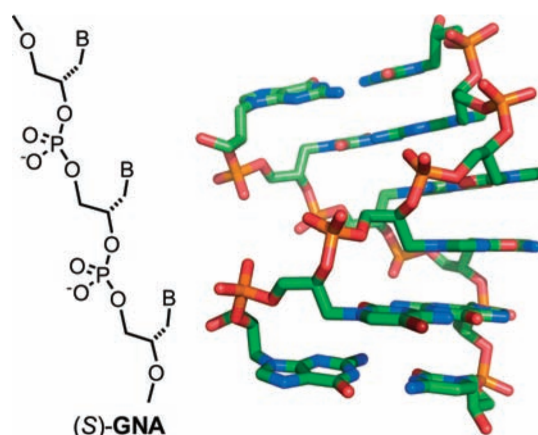
ACCOUNTS
of chemical research

Synthesis and Properties of the Simplified Nucleic Acid Glycol Nucleic Acid

ERIC MEGGERS* AND LILU ZHANG
Fachbereich Chemie, Philipps-Universität Marburg, Hans-Meerwein-Strasse, 35032 Marburg, Germany

RECEIVED ON JANUARY 5, 2010

CONSPECTUS



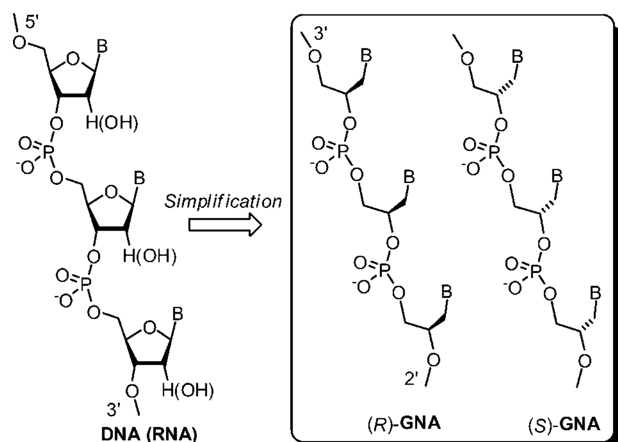
The nucleosides of glycol nucleic acid (GNA), with the backbone comprising just the three carbons and one stereocenter of propylene glycol (1,2-propanediol), probably constitute the simplest possible building blocks for a chemically stable nucleic acid that contains phosphodiester bonds. However, it was not until 2005 that the astonishing duplex formation properties of GNA homoduplexes were discovered in our laboratory. The *R*- and *S*-enantiomers of GNA, (*R*)-GNA and (*S*)-GNA, pair in like-symmetric combinations to form highly stable antiparallel duplexes in a Watson−Crick fashion, with thermal and thermodynamic stabilities exceeding those of analogous duplexes of DNA and RNA. Interestingly, (*R*)-GNA and (*S*)-GNA do not significantly cross-pair with each other, either in a parallel or antiparallel fashion. GNA discriminates strongly in favor of the Watson−Crick base-pairing scheme, with only slightly lower fidelity than DNA. Two (*S*)-GNA homoduplex structures recently determined by X-ray crystallography, one a brominated 6-mer duplex and the other an 8-mer duplex containing two copper(II) ions, reveal that the overall GNA double helix is distinct from canonical A- and B-form nucleic acids. The structure is perhaps best described as a helical ribbon loosely wrapped around the helix axis. Within the backbone, the propylene glycol nucleotides adopt two different conformations, *gauche* and *anti*, with respect to the torsional angles between the vicinal C3′−O and C2′−O bonds. A strikingly large backbone−base inclination results in extensive zipper-like interstrand and reduced intrastrand base−base interactions. This strong backbone−base inclination might explain the observation that neither the *R*- nor *S*-enantiomer of GNA cross-pairs with DNA, whereas (*S*)-GNA can interact with RNA strands that are devoid of G:C base pairs. Given the combination of structural simplicity, straightforward synthetic accessibility, and high duplex stability of GNA duplexes, GNA affords a promising nucleic acid scaffold for biotechnology and nanotechnology. Along these lines, we describe the functionalization of GNA duplexes through the incorporation of metal-ion-mediated base pairs. Finally, the properties of GNA discussed here reinforce its candidacy as one of the initial genetic molecules formed during the origins of life on Earth.

Introduction

Unrelated to its function in biology, DNA has emerged as an ideal building block for nanotechnology due to its ability to self-assemble into predictable structures in a programmable fashion.[1] Precisely designed two- and three-dimensional DNA structures hold promise as scaffolds for nanoscale devices such as sensors, switches, tweezers, diodes, and transistors. However, applications along these lines will rely on the controlled mod-

Published on the Web 04/20/2010 www.pubs.acs.org/acr
10.1021/ar900292q © 2010 American Chemical Society

SCHEME 1. Constitutions of DNA, RNA, and *S*- and *R*-Enantiomers of GNA



ification of DNA in order to introduce the required functionality.[2] Modified nucleic acids are nowadays readily accessible through standard automated solid phase synthesis using nucleoside phosphoramidites as the building blocks, enabling the synthesis of oligonucleotides greater than 100 bases. Unfortunately, in particular due to the stereocenters in the deoxyribose moiety and especially the anomeric C1′-stereocenter, the chemical synthesis of modified nucleotide building blocks for automated nucleic acid synthesis is often tedious and lengthy, includes the chromatographic separation of diastereomers, and therefore proves challenging to scale-up. Without doubt, the chemical synthesis of modified nucleotides is a critical bottleneck for progress in this and related areas of research.

With this challenge in mind, we started out 6 years ago with the goal to design an artificial nucleic acid that is composed of structurally simplified and synthetically easily accessible building blocks, while retaining the desired base-pairing properties. In this Account, we describe our discovery that a glycol nucleic acid (GNA),[3] comprised of a simplified acyclic propylene glycol phosphodiester backbone (Scheme 1), can form stable antiparallel duplexes in a Watson−Crick fashion.[4] GNA uniquely combines atom economy, structural simplicity, and high duplex stability.

Background and Related Studies

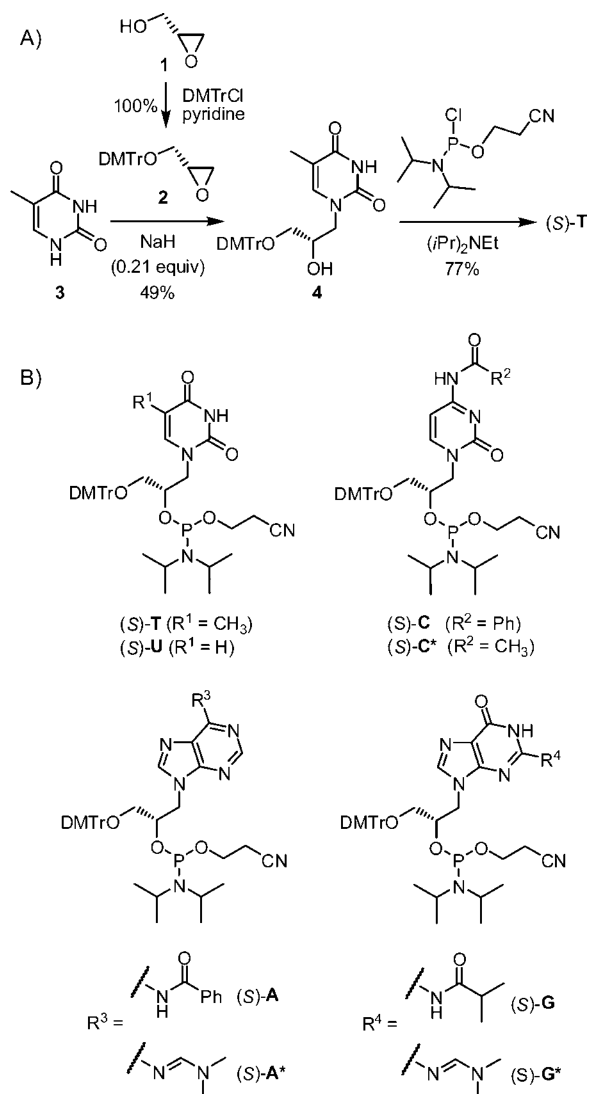
Compared with 2′-deoxyribonucleosides, acyclic nucleosides are structurally simplified due to the reduced number of stereocenters, and they should therefore serve in principle as ideal building blocks for simplified artificial nucleic acids. Among acyclic nucleosides, 2,3-dihydroxypropyl-modified nucleobases appear to be particularly attractive building blocks due to the minimal number of backbone carbons and stereocenters.

Racemic *N*-(2,3-dihydroxypropyl) derivatives of nucleobases were first synthesized by Ueda et al.[5] and shortly thereafter by the group of Imoto.[6] Imoto's group subsequently also reported the synthesis of oligomers by condensation of dihydrogenphosphates with *N*,*N*′-dicyclohexylcarbodiimide.[7] Aware of the importance of chiral recognition in biology, enantiomerically pure 2,3-dihydroxypropyl nucleosides and derivatives were first synthesized and extensively investigated by the group of Holý.[8−11]

In 1995, Cook et al. from Isis Pharmaceuticals filed a patent regarding the synthesis of glycol-based phosphate-linked oligomers.[12] Racemic and *S*-enantiomeric 2,3-dihydroxypropyl nucleosides were synthesized starting from racemic glycidol and (*R*)-(+)-glycidol, respectively.[13] Phosphodiester oligomers with natural and unnatural nucleobases were prepared from phosphoramidite building blocks and investigated as scaffolds for the design of enzyme inhibitors.[14] However, duplex formation properties of such artificial oligonucleotides were not reported. The same year, Wengel's group disclosed the investigation of pairing properties of oligodeoxynucleotides containing one or two (*S*)-1-(2,3-dihydroxypropyl)thymines, and they found pronounced destabilizations of the evaluated duplexes containing this and other acyclic nucleoside analogues.[15,16] In fact, during the 1980s and 1990s, several groups reported the synthesis and investigation of oligodeoxynucleotides containing acyclic nucleosides.[15,17] In all published studies, the incorporation of single and multiple acyclic nucleosides into oligodeoxynucleotides afforded pronounced reductions in duplex stabilities. Based on these results, it was concluded that acyclic nucleosides are too flexible to produce the required conformational preorganization of the individual strands for stable duplex formation.

In the course of studies toward the chemical etiology of nucleic acid structure,[18] Eschenmoser and co-workers reported in 2000 that L-α-threofuranosyl oligonucleotides (TNA) are capable of forming stable antiparallel duplexes based on Watson−Crick base pairing and can even form heteroduplexes with complementary strands of DNA and RNA.[19] This exciting discovery demonstrated that the "six-bonds-per-backbone" rule can be violated by having five bonds instead in an arrangement of vicinal phosphodiester groups. Based on the surprising results with TNA and subsequent experiments with the backbone-shortened L-α-lyxopyranosy-(4′→3′)-oligonucleotide system, Eschenmoser raised the question whether base-pairing capability might be even compatible with acyclic backbones containing vicinal phosphodiester groups.[20]

SCHEME 2. (*S*)-Phosphoramidite Building Blocks for the Synthesis of GNA Oligonucleotides: (A) Synthesis of (*S*)-**T**; (B) First (**A**, **G**, **C**, **T**, **U**) and Improved Second Generation (**A***, **G***, **C***) Phosphoramidite Building Blocks



GNA Synthesis

The synthesis of GNA oligonucleotides from phosphoramidite building blocks was first reported by Isis Pharmaceuticals[12−14] and was later improved and modified by our group.[21,22] The propylene glycol nucleosides are easily accessible starting from commercially available enantiomerically pure "spring-loaded" (*R*)-(+)- and (*S*)-(−)-glycidol. The representative synthesis of the phosphoramidite (*S*)-**T** is shown in Scheme 2A. Tritylation of (*R*)-(+)-glycidol (**1**) with DMTrCl provides quantitatively (*S*)-glycidyl 4,4′-dimethoxytrityl ether (**2**). The following reaction of epoxide **2** with thymine (**3**) in the presence of around 0.2 equiv of sodium hydride affords in a regioselective and stereospecific fashion the ring-opened compound **4** (49%), which is subsequently converted to the phosphoramidite (*S*)-**T** by reaction with 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite in the presence of Hünig's base (77%). Thus, (*S*)-**T** is synthetically accessible from commercially available starting materials in just three steps with an overall yield of 38%.[21] In a similar fashion, the phosphoramidite building blocks (*S*)-**U**, (*S*)-**C**, and (*S*)-**C*** are accessible in three steps with overall yields of 40%, 39%, and 44%, respectively (Scheme 2B).[21,22] The purine phosphoramidites (*S*)-**A** and (*S*)-**G** can be synthesized in four (overall yield of 26%) and five steps (overall yield of 8%), respectively (Scheme 2B).[21] The low overall yield of the (*S*)-**G** synthesis in addition to its slight instability, which prohibits a purification by column chromatography or storage at −20 °C for longer times, prompted us to investigate alternative protection groups for the exocyclic amino group and resulted in the development of the phosphoramidites (*S*)-**G*** and (*S*)-**A*** with *N*-dimethylformamidine protection groups (Scheme 2B).[22] N^2-Dimethylformamidine phosphoramidite (*S*)-**G***, unlike its N^2-isobutyryl counterpart (*S*)-**G**, is stable to flash chromatography over silica gel and has a significantly improved overall synthetic yield of 30% over five steps. Thus, phosphoramidite building blocks for automated solid phase GNA synthesis are readily available.

GNA oligonucleotides were synthesized on CPG supports with standard protocols for 2-cyanoethyl phosphoramidites, except that the coupling time was increased to 3 min.[21] Typically, GNA oligonucleotides synthesized from (*S*)-**A**, (*S*)-**G**, and (*S*)-**C**, along with (*S*)-**T**, (*S*)-**U**, or the *R*-enantiomeric counterparts, were synthesized in the trityl-on mode and cleaved from the resin with concentrated ammonia at 55 °C for 12 h. GNA oligonucleotides synthesized from the second generation phosphoramidites (*S*)-**G***, (*S*)-**A***, and (*S*)-**C***, along with (*S*)-**T**, were subjected to a quicker deprotection procedure by using a 1:1 mixture of 40% aqueous methylamine and 25% aqueous ammonium hydroxide at 55 °C for only 15−20 min.[22] Purification and detritylation of the GNA oligonucleotides were performed according to published standard procedures.[21,22]

Due to the vicinal arrangement of phosphodiester bridges in GNA, in combination with a rotational freedom in the acyclic backbone, we were initially concerned about the potential hydrolytic instability of GNA strands. In analogy to the hydrolysis of RNA, the free terminal OH groups are well positioned to attack adjacent internal phosphodiester linkages in a cleavage−transesterification mechanism, resulting in the release of cyclic phosphates at the termini. Gratifyingly, HPLC-monitored stability tests of GNA oligonucleotides demonstrated their high stability in buffered solution (pH 7.0) at room

TABLE 1. Thermal Stabilities of Duplexes of GNA, DNA, and DNA/GNA hybrids[a]

Entries	Duplexes[b]	T_m (°C)[c]
1	5′-CACATTATTGTTGTA-3′ 3′-GTGTAATAACAACAT-5′	47
2	5′-CACATTATTGTTGTA-3′ 3′-GTGTAATAACAACAT-5′	34 (40)
3	5′-CACATTATTGTTGTA-3′ 3′-GTGTAATAACAACAT-5′	29 (28)
4	3′-CACATTATTGTTGTA-2′ 3′-GTGTAATAACAACAT-5′	no T_m (no T_m)
5	5′-CACATTATTGTTGTA-3′ 3′-GTGTAATAACAACAT-5′	33 (33)
6	5′-CACATTATTGTTGTA-3′ 3′-GTGTAATAACAACAT-5′	39 (39)
7	3′-CACATTATTGTTGTA-2′ 2′-GTGTAATAACAACAT-3′	71

[a] Measured in 10 mM sodium phosphate (pH 7.0) with 100 mM NaCl and 2 μM of each strand. [b] GNA nucleotides are highlighted. [c] Shown for (*S*)-nucleotides, with (*R*)-nucleotides in parentheses.

temperature.[21] Furthermore, Szostak and co-workers reported that GNA phosphodiester linkages are stable at pH 9.0 (20 °C) and only slowly degrade at pH 13 with a half-life of 141 h.[23] In any instance, this risk of base-catalyzed decomposition can be completely circumvented by capping the free terminal OH groups. However, since GNA strands are very stable in buffer and because we could not observe any difference in synthesis yields and final purities of GNA oligonucleotides bearing free OH termini compared with GNA strands synthesized with 3′- and 2′-hydroxypropyl caps, all our experiments with GNA described in this Account were performed with uncapped GNA strands.

From Individual GNA Nucleotides in DNA to a Completely Acyclic Backbone

In order to understand the relationship between properties of single incorporated glycol nucleotides in DNA and completely acyclic GNA strands, we investigated different DNA/GNA hybrid systems by UV-melting as shown in Table 1.[24] We first incorporated single and multiple *R*- and *S*-enantiomeric GNA nucleotides into DNA and found that they strongly compromise the thermal DNA duplex stabilities (Table 1, entries 1−4). These results are consistent with a report by Wengel's group more than a decade ago in which 17mer oligonucleotides containing one or two (*S*)-1-(2,3-dihydroxypropyl)thymine nucleotides in middle positions afforded a pronounced reduction in duplex stability.[15] However, incorporated single and three consecutive GNA base pairs resulted in a diminished destabilization of only 7 and 1.3 °C per GNA nucleotide, respectively (Table 1, entries 5 and 6). Replacing finally all deoxyribose nucleotides of the 15mer by *S*-enantiomeric glycol nucleotides provided an extremely stable duplex with a melting point of 71 °C, exceeding the stability of the analogous DNA duplex by 24 °C (Table 1, entry 7, and Figure 1A). CD melting experiments confirmed the formation of a very thermally stable GNA duplex (Figure 1B).

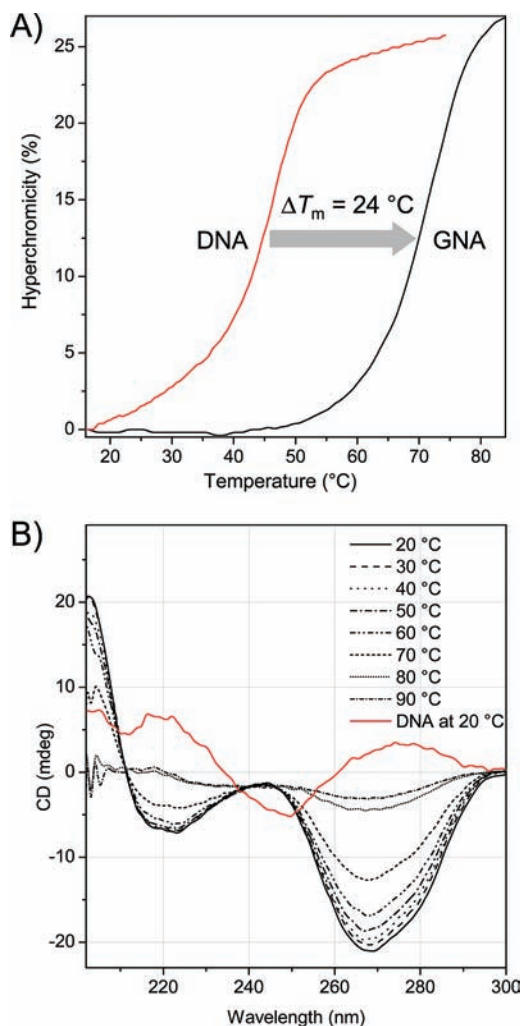


FIGURE 1. UV and CD melting of a 15mer (*S*)-GNA duplex demonstrating the high thermal stability of GNA duplexes. A DNA duplex of the same sequence is shown in red as a reference. Experiments were performed in 10 mM sodium phosphate (pH 7.0) with 100 mM NaCl. For strand sequences, see Table 1, entries 1 and 7.

These melting experiments with DNA/GNA hybrid strand systems thus clearly demonstrate that the destabilization of propylene glycol nucleotides in DNA duplexes is not caused by a lack of conformational preorganization of the acyclic nucleotide but is rather due to the incompatibility of the DNA and the GNA backbone. Neither the (*S*)-GNA nor the enantiomeric (*R*)-GNA strands form heteroduplexes with DNA.

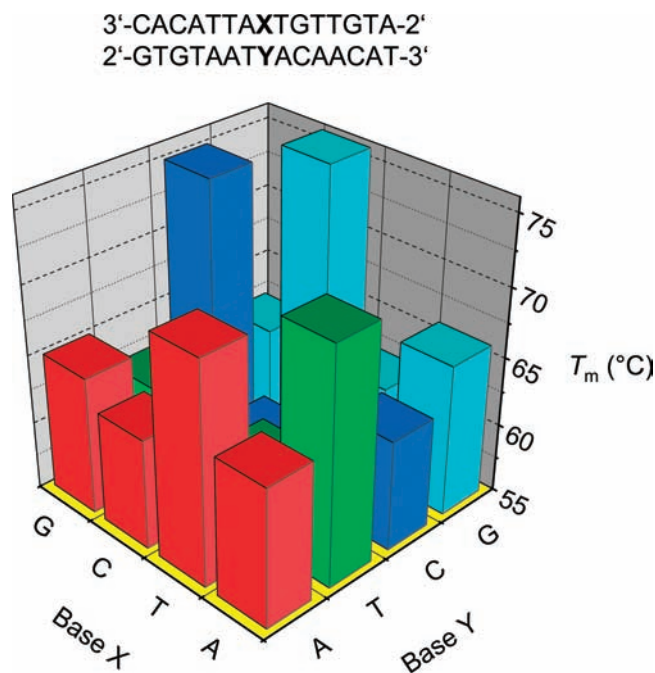


FIGURE 2. Watson−Crick base-pairing fidelity in GNA duplexes. T_m values were determined with 2 μM duplex concentration in 10 mM sodium phosphate and 100 mM NaCl.

Duplex Formation Properties of GNA

In analogy to DNA and RNA, homoduplex formation of GNA relies on Watson−Crick base pairing with antiparallel strand complementarity and a 1:1 strand stoichiometry.[4,24] A:T and G:C display significantly higher stabilities compared with any other base pair combinations with G:C being approximately 4 °C more stable than A:T. The thermal destabilization of mismatches in the investigated GNA duplex in Figure 2 ranges from $\Delta T_m = -6$ to -18 °C, compared with mismatches in the reference DNA duplex of the same sequence ranging from $\Delta T_m = -10$ to -23 °C.[24] These data demonstrate that GNA discriminates strongly in favor of the Watson−Crick base-pairing scheme with only slightly lower fidelity than DNA.

Over the last several years, we investigated a large number of (*S*)-GNA/(*S*)-GNA and a few (*R*)-GNA/(*R*)-GNA homoduplexes of different lengths and sequences (Table 2).[4,21,24] In all analyzed systems, the comparison of duplexes of GNA and DNA of the same sequence revealed that the acyclic backbone of GNA supports duplexes that are thermally significantly more stable ($\Delta T_m = +18$ to $+25$ °C). The higher thermal stabilities of GNA duplexes correlate with higher thermodynamic stabilities at room temperature ($\Delta\Delta G = -2.6$ to -5.9 kcal/mol at 298 K).[24,25] To our surprise, van't Hoff plots revealed that compared with DNA, GNA duplex formation is less exothermic but is entropically significantly more favorable.[25] This lower entropic penalty for GNA is counterintuitive at first glance since one would expect that an acyclic backbone is more flexible and thus entropically disfavored for duplex formation. However, recent studies from our group suggest that this entropic advantage is at least in part the consequence of a strong conformational preorganization of the GNA single strands in combination with especially favorable stacking interactions in the corresponding duplex.[25]

Stable Watson−Crick duplex formation is exclusively observed with homoduplexes of (*S*)-GNA/(*S*)-GNA and (*R*)-GNA/(*R*)-GNA (Figure 3).[4,24] (*S*)-GNA is incompatible with (*R*)-GNA in an antiparallel as well as parallel strand orientation, and neither (*S*)- nor (*R*)-GNA pairs with DNA. However, (*S*)-GNA, but not (*R*)-GNA, can pair with RNA in an antiparallel fashion in duplexes that are devoid of G and C, indicating that this interaction is not based on Watson−Crick base pairing.[24,26] It is also noteworthy that Chaput and co-workers recently reported that (*S*)-GNA is incapable of cross-pairing with TNA although both nucleic acid systems contain a shortened backbone with vicinal phosphodiester bridges and (*S*)-GNA is formally derived from TNA by removing a CH_2O unit from the tetrahydrofuran ring.[26] This surprising result demonstrates that GNA and TNA must adopt different backbone conformations in double helices.

GNA Duplex Structure

We recently determined the atomic resolution structure (0.965 Å) of an (*S*)-GNA duplex formed from the self-complementary strand 3′-G^{Br}CGCGC-2′ by X-ray crystallography.[27] A 5-bromocytosine nucleobase (^{Br}C) was used for phasing the crystallographic data. The employed phosphoramidite (*S*)-^{Br}**C** was synthesized from 5-bromocytosine (**5**) in just three steps with an overall yield of 45% (**5** → **6** → **7** → (*S*)-^{Br}**C**, Scheme 3).

The self-complementary (*S*)-GNA strand 3′-G^{Br}CGCGC-2′ forms a right-handed antiparallel double helix with a helical pitch of 26 Å and 10 base pairs per turn (Figure 4A). The base pairs are displaced from the helix axis (*x*-displacement) by 5.4−6.8 Å, resulting in a large hollow core. The GNA duplex possesses only one large groove, corresponding to the canonical minor groove, whereas the canonical major groove is a convex surface instead. With these features, the GNA helix structure differs significantly from the canonical A- and B-form nucleic acid duplexes and might be best described as a helical ribbon loosely wrapped around the helix axis (Figure 4A). All base pairs form standard Watson−Crick hydrogen-bonding patterns with the 5-bromocytosine nucleotide appearing to have little, if any, distorting effect. The propylene glycol backbone adopts two different conformations with respect to the torsional angles between the vicinal C3′−O and C2′−O bonds, displaying alternating *gauche* and *anti* conformations, and

TABLE 2. Thermal and Thermodynamic Stabilities of (*S*)-GNA and DNA Duplexes[a]

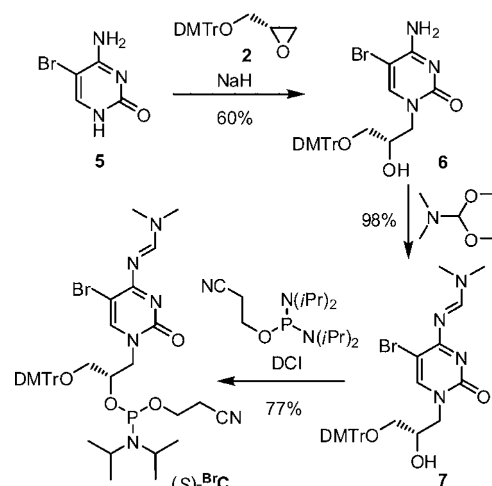
strand size	NaCl concn (mM)	duplexes	T_m (°C)[b]	ΔG (kcal/mol, 298 K)
6mer	1000	3′-ATTAAT-2′ 2′-TAATTA-3′	<20	*c*
8mer	1000	3′-AAATATTT-2′ 2′-TTTATAAA-3′	25	*c*
8mer	150	3′-AACTAGTT-2′ 2′-TTGATCAA-3′	38	*c*
8mer	500	3′-CGAATTCG-2′ 2′-GCTTAAGC-3′	54 (36)	−12.2(−9.4)
8mer	150	3′-ATGCGCAT-2′ 2′-TACGCGTA-3′	57	*c*
15mer	100	3′-CACATTATTGTTGTA-2′ 2′-GTGTAATAACAACAT-3′	71 (47)	−21.1(−15.2)
15mer	500	3′-AATATTATTATTTTA-2′ 2′-TTATAATAATAAAAT-3′	59 (41)	−16.2(−12.4)
15mer	150	3′-A$_{15}$-2′ 2′-T$_{15}$-3′	65 (40)	−15.1 (−12.5)
15mer	150	3′-A$_7$CA$_7$-2′ 2′-T$_7$GT$_7$-3′	64	*c*
15mer	150	3′-A$_7$GA$_7$-2′ 2′-T$_7$CT$_7$-3′	65	*c*
15mer	150	3′-A$_6$C$_3$A$_6$-2′ 2′-T$_6$G$_3$T$_6$-3′	73	*c*
18mer	200	3′-TAAAATTTATATTATTAA-2′ 2′-ATTTTAAATATAATAATT-3′	63 (41)	*c*
18mer	200	3′-TTTTAAATTTTAATATAT-2′ 2′-AAAATTTAAAATTATATA-3′	63	*c*

[a] Measured in 10 mM sodium phosphate (pH 7.0), 2 μM of each strand and the indicated NaCl concentrations. See refs 24 and 25 for more details. [b] DNA values in parentheses. [c] Not determined.

	(*S*)-GNA	(*R*)-GNA
(*S*)-GNA	Yes	No
(*R*)-GNA	No	Yes
DNA	No	No
RNA	Limited	No

FIGURE 3. Antiparallel cross-pairing of (*R*)- and (*S*)-GNA. "Limited" indicates that (*S*)-GNA pairs with RNA in an antiparallel fashion in duplexes that are mainly composed of A:T(U) base pairs.

SCHEME 3. Synthesis of the Phosphoramidite (*S*)-**^{Br}C** for Automated Solid Phase Synthesis of ^{Br}C-Containing GNA Strands[a]



[a] DCI = 4,5-dicyanoimidazole.

interestingly in a fashion so that each base pair contains one nucleotide in *gauche* and one in *anti* conformation (Figure 5). Table 3 lists the average helical parameters for this (*S*)-GNA duplex in comparison to standard B-form and A-form DNA duplexes.

The most striking feature of the (*S*)-GNA duplex is the large average slide of around 3.4 Å between neighboring base pairs which is due to a large backbone−base inclination ranging from −46° to −53° compared with 0° for B-DNA. This strong inclination results in extensive zipper-like interstrand and at the same time reduced intrastrand base−base stacking interactions. The latter are apparently compensated by intrastrand backbone−base hydrophobic interactions, in particular the stacking of the C1′-methylene groups of the propylene glycol backbone against neighboring nucleobases.

Strong inclinations between backbone and base pair axes have also been observed in the unnatural nucleic acids pyranosyl-RNA,[28,29] homo-DNA,[29,30] and hexitol nucleic acid.[31] It has been proposed by Eschenmoser that the backbone−base inclination is an important parameter to assess cross-pairing between different nucleic acids systems.[20] Accordingly, cross-pairing with B-DNA would require a nucleic acid system with a small or absent backbone−base inclination, and this could explain the complete inability of GNA to cross-pair with DNA.

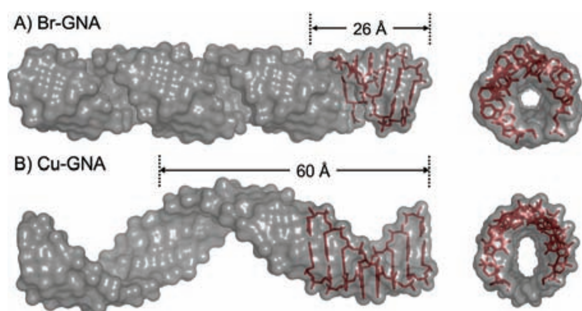


FIGURE 4. Overall determined structures of (*S*)-GNA duplexes: (A) duplex derived from 3′-G^{Br}CGCGC-2′ (Br-GNA); (B) copper-GNA duplex, 3′-CGHATHCG-2′ (Cu-GNA), where H = Cu(II)-dependent hydroxypyridone nucleobase. Single duplexes are shown in red sticks. On the right are views along the helix axis. Reproduced with small changes from ref 27 with permission of the Royal Society of Chemistry.

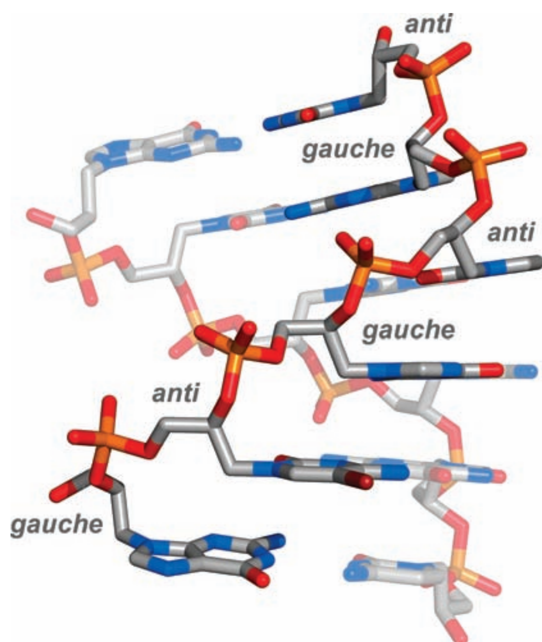


FIGURE 5. Structure of a single (*S*)-GNA duplex from 3′-G^{Br}CGCGC-2′ (PDB code 2WNA). *Gauche* and *anti* refer to the torsional angles between the vicinal bonds C2′−O and C3′−O. Reproduced from ref 27 with permission of the Royal Society of Chemistry.

Metal-Ion-Mediated Base Pairing in GNA

Due to the straightforward synthetic accessibility of GNA nucleotide building blocks in combination with high duplex stabilities, GNA comprises an appealing scaffold for future functional nucleic acid based nanostructures. In order to tune the electronic and magnetic properties of nucleic acids we became interested in studying metal-ion-mediated base pairing in the context of GNA duplexes. For our initial studies, we chose a copper(II)-mediated hydroxypyridone homobase pair (H−Cu−H, Figure 6) and a nickel(II)-mediated pyridylpurine homobase pair (P−Ni−P, Figure 6) scheme,[32] both previously

TABLE 3. Comparison of Average Helical Parameters for (*S*)-GNA Duplexes Derived from 3′-G^{Br}CGCGC-2′ (Br-GNA) and 3′-CGHATHCG-2′ (Cu-GNA) with B-DNA and A-DNA as References[a]

	Br-GNA	Cu-GNA	B-DNA	A-DNA
helical sense	right	right	right	right
residues per turn	10	16	10	12
helical pitch (Å)	26	60	34	34
helical rise (Å)	2.6	3.8	3.4	2.9
x-displacement (Å)	−6.0	−7.0	0.1	−4.2
tilt (deg)[b]	0.5	0.0	0.1	−0.1
roll (deg)[b]	6.4	−2.7	0.6	8.0
twist (deg)[b]	35.7	23.5	36.0	31.0
slide (deg)[b]	−3.4	−3.5	0.2	−1.5
P−P distance (Å)[c]	5.4	5.4	7.0	5.9

[a] See ref 27 for more details. [b] Local base pair step parameters. [c] Intrastrand P−P distances.

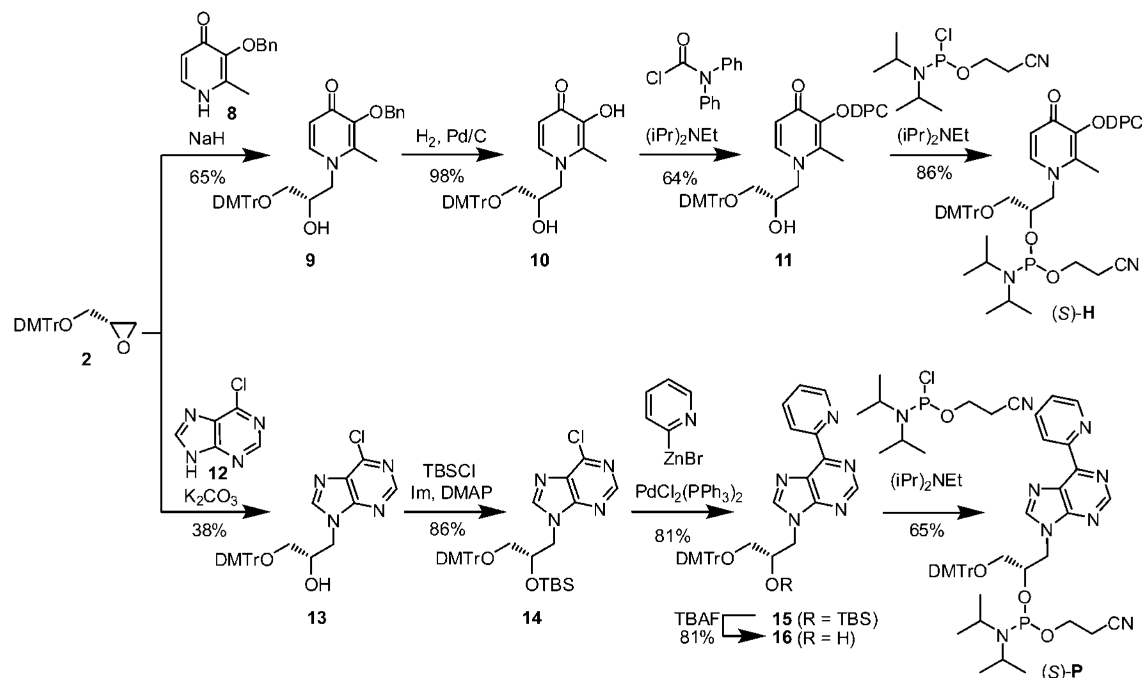
developed for metal-mediated base pairing in DNA by Shionoya et al. and Switzer et al., respectively.[33,34]

Synthesis of Chelating Phosphoramidite Building Blocks. Scheme 4 shows the straightforward synthetic routes to the phosphoramidite building blocks (*S*)-**H** and (*S*)-**P** for automated solid phase synthesis of metallo-GNA strands. Starting from (*S*)-glycidyl 4,4′-dimethoxytrityl ether (**2**), the reaction with 3-benzyloxy-2-methyl-1*H*-pyridin-4-one (**8**) in the presence of catalytic amounts of NaH, afforded in a regio- and stereospecific fashion the ring-opened glycol nucleoside **9**, which was converted after protection group exchange (**9** → **10** → **11**) to the final phosphoramidite building block (*S*)-**H**. Overall (*S*)-**H** was synthesized from commercially available (*R*)-(+)-glycidol in five steps in an overall yield of 34%. This is a significant improvement over the corresponding 2′-deoxynucleotide for which a nine step synthesis plus one additional step for the separation of an anomeric mixture was reported with an overall yield of 8.0%.[33]

Phosphoramidite (*S*)-**P**, leading to pyridylpurine-containing base pairs, was synthesized starting from epoxide **2** by ring-opening with 6-chloropurine (**12**) in the presence of catalytic amounts of K_2CO_3 to afford the glycol nucleoside **13**. After protection of the free 2′-OH group with *tert*-butyldimethylsilyl (TBS) (**14**), a Negishi coupling with 2-pyridylzincbromide afforded the pyridylpurine chelate **15**, followed by TBAF deprotection to **16** and conversion to the phosphoramidite (*S*)-**P**. Starting from (*R*)-(+)-glycidol, (*R*)-**P** was synthesized in five steps with an overall yield of 17% compared with the analogous 2′-deoxynucleotide for which a six step synthesis starting with 2′-deoxy-D-ribose with an overall yield of 26% was reported.[34]

Metal Ion Dependent Duplex Formation. The chelating nucleotides were introduced in the middle position of a 15mer duplex (Figure 6), and the metal ion-dependent duplex stabilities were investigated by UV-melting.[32] The pyridylpu-

SCHEME 4. Synthesis of the Phosphoramidites (*S*)-**H** and (*S*)-**P** for Solid Phase Metallo-GNA Oligonucleotide Synthesis[a]



[a] DPC = diphenylcarbamoyl; DMTr = dimethoxytrityl. See ref 32 for more details.

rine homobase pair P:P in the context of GNA is most strongly stabilized by Ni^{2+} ions ($\Delta T_m = +17.9$ °C with 2 equiv of $NiCl_2$) Figure 6B, displaying a metal ion preference order of $Ni^{2+} > Ag^{+} > Cu^{2+} \approx Co^{2+}$. The Ni^{2+}-induced stabilization of the pyridylpurine base pair in GNA is comparable to the stabilization of this base pair in the context of a DNA duplex reported by Switzer et al., whereas the order of preference for other metal ions is modulated.[34]

Whereas the pyridylpurine base pair displays a preference for Ni^{2+} ions, the hydroxypyridone base pair is stabilized most strongly with Cu^{2+} ions. The addition of 2 equiv of $CuSO_4$ results in an increase of the T_m value by 33.2 °C (Figure 6A). With a T_m value of 70.2 °C, the copper(II)-containing H:H base pair is 19.7 °C more stable than an A:T Watson−Crick base pair at the same position. No significant stabilization can be observed with $Pd(NO_3)_2$, $AgNO_3$, $AuCl_3$, $Cd(NO_3)_2$ or $Ni(NO_3)_2$, but $Co(NO_3)_2$ and $ZnCl_2$ stabilize the H:H base pair by 5.1 and 15.9 °C, respectively. The copper(II)-induced stabilization of the hydroxypyridone base pair in GNA is surprisingly strong and significantly exceeds the reported stabilization for the analogous 2′-deoxynucleotide hydroxypyridone base pair in DNA for which for a stabilization of only 13.1 °C was reported in a 15mer duplex.[33] This indicates that the Cu^{2+}-mediated H:H base pair is particularly well accommodated in the GNA duplex compared with DNA.

We also investigated the metal-mediated cross-pairing of the hydroxypyridone and pyridylpurine chelates (Figure 6C). Overall, this heterobase pair H:P behaves similarly to the hydroxypyridone homobase pair H:H in its metal selectivities. In fact, the stabilization of H:P by Cu^{2+} surpasses the copper(II)-induced stabilization of H:H, affording a duplex with a T_m value of 74.4 °C. Thus, H:P is 23.9 °C more stable than an A:T base pair at the same position in the presence of Cu^{2+} ions. This higher stability can be rationalized by the higher hydrophobicity of P compared with H, whereas the steric problems within metal-coordinated P:P are absent in H:P. Cu^{2+}-mediated H:P represents to date the most stable base pair known for GNA.

Structure of the Copper-Mediated Hydroxypyridone Base Pair in GNA. We determined the 1.3 Å crystal structure of a metallo-GNA duplex formed from the Cu^{2+}-mediated pairing of the self-complementary hydroxypyridone-nucleobase-containing strand 3′-CGHATHCG-2′.[35] This is only the second reported nucleic acid duplex structure containing a designed metal-ion-mediated base pair.[36] Conveniently, with this method site-selectively introduced Cu^{2+} ions served as handles for phasing the crystallographic data.[35,36]

Within the crystal, individual duplexes are coaxially stacked, thereby forming a continuous helix with a periodic pattern of Cu^{2+} ions along the helix axis with Cu−Cu distances of 14.5 Å within a duplex and 23.7 Å between adjacent duplexes (Figure 7A,B). Within the metallo-base pair, a copper ion is coordinated by two hydroxypyridone nucleobases in an almost perfectly square planar fashion (Figure 7D) with

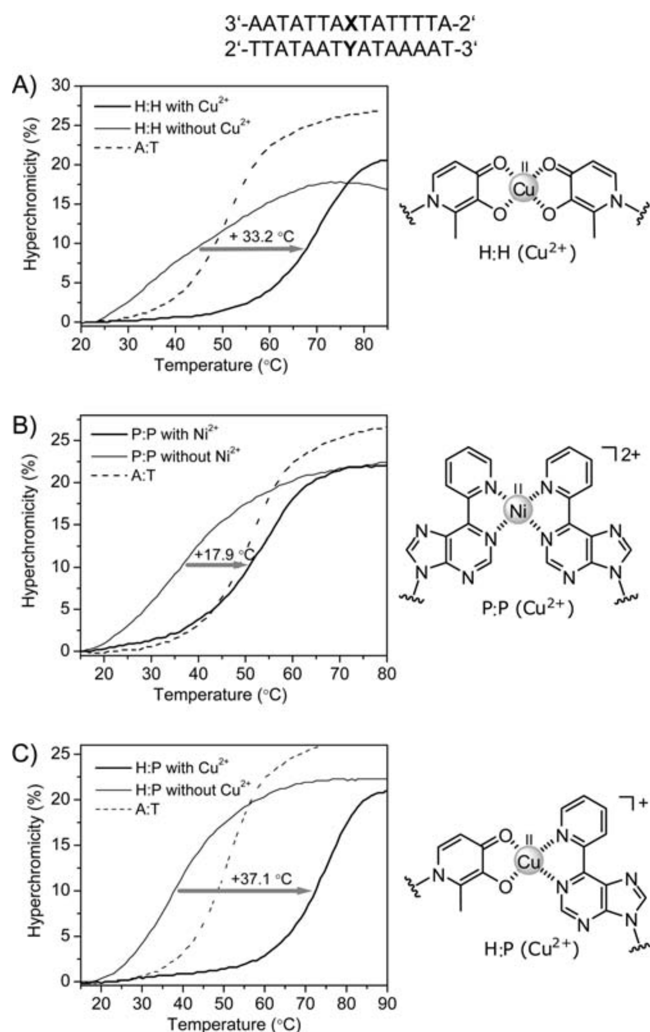


FIGURE 6. Metal-dependent stabilities of GNA duplexes containing hydroxypyridone (H) and pyridylpurine (P) homo- and heterobase pairs determined by UV melting: (A) Cu^{2+}-dependence of H:H base pairing; (B) Ni^{2+}-dependence of P:P base pairing; (C) Cu^{2+}-dependence of H:P base pairing (X = H, Y = P). Measurements were performed in 10 mM sodium phosphate, 100 mM $NaNO_3$, pH = 7.0, with 2 μM of each 15mer GNA strand. $CuSO_4$ or $NiCl_2$ were added to a concentration of 4 μM. See ref 32 for more details.

only a slight propeller twist of around 15° between the two chelating ligands. Coordinative bonds between the copper and oxygen ligands vary between 1.8 and 2.0 Å.

An important common feature of this metallo-GNA and the described brominated GNA duplex composed entirely out of Watson–Crick base pairs is the large backbone–base inclination resulting in a large slide between neighboring base pairs (Figure 7A,B). However, the metallo-GNA structure shows two significant differences from the brominated GNA duplex. First, the duplex is more elongated and less twisted along the *z*-axis with a helical pitch of 60 Å and 16 residues per turn (Figure 4, Table 3). Second, in the metallo-GNA duplex structure, all nucleotides of Watson–Crick base pairs adopt a *gauche* conformation with respect to the torsional angles between C2′–O and C3′–O, whereas the hydroxypyridone glycol nucleotides adopt *anti* conformations (Figure 7C). This reveals that the GNA backbone can adjust its conformation as a function of sequence and the nature of the base pairs. In fact, the C1′–C1′ distance in the Cu^{2+}-mediated H:H base pair is 12.7 Å and thus exceeds the analogous C1′–C1′ distance of the natural A:T and G:C base pairs by around 2.0 Å (Figure 7E). The GNA backbone appears to accommodate this extended size of the metallo-base pair by assuming the *anti* conformation of the vicinal C–O bonds in order to reduce the distances between the backbone oxygen atoms within the hydroxypyridone base pair.

Conclusions and Future Prospects

We have herein summarized the state-of-the-art regarding synthesis, derivatization,[37,38] and duplex formation properties of the simplified nucleic acid GNA. Propylene glycol nucleosides, with the backbone moiety being composed of just three carbon atoms and one stereocenter, probably constitute the simplest possible building blocks for a chemically stable phosphodiester-bond-containing nucleic acid backbone. Due to its simplicity and the synthetic accessibility of glycol nucleotides, in combination with the high duplex stability of GNA duplexes, GNA should comprise a promising scaffold for future structural and functional nucleic acid nanotechnology. In this respect, it is also appealing that both enantiomers of GNA are synthetically accessible, which will allow synthesis of two- and three-dimensional architectures that are mirror-imaged and thus will display different properties in their interaction with biological macromolecules and other chiral compounds. In fact, in a first step toward this direction, Chaput and co-workers recently reported mirror-imaged four-helix junctions composed of (*R*)- and (*S*)-GNA.[39]

In a different but not unrelated direction, it will be exciting to investigate the ability of GNA to form sequence-dependent functional oligomers, similar to aptamers and catalysts composed of RNA and DNA. For this, it will be necessary to find a reliable and efficient way to amplify GNA in order to apply the technique of *in vitro* evolution. Along these lines, McLaughlin and Szostak recently developed the synthesis of GNA nucleoside triphosphates as polymerase substrates[40] and Szostak and co-workers discovered that *Bst* DNA polymerase can catalyze full-length DNA synthesis on a dodecamer GNA template, surprisingly, even in the absence of a stable duplex formation between GNA and DNA.[23]

Last but not least, the structural simplicity of GNA, in combination with a high stability of antiparallel GNA duplex for-

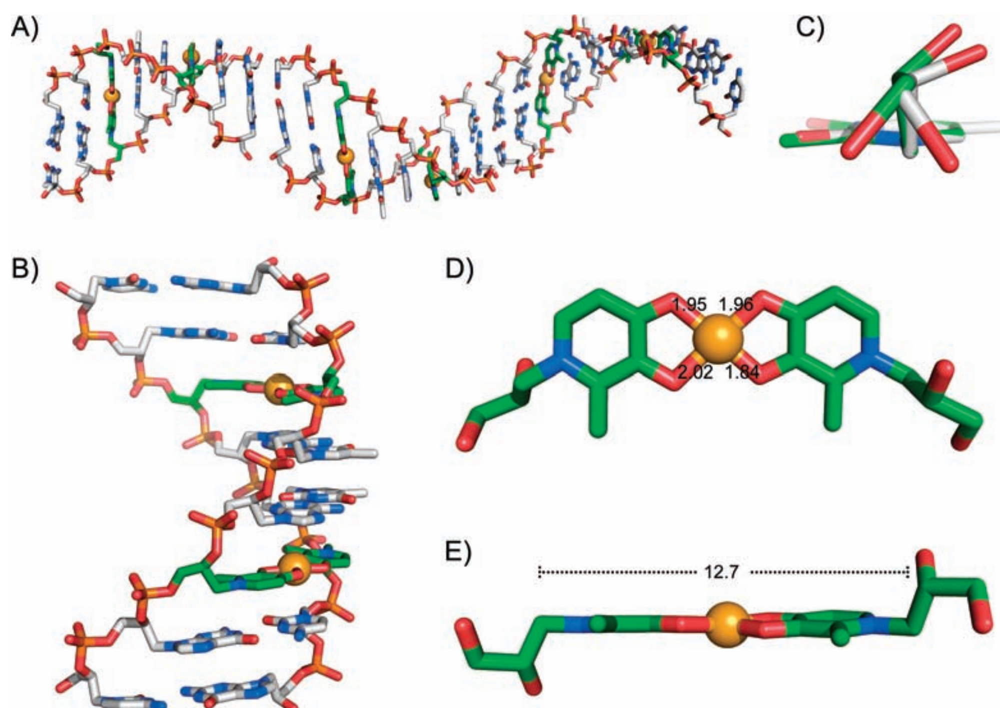


FIGURE 7. Crystal structure of an 8mer GNA duplex from the self-complementary strand 3′-CGHATHCG-2′ in the presence of copper(II). The metallo-base pairs are shown in green: (A) coaxially stacked 8mer duplexes form a continuous helix in the crystal with defined pattern of copper(II) ions; (B) an individual 8mer duplex; (C) superimposed hydroxypyridone and thymine nucleoside from the crystal structure indicating the differences in the conformation along C2′−C3′; (D, E) copper(II) coordination to hydroxypyridone nucleobases of opposite strands. Indicated are the coordinative bond length and the C1′−C1′ distance (in Å). Reproduced form ref 32 with permission of the Royal Society of Chemistry.

mation and high fidelity Watson−Crick base pairing, renders GNA a potential initial genetic material during the origin of life on the way to the transient proposed "RNA world".[41] Starting with glycerol, GNA monomers are directly synthetically accessible in one step by Lewis acid or proton acid catalyzed substitution of a primary OH group of glycerol by nucleobases or through first conversion of glycerol to glycidol, followed by the reaction with nucleobases under basic conditions. Glycerol has been found in meteorites in amounts comparable to simple amino acids, and it therefore must have been present on the early Earth and thus available for incorporation into the first molecules of life.[42] It seems thus very likely that GNA monomers were formed under prebiotic conditions. Activated GNA nucleotides or cyclic GNA phosphates[43] could then have oligomerized nonenzymatically, maybe even reversibly, to form the first GNA oligonucleotide pools from which a dynamic selection could have evolved initial functional GNA molecules used for information storage and catalysis. In first successful efforts toward reading information from GNA templates nonenzymatically, Szostak and co-workers recently demonstrated that a GNA analogue with N2′ → P3′ phosphoramidite linkages (npGNA) can be formed by GNA-template-directed ligations of 3′-imidazole-activated-2′-amino GNA dinucleotides.[44]

With this Account, we intended to share our fascination with the unique combination of simplicity and excellent duplex formation properties of GNA that will hopefully encourage more researchers to start working with this simplified nucleic acid as a molecular scaffold in biotechnology and nanotechnology and to investigate its potential involvement during the evolution of life on Earth.

The authors would like to thank present and past laboratory members for their contributions to this project. We are also especially grateful to our collaborator Prof. Dr. L.-O. Essen (Philipps-University Marburg) for his important contributions regarding the crystal structure determination of GNA duplexes.

BIOGRAPHICAL INFORMATION

Eric Meggers is Professor and Chair of Chemical Biology at the Chemistry Department of the Philipps-University Marburg, Germany.

Lilu Zhang is Junior-Professor at the Chemistry Department of the Philipps-University Marburg, Germany.

FOOTNOTES

*To whom correspondence should be addressed. E-mail: meggers@chemie.uni-marburg.de.

REFERENCES

1 Seeman, N. C. DNA in a Material World. *Nature* **2003**, *421*, 427–431.

2 Gothelf, K. V.; LaBean, T. H. DNA-Programmed Assembly of Nanostructures. *Org. Biomol. Chem.* **2005**, *3*, 4023–4037.

3 GNA has also been referred to as “glycerol nucleic acid” since the nucleoside building blocks are formally derived from glycerol by replacing a primary OH group for a nucleobase. However, we named it glycol nucleic acid in order to emphasize the unusual glycolic (vicinal) arrangement of the two nucleosidic C−O bonds, which leads to a simplified and also shortened backbone.

4 Zhang, L.; Peritz, A.; Meggers, E. A Simple Glycol Nucleic Acid. *J. Am. Chem. Soc.* **2005**, *127*, 4174–4175.

5 Ueda, N.; Kawabata, T.; Takemoto, K. Synthesis of *N*-(2,3-Dihydroxylpropyl) Derivatives of Nucleic Bases. *J. Heterocycl. Chem.* **1971**, *8*, 827–829.

6 Seita, T.; Yamauchi, K.; Kinoshita, M.; Imoto, M. The Synthesis of Nucleoside and Nucleotide Analogs. *Bull. Chem. Soc. Jpn.* **1972**, *45*, 926–928.

7 Seita, T.; Yamauchi, K.; Kinoshita, M.; Imoto, M. Condensation Polymerization of Nucleotide Analogues. *Makromol. Chem.* **1972**, *154*, 255–261.

8 Holý, A.; Ivanova, G. S. Aliphatic Analogues of Nucleotides: Synthesis and Affinity towards Nucleases. *Nucleic Acids Res.* **1974**, *1*, 19–34.

9 Holý, A. Aliphatic Analogues of Nucleosides, Nucleotides, and Oligonucleotides. *Collect. Czech. Chem. Commun.* **1975**, *40*, 187–214.

10 Holý, A. Synthesis of Some 2,3-Dihydroxypropyl Derivatives of Purine Bases. *Collect. Czech. Chem. Commun.* **1978**, *43*, 3103–3117.

11 Holý, A.; Rosenberg, I.; Dvořáková, H. Synthesis of (3-Hydroxy-2-Phosphonylmethoxypropyl) Derivatives of Heterocyclic Bases. *Collect. Czech. Chem. Commun.* **1989**, *54*, 2470–2485.

12 Cook, P. D.; Acevedo, O. L.; Davis, P. W.; Ecker, D. J.; Normand, H. Phosphate Linked Oligomers. U.S. Patent 5,886,177, 1999.

13 Acevedo, O. L.; Andrews, R. S. Synthesis of Propane-2,3-diol Combinatorial Monomers. *Tetrahedron Lett.* **1996**, *37*, 3931–3934.

14 Davis, P. W.; Vickers, T. A.; Wilson-Lingardo, L.; Wyatt, J. R.; Guinosso, C. J.; Sanghvi, Y. S.; DeBaets, E. A.; Acevedo, O. L.; Cook, P. D.; Ecker, D. J. Drug Leads from Combinatorial Phosphodiester Libraries. *J. Med. Chem.* **1995**, *38*, 4363–4366.

15 Nielsen, P.; Dreiøe, L. H.; Wengel, J. Synthesis and Evaluation of Oligodeoxynucleotides Containing Acyclic Nucleosides: Introduction of Three Novel Analogues and a Summary. *Bioorg. Med. Chem.* **1995**, *3*, 19–28.

16 Related aminopropyl nucleosides incorporated into DNA also led to a reduction of thermal duplex stability: Zhou, D.; Lagoja, I. M.; Rozenski, J.; Busson, R.; Van Aerschot, A.; Herdewijn, P. Synthesis and Properties of Aminopropyl Nucleic Acids. *ChemBioChem* **2005**, *6*, 2298–2304.

17 Leumann, C. J. DNA Analogues: From Supramolecular Principles to Biological Properties. *Bioorg. Med. Chem.* **2002**, *10*, 841–854.

18 Eschenmoser, A. Chemical Etiology of Nucleic Acid Structure. *Science* **1999**, *284*, 2118–2124.

19 Schöning, K.-U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. Chemical Etiology of Nucleic Acid Structure: The α-Threofuranosyl-(3′→2′) Oligonucleotide System. *Science* **2001**, *290*, 1347–1351.

20 Wippo, H.; Reck, F.; Kudick, R.; Ramaseshan, M.; Ceulemans, G.; Bolli, M.; Krishnamurthy, R.; Eschenmoser, A. Pentopyranosyl Oligonucleotide Systems. Part 11: Systems with Shortened Backbones: (D)-β-Ribopyranosyl-(4′→3′)- and (L)-α-Lyxopyranosyl-(4′→3′)-oligonucleotides. *Bioorg. Med. Chem.* **2001**, *9*, 2411–2428.

21 Zhang, L.; Peritz, A. E.; Carroll, P. J.; Meggers, E. Synthesis of Glycol Nucleic Acids. *Synthesis* **2006**, 645–653.

22 Schlegel, M. K.; Meggers, E. Improved Phosphoramidite Building Blocks for the Synthesis of the Simplified Nucleic Acid GNA. *J. Org. Chem.* **2009**, *74*, 4615–4618.

23 Tsai, C.-H.; Chen, J.; Szostak, J. W. Enzymatic Synthesis of DNA on Glycerol Nucleic Acid Templates without Stable Duplex Formation between Product and Template. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 14598–14603.

24 Schlegel, M. K.; Peritz, A. E.; Kittigowittana, K.; Zhang, L.; Meggers, E. Duplex Formation of the Simplified Nucleic Acid GNA. *ChemBioChem* **2007**, *8*, 927–932.

25 Schlegel, M. K.; Xie, X.; Zhang, L.; Meggers, E. Insight into the High Duplex Stability of the Simplified Nucleic Acid GNA. *Angew. Chem., Int. Ed.* **2009**, *48*, 960–963.

26 Yang, Y.-W.; Zhang, S.; McCullum, E. O.; Chaput, J. C. Experimental Evidence that GNA and TNA Were Not Sequential Polymers in the Prebiotic Evolution of RNA. *J. Mol. Evol.* **2007**, *65*, 289–295.

27 Schlegel, M. K.; Essen, L.-O.; Meggers, E. Atomic Resolution Duplex Structure of the Simplified Nucleic Acid GNA. *Chem. Commun.* **2010**, *46*, 1094–1096.

28 Schlönvogt, I.; Pitsch, S.; Lesueur, C.; Eschenmoser, A.; Jaun, B. Pyranosyl-RNA (‘p-RNA’): NMR and Molecular-Dynamics Study of the Duplex Formed by Self-pairing of Ribopyranosyl-(C-G-A-A-T-T-C-G). *Helv. Chim. Acta* **1996**, *79*, 2316–2345.

29 Micura, R.; Kudick, R.; Pitsch, S.; Eschenmoser, A. Opposite Orientation of Backbone Inclination in Pyrasonyl-RNA and Homo-DNA Correlates with Opposite Directionality of Duplex Properties. *Angew. Chem., Int. Ed.* **1999**, *38*, 680–683.

30 Egli, M.; Pallan, P. S.; Pattanayek, R.; Wilds, C. J.; Lubini, P.; Minasov, G.; Dobler, M.; Leumann, C. J.; Eschenmoser, A. Crystal Structure of Homo-DNA and Nature’s Choice of Pentose over Hexose in the Genetic System. *J. Am. Chem. Soc.* **2006**, *128*, 10847–10856.

31 Declercq, R.; Van Aerschot, A.; Read, R. J.; Herdewijn, P.; Van Meervelt, L. Crystal Structure of Double Helical Hexitol Nucleic Acids. *J. Am. Chem. Soc.* **2002**, *124*, 928–933.

32 Schlegel, M. K.; Zhang, L.; Pagano, N.; Meggers, E. Metal-Mediated Base Pairing within the Simplified Nucleic Acid GNA. *Org. Biomol. Chem* **2009**, *7*, 476–482.

33 Tanaka, K.; Tengeiji, A.; Kato, T.; Toyama, N.; Shiro, M.; Shionoya, M. Efficient Incorporation of a Copper Hydroxypyridone Base Pair in DNA. *J. Am. Chem. Soc.* **2002**, *124*, 12494–12498.

34 Switzer, C.; Sinha, S.; Kim, P. H.; Heuberger, B. D. A Purine-like Nickel(II) Base Pair for DNA. *Angew. Chem., Int. Ed.* **2005**, *44*, 1529–1532.

35 Schlegel, M. K.; Essen, L.-O.; Meggers, E. Duplex Structure of a Minimal Nucleic Acid. *J. Am. Chem. Soc.* **2008**, *130*, 8158–8159.

36 Atwell, S.; Meggers, E.; Spraggon, G.; Schultz, P. G. Structure of a Copper-Mediated Base Pair in DNA. *J. Am. Chem. Soc.* **2001**, *123*, 12364–12367.

37 For a recent example of derivatized GNA, see also: Zhou, H.; Ma, X.; Wang, J.; Zhang, L. Pyrene Acetylide Nucleotides in GNA: Probing Duplex Formation and Sensing of Copper(II) Ions. *Org. Biomol. Chem.* **2009**, *7*, 2297–2302.

38 For the incorporation of modified GNA building blocks into DNA, see, for example: Baumstark, D.; Wagenknecht, H.-A. Fluorescent Hydrophobic Zippers inside Duplex DNA: Interstrand Stacking of Perylene-3,4:9,10-tetracarboxylic Acid Bisimides as Artificial DNA Base Dyes. *Chem.—Eur. J.* **2008**, *14*, 6640–6645.

39 Zhang, R. S.; McCullum, E. O.; Chaput, J. C. Synthesis of Two Mirror Image 4-Helix Junctions from Glycerol Nucleic Acid. *J. Am. Chem. Soc.* **2006**, *130*, 5846–5847.

40 Horhota, A. T.; Szostak, J. W.; McLaughlin, L. W. Glycerol Nucleoside Triphosphates: Synthesis and Polymerase Substrate Activities. *Org. Lett.* **2006**, *8*, 5345–5347.

41 Orgel, L. E. Prebiotic Chemistry and the Origin of the RNA World. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 99–123.

42 Cooper, G.; Kimmich, N.; Belisle, W.; Sarinana, J.; Brabham, K.; Garrel, L. Carbonaceous Meteorites As a Source of Sugar-Related Organic Compounds for the Early Earth. *Nature* **2001**, *414*, 879–883.

43 For the template-directed oligomerization of cyclic phosphates, see: Uesugi, S.; Ikehara, M. Synthesis and Template-Directed Polymerization of Adenylyl(3′-5′)adenosine Cyclic 2′,3′-Phosphate. *Biochemistry* **1977**, *16*, 493–498.

44 Chen, J. J.; Cai, X.; Szostak, J. W. N2′→P3′ Phosphoramidite Glycerol Nucleic Acid as a Potential Alternative Genetic System. *J. Am. Chem. Soc.* **2008**, *131*, 2119–2121.