Modified Nucleic Acids

DOI: 10.1002/anie.201300795

Base-Pairing Properties of a Structural Isomer of Glycerol Nucleic Acid**

*Phaneendrasai Karri, Venkateshwarlu Punna, Keunsoo Kim, and Ramanarayanan Krishnamurthy**

The investigation of constitutionally simple oligomeric information systems is needed from the viewpoint of knowing how structurally simple, and minimal, an oligomeric system can become and still be functional (with respect to information storage and processing). Herein, we report on the synthesis and investigation of the base-pairing properties of one such system, a glycerol-(3′→1′)-linked oligonucleotide.

Our interest in minimal, oligomeric information systems was motivated by the experimental mapping of potentially primordial oligomeric information systems,[1] which contain backbones, recognition elements, and linker groups that are structurally quite different from conventional nucleic acids, an approach pioneered by Eschenmoser and co-workers.[2] In this pursuit, we used a "qualitative conformational analysis"[3] approach—a simple analysis of the capability of the oligomeric system to adopt regular and repetitive conformations, which is a prerequisite for base-pairing. Such a method can also be used to design novel synthetic oligomeric structures.[3] Employing this approach, we derived from the A-form of DNA and RNA, a glycerol based acyclic-(3′→1′)-linked oligonucleotide system that is tagged at the 2′-position with canonical nucleobases (Figure 1).

This acyclic oligomer (Figure 2c) shows a resemblance to the first acyclic system—glycerol-derived acyclic nucleosides—which was considered by Joyce et al.[4a] when making the case for an ancestral genetic system involving simpler acyclic analogues of RNA and DNA (Figure 2a). Subsequently, this acyclic system (FNA) was shown to have very weak base-pairing abilities.[5] Later, Meggers and co-workers[6] showed that GNA,[7] in spite of having an acyclic backbone, had strong base-pairing properties (Figure 2b). This led to further study of acyclic nucleic acids.[8] Our glycerol-based system, depicted in Figure 1 and Figure 2, is a structural isomer of GNA (isoGNA)—with the nucleobase at the 2′-position (instead of the 1′-position as in GNA) and has 3′→1′-linked phosphodiester bonds (instead of the 3′→2′ linkages in GNA). When compared to FNA,[4,5] our system lacks the oxymethylene linker between the nucleobase and the backbone—placing the recognition element directly on the backbone; however at the monomeric level, the backbone units of both systems are achiral, unlike GNA. Therefore, we investigated the base-pairing potential of isoGNA, which we considered to have a structure that is in between that of FNA and GNA.

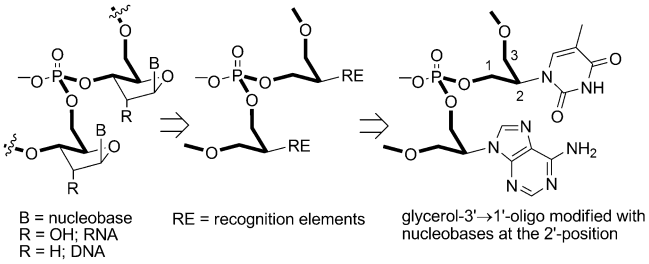

Figure 1. Derivation of the 3′→1′-glycerolphosphate linked oligonucleotide based on qualitative conformational analysis[3] of an idealized A-form RNA/DNA backbone.

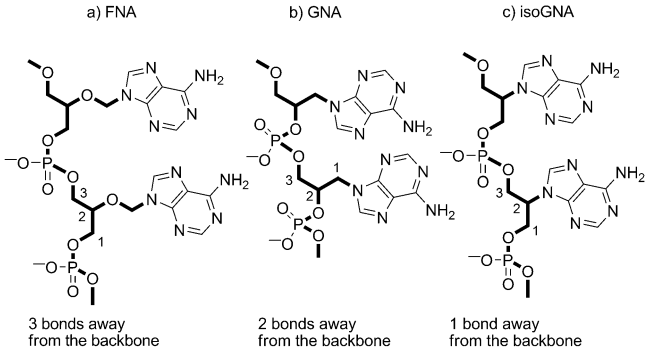

Figure 2. Comparison of the three, flexible, acyclic glycerol-backbone-based oligonucleotides.

The appropriate isoGNA phosphoramidites **6a–d** (Figure 3) were synthesized starting from the commercially available (*s*)-solketal (Supporting Information, Scheme S1). The synthesis of isoGNA oligonucleotides followed standard methods from the literature with some adjustments (Table S3). IsoGNA (^{i}g) oligomers with one end-protecting group seemed to be stable under the conditions investigated, and all our sequences had a 1′-O- (and/or 3′-O-) cap, either in the form of a phosphate or a DNA or RNA nucleotide. Base-pairing properties of the isoGNA sequences (Table S4) were investigated by temperature dependent UV absorption and circular dichroism (CD) spectroscopy in phosphate buffer

[*] Dr. P. Karri, Dr. V. Punna, Dr. K. Kim, Prof. Dr. R. Krishnamurthy
Department of Chemistry, The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)
E-mail: rkrishna@scripps.edu
Homepage: http://www.scripps.edu/krishnamurthy/

[**] This work was supported by NASA Astrobiology: Exobiology and Evolutionary Biology Program (Grant NNX09AM96G), and was jointly supported by NSF and NASA Astrobiology Program under the NSF Center for Chemical Evolution, Grant CHE-1004570. We thank Professors A. Eschenmoser and G. Joyce, and Dr. Matthias Stoop for constructive feedback on the manuscript.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201300795.

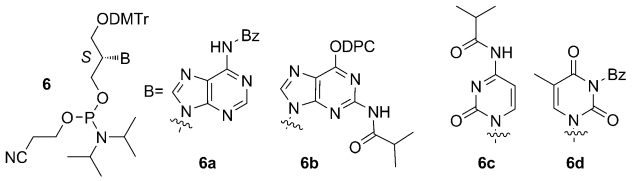

Figure 3. The nucleobase-modified glycerol phosphoroamidites (**6**) used in our study. DMTr = dimethoxytrityl; B = nucleobase; DPC = diphenylcarbamoyl.

(pH 7.0, 1 M NaCl). Initially, we studied $^{i}g(T)_{12}$ and observed no cooperative melting behavior with corresponding complementary RNA or DNA sequences. Therefore, we decided to investigate the properties of hexadecamers of isoGNA.

No self-association was found for $^{i}g(T)_{16}$; not unexpectedly, $^{i}g(A)_{16}$ showed weak self-pairing,[9] which was supplanted by reasonably strong heteroduplex formation in the presence of complementary $^{i}g(T)_{16}$ (Figure 4; Figure S3).

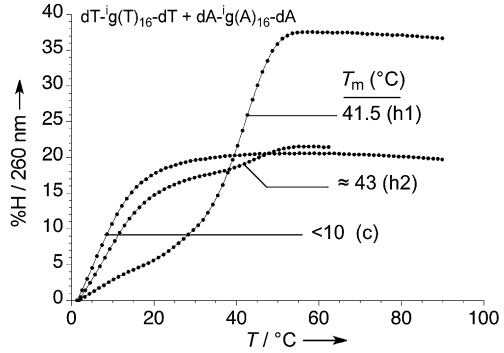

Figure 4. Thermal melting curves of duplexes formed from dA-$^{i}g(A)_{16}$-dA and dT-$^{i}g(T)_{16}$-dT oligonucleotides measured by UV spectroscopy; ^{i}g = isoGNA; d = DNA; h1 = reheating of a sample (after the first measurement) kept at 4 °C for 6 days; c = cooling of the same sample; h2 = second heating cycle (after waiting at 4 °C for 1 min); %H = percent hyperchromicity. Conditions: 5 μM of each oligonucleotide, 1 M NaCl, 10 mM Na_2HPO_4, 0.1 mM Na_2EDTA, pH 7.0.

However, the hybridization behavior showed strong dependence on time. A freshly prepared sample showed triplex-like behavior. The lower transition was attributed to self-association of $^{i}g(A)_{16}$—a kinetically preferred process over the thermodynamically more stable mixed duplex $^{i}g(A)_{16}/^{i}g(T)_{16}$. This was confirmed when the sample was allowed to equilibrate at 4 °C for six days; only a single melting-temperature (T_m) transition (h1 in Figure 4) was seen for the duplex between $^{i}g(A)_{16}$ and $^{i}g(T)_{16}$, indicating that the kinetics of its formation is relatively slow. Variation in the T_m values with a change in the oligomer and salt concentrations or replacement of NaCl with $MgCl_2$, corroborated the formation of interstrand duplexes. The corresponding temperature dependent CD-spectrum showed weak signals indicating that the duplex has little secondary structure (Figure S3). This CD-behavior (and duplex stability) was independent of the nature of the final capping residue on the isoGNA oligomer (RNA, DNA, or free OH group).

Surprisingly, the heterobase-containing sequences (isoGNA-(A,T)) showed no base-pairing. Both alternating sequences $^{i}g(TA)_8$ and $^{i}g(T_8A_8)$ showed no sigmoidal behavior in their UV-melting curves. Moreover, no duplex formation was found for any of the irregular mixed isoGNA-(A,T) 16-mer sequences investigated (Entries 10–11 in Table S4). Mixed 8-mer as well as 16-mer sequences containing A, T, G, and C residues also showed no base-pairing (UV- and CD-measurements). Such contrasting behavior where the homogeneous sequences show base-pairing but the heterogeneous sequences do not is unprecedented within a given backbone system.

The base-pairing behavior of isoGNA with the corresponding complementary RNA and DNA sequences was, once again, in contrast with GNA. For example, the homopurinic $^{i}g(A)_{16}$ oligomer base-paired with both complementary DNA and RNA sequences (Figure 5; Entries 1–6 in Table S5). This is in contrast to GNA, where there is interaction with RNA but not with DNA. In the isoGNA series, the RNA backbone also seems to be preferred over a DNA backbone (Figure 5a). A 1:1 stoichiometry of duplex formation was indicated by a job plot (Figure S5), although the formation of a higher order complex (2:1) was also found during the mixing experiment. There is almost no hysteresis for these pairings with DNA or RNA, contrary to the curves from self-pairing of isoGNA (Figure S4). This indicates that the nature of the phosphate backbone dictates the kinetics of duplex formation. The inverse combination, the dT-$^{i}g(T)_{16}$-dT with d-$(A)_{18}$ (or poly-dA) showed no duplex formation (as in

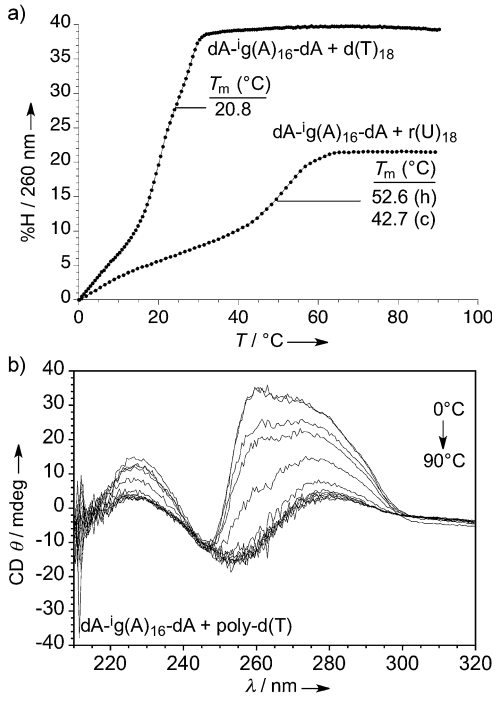

Figure 5. Base-pairing behavior of a duplex formed by dA-$^{i}g(A)_{16}$-dA with complementary DNA (d = DNA) and RNA (r = RNA). a) Thermal melting curves; for conditions see footnote of Table S5. b) Temperature dependent CD spectra; 5 μM of each oligonucleotide, 1 M NaCl, 10 mM Na_2HPO_4, 0.1 mM Na_2EDTA, pH 7.0.

FNA).[5] Once again, none of the heterogeneous sequences of isoGNA, containing any combination of A, T, G, and C residues, formed a duplex with the corresponding complementary RNA or DNA sequences.

The fact that $^{i}g(T)_n$ sequences pair with $^{i}g(A)_n$ but not with complementary homopurinic sequences in the DNA and RNA series, suggested that the isoGNA backbone may have a preference for acyclic backbones, a phenomenon that could be related to the flexibility of the backbone and inclination between base-pairs on the backbone.[1d,10,11] Therefore, we investigated the duplex formation of isoGNA sequences with complementary alternative nucleobases on various acyclic dipeptide backbones that were investigated previously in our laboratory[1a–c] (Figure 6; Entries 10–16 in Table S5).

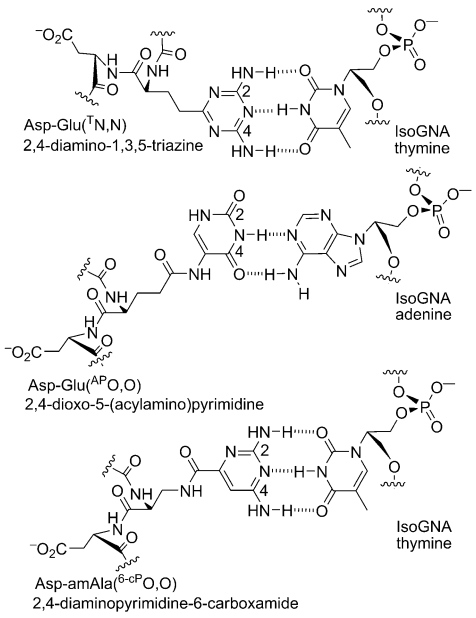

Figure 6. Inter-system base-pairing of $^{i}g(A)_{16}$ and $^{i}g(T)_{16}$ with 2,4-diamino-1,3,5-triazine, 2,4-dioxo-5-(acylamino)pyrimidine and 2,4-diaminopyrimidine-6-carboxamide on acyclic dipeptide backbones.

Moderate to strong duplex formation was observed when $^{i}g(T)_n$ was mixed with the corresponding complementary 12- and 16-mer 2,4-diamino-1,3,5-triazines (^{T}NN) on a peptide Asp-Glu backbone (Figure 7a). A clear dependence on the length of the oligomer in the T_m confirmed an interstrand association. The fact that even a 12-mer oligonucleotide showed duplex formation—while this was not seen at the same length for isoGNA with DNA or RNA—suggests the important roles the nature of backbone, the orientation of the hydrogen-bonding face, and charges in the backbone play in determining ability to form a duplex.[12] Weak duplex formation between $^{i}g(T)_{16}$ and a 16-mer of 2,4-diaminopyrimidine-6-carboxamide on a Asp-amAla backbone was observed (Table S5). In the reverse combination, $^{i}g(A)_{16}$ was also found to form a duplex with a 2,4-dioxopyrimidine modified Asp-Glu backbone (Figure 7b). At 260 nm and below, the association of the heteroduplex along with

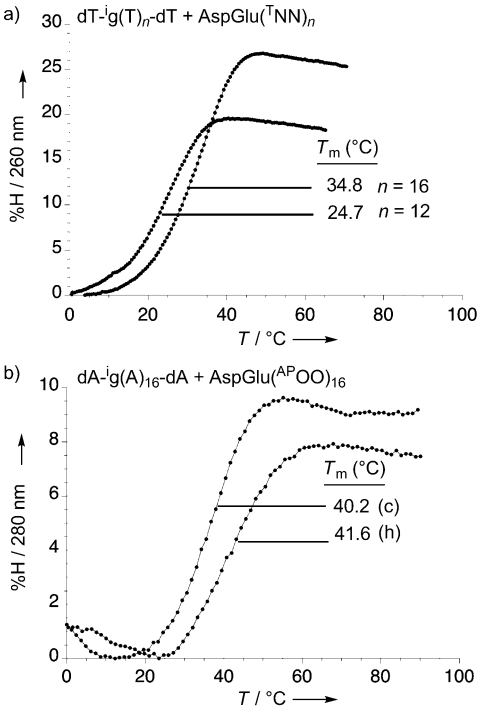

Figure 7. Thermal melting curves showing the duplex formation of a) dT-$^{i}g(T)_{12}$-dT and dT-$^{i}g(T)_{16}$-dT with the corresponding 2,4-diamino-1,3,5-triazine modified AspGlu-dipeptide backbone; b) dA-$^{i}g(A)_{16}$-dA with a 16-mer 2,4-dioxo-5-(acylamino)pyrimidine modified AspGlu-dipeptide backbone at 280 nm. Conditions: 5 μM of each oligonucleotide, 1 M NaCl, 10 mM Na_2HPO_4, 0.1 mM Na_2EDTA, pH 7.0.

interference from the self-association of $^{i}g(A)_{16}$ was observed (Figure S6). However, at 280 nm, only the melting of the duplex could be observed (Figure 7b). In all of these acyclic backbone intra-system pairings, little or no hysteresis was observed, in contrast to the intra-system isoGNA pairings.

The interaction of GNA with isoGNA was examined using two heterogeneous GNA sequences CACATTATTGTTGTA and TACAACAATAATGTG,[13] and no base-pair mediated interaction was seen. No interaction between isoGNA and complementary homogeneous PNA sequences was observed (Table S5), in contrast to acyclic dipeptide backbones with alternative heterocycles (Figure 6 and Figure 7).

The striking difference in the intra- and inter-system base-pairing of isoGNA, in combination with other studies,[4–6,14–16] reinforces the notion that the backbone plays a critical role in fine-tuning the nature of the base-pairing for the same set of canonical nucleobases (found in nature). The acyclic nucleic acid (FNA) system exhibits no intra-system base-pairing and shows very limited duplex formation with RNA and DNA.[5] GNA shows strong self-pairing but limited hybridization with RNA and none with DNA,[6] and, isoGNA self- and cross-pairing seems to be limited to homogeneous base sequences.[17,18] A shift in the position of the nucleobase in the acyclic backbone (from the 1′- to the 2′-position), or variation in the linker length (Figure 2), results in a drastic variation of base-pairing capabilities, even though these systems contain the same canonical nucleobases and closely related backbones.

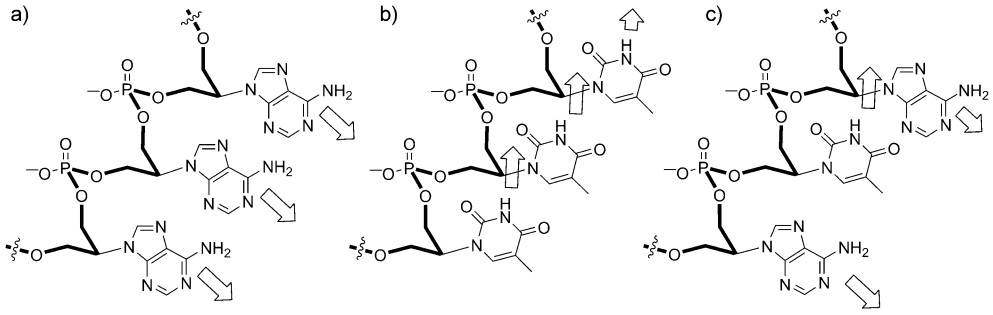

Figure 8. Hypothesis to account for the unexpected dichotomous behavior between the homo- and heterogeneous nucleobase sequences of isoGNA (arrows indicate the Watson–Crick base-pairing face). In the homogeneous sequences, an all *anti*-purine (a) or an all *syn*-pyrimidine (b) arrangement (or the other way around) allows for hydrogen-bond-mediated base-pairing. In a mixed sequence (c), the Watson–Crick faces are pointing in different directions, thus hindering hydrogen-bond mediated associations.[19]

To account for the unexpected lack of base-pairing between heterogeneous isoGNA sequences, we propose that the purine and pyrimidine bases have different *syn*- and *anti*-orientations (or the other way around) with respect to the backbone (Figure 8).[18] Such different orientations would allow the homo-sequences to form base-pairs because the acceptor face of the pyrimidine from one strand would be exposed to the donor face of the purine base. However, if they are in the same (heterogeneous) sequence, then the hydrogen-bonding faces of the pyrimidine and purine bases would be oriented in different directions, and one of the bases would have to flip from the *syn* to the *anti* conformation (or the other way around) to bring together uniform faces for base-pairing.[19] If the energy gained by duplex formation is not greater than the energy needed for such a change in base orientation, then duplex formation will not be favored.

Such loss-of-base-pairing capability is also observed in GNA with mixed sequences when paired with RNA. While GNA A,T-sequences form duplexes with RNA, adding G and C residues to the sequence causes GNA to lose this ability.[6b] These results, when combined with ours, show that the nature of the backbone is critical for determining, and limiting, the recognition elements and the modes of base-pairing that can be mediated by the canonical nucleobases. These observations caution against extrapolating base-pairing expectations simply based on the behavior of homogeneous base sequences, at least for acyclic backbones.

Overall, the base-pairing properties of the isoGNA system stand in stark contrast to GNA and other flexible nucleic acid backbones, necessitating discretion regarding the structural requirements for a simple oligomeric information system containing only the canonical nucleobases. This is relevant to the evolution of information-containing polymers—especially in the context of the origins of life.[4,20] The search for structurally minimal ("simple") informational systems need not be constrained to the set of natural nucleobases as the recognition elements. These canonical nucleobases may represent a "functional optimum"[21] with respect to the phosphate backbone, as the product of evolutionary selection,[22] but this also allows for the possibility that there may be other primordial recognition elements that could function in conjunction with simple, prebiotic backbones.[1] Thus, the search for—and mapping of—primordial information systems could encompass a wider variety of constituents, which are compatible with heterotrophic and/or autotrophic environments.[2,23]

Received: January 29, 2013
Revised: March 19, 2013
Published online: April 18, 2013

Keywords: acyclic nucleic acids · DNA · glycerol nucleic acid · oligonucleotides · RNA

[1] a) G. K. Mittapalli, K. R. Reddy, H. Xiong, O. Munoz, B. Han, F. De Riccardis, R. Krishnamurthy, A. Eschenmoser, *Angew. Chem.* **2007**, *119*, 2522; *Angew. Chem. Int. Ed.* **2007**, *46*, 2470; b) G. K. Mittapalli, Y. M. Osornio, M. Guerrero, K. R. Reddy, R. Krishnamurthy, A. Eschenmoser, *Angew. Chem.* **2007**, *119*, 2530; *Angew. Chem. Int. Ed.* **2007**, *46*, 2478; c) X. Zhang, R. Krishnamurthy, *Angew. Chem.* **2009**, *121*, 8268; *Angew. Chem. Int. Ed.* **2009**, *48*, 8124; d) M. Hernández-Rodríguez, J. Xie, Y. M. Osornio, R. Krishnamurthy, *Chem. Asian J.* **2011**, *6*, 1252.

[2] a) A. Eschenmoser, *Origins Life Evol. Biosphere* **2004**, *34*, 277; b) A. Eschenmoser, *Tetrahedron* **2007**, *63*, 12821.

[3] A. Eschenmoser, *Chimia* **2005**, *59*, 836.

[4] a) G. F. Joyce, A. W. Schwartz, S. L. Miller, L. E. Orgel, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 4398; b) M. Tohidi, L. E. Orgel, *J. Mol. Evol.* **1989**, *28*, 367.

[5] a) K. C. Schneider, S. A. Benner, *J. Am. Chem. Soc.* **1990**, *112*, 453; b) F. Vandendriessche, K. Augustyns, A. Van Aerschot, R. Busson, J. Hoogmartens, P. Herdewijn, *Tetrahedron* **1993**, *49*, 7223; c) P. Nielsen, L. H. Dreiøe, J. Wengel, *Bioorg. Med. Chem.* **1995**, *3*, 19; d) Y. Merle, E. Bonneil, L. Merle, J. Sagi, A. Szemzo, *Int. J. Biol. Macromol.* **1995**, *17*, 239; e) B. M. Dahl, U. Henriksen, O. Dahl, *Org. Biomol. Chem.* **2006**, *4*, 1115.

[6] a) L. Zhang, A. Peritz, E. Meggers, *J. Am. Chem. Soc.* **2005**, *127*, 4174; b) M. K. Schlegel, A. E. Peritz, K. Kittigowittana, L. Zhang, E. Meggers, *ChemBioChem* **2007**, *8*, 927.

[7] GNA was derived from threofuranosyl nucleic acid (TNA): K. U. Schöning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, A. Eschenmoser, *Science* **2000**, *290*, 1347.

[8] S. Zhang, C. Switzer, J. C. Chaput, *Chem. Biodiversity* **2010**, *7*, 245.

[9] For strong homo-adenine self-pairing see: J. Hunziker, H.-J. Roth, M. Böhringer, A. Giger, U. Diedrichsen, M. Göbel, R. Krishnan, B. Jaun, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1993**, *76*, 259.

[10] H. Wippo, F. Reck, R. Kudick, M. Ramaseshan, G. Ceulemans, M. Bolli, R. Krishnamurthy, A. Eschenmoser, *Bioorg. Med. Chem.* **2001**, *9*, 2411.

[11] P. S. Pallan, P. Lubini, M. Bolli, M. Egli, *Nucleic Acids Res.* **2007**, *35*, 6611.

[12] GNA does not base-pair with TNA; see Y.-W. Yang, S. Zhang, E. McCullum, J. C. Chaput, *J. Mol. Evol.* **2007**, *65*, 289.

[13] We are grateful to Prof. E. Meggers for providing these two GNA sequences.

[14] a) P. E. Nielsen, M. Egholm, R. H. Berg, O. Burchardt, *Science* **1991**, *254*, 1497; b) P. E. Nielsen, *Origins Life Evol. Biospheres* **1993**, *23*, 323.
[15] a) H. Asanuma, T. Toda, K. Murayama, X. Liang, H. Kashida, *J. Am. Chem. Soc.* **2010**, *132*, 14702; b) H. Kashida, K. Murayama, T. Toda, H. Asanuma, *Angew. Chem.* **2011**, *123*, 1321; *Angew. Chem. Int. Ed.* **2011**, *50*, 1285.
[16] A. B. Petersen, M. A. Petersen, U. Henriksen, S. Hammerum, O. Dahl, *Org. Biomol. Chem.* **2003**, *1*, 3293.
[17] An isoGNA-$(G)_n$ sequence was not studied owing to anticipated complex associations.
[18] Collaborative studies are underway with Prof. N. Hud at Georgia Institute of Technology (Atlanta) to understand this dichotomous behavior.
[19] Modeling studies suggest that other modes of base-pairing (e.g. Hoogsteen) are disfavored because of the shortening of the interstrand phosphate distance (increased electrostatic repulsion).
[20] a) A. W. Schwartz, L. E. Orgel, *Science* **1985**, *228*, 585; b) A. W. Schwartz, J. Visscher, C. G. Bakker, J. Niessen, *Origins Life Evol. Biospheres* **1987**, *17*, 351; c) J. Visscher, A. W. Schwartz, *J. Mol. Evol.* **1988**, *28*, 3; d) J. C. Chaput, C. Switzer, *J. Mol. Evol.* **2000**, *51*, 464.
[21] R. Krishnamurthy, *Acc. Chem. Res.* **2012**, *45*, 2035.
[22] It has been argued that the nucleobases in DNA have been selected and optimized for DNA, from an ancestral RNA world: A. C. Rios, Y. Tor, *Astrobiology* **2012**, *12*, 884.
[23] For a discussion see pages 12451 – 12453 and 12459 – 12462 in A. Eschenmoser, *Angew. Chem.* **2011**, *123*, 12618; *Angew. Chem. Int. Ed.* **2011**, *50*, 12412.

