Supplementary Material (ESI) for Chemical Communications

Electronic Supplementary Information

Atomic Resolution Duplex Structure of the Simplified Nucleic Acid GNA

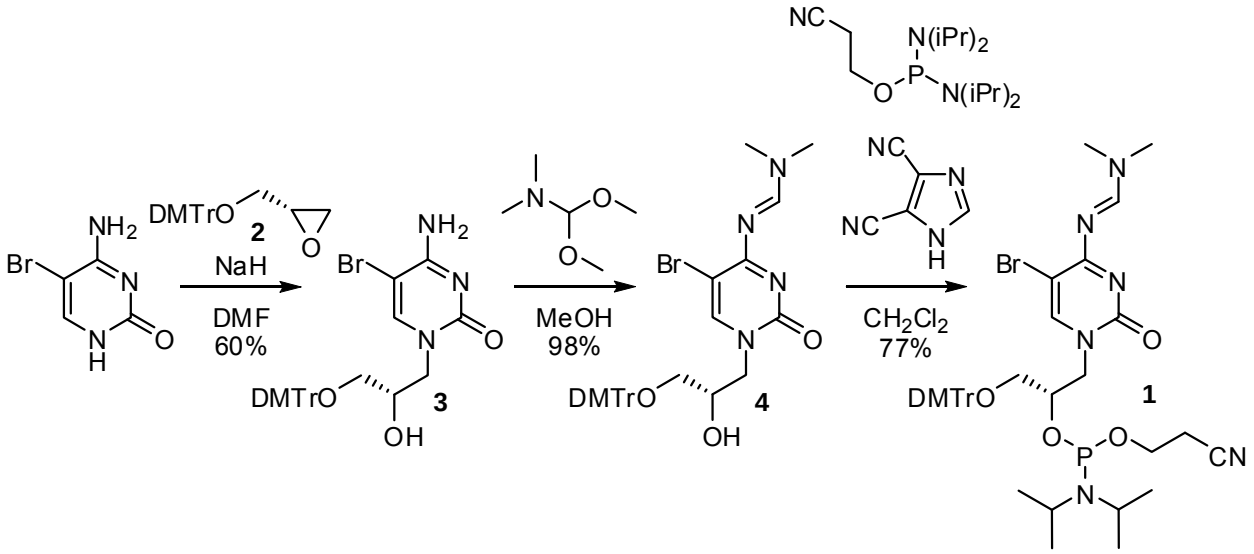
Mark K. Schlegel, Lars-Oliver Essen and Eric Meggers

Fachbereich Chemie, Philipps-Universität Marburg, Hans-Meerwein Strasse, 35032 Marburg, Germany

Content:

1.) Synthesis of the 5-Bromocytosine Phosphoramidite for Automated Solid Phase Synthesis

General Procedures and Reagents. NMR spectra were recorded on a Bruker DMX-300 (300 MHz) spectrometer. High-resolution mass spectra were obtained with a Thermo LTQ-FT instrument using ES ionization. Infrared spectra were recorded on a Bruker Alpha spectrometer. Solvents and reagents were used as supplied from Aldrich, Acros, or Hampton Research. Reactions were performed under an atmosphere of nitrogen unless otherwise specified. Compound **2** was prepared as described previously.[1]



Scheme S1. Synthesis of phosphoramidite **1**.

Compound 3. To a suspension of 5-bromocytosine (1.00 g, 5.26 mmol) in anhydrous DMF (10.5 mL) under argon was added NaH (42 mg, 1.1 mmol, 60% in mineral oil) and the solution was allowed to stir under argon for one hour. A solution of (*S*)-glycidyl 4,4'-dimethoxytrityl ether (**2**) (1.88 g, 4.99 mmol) in DMF (10.5 mL) was added to the first solution and then heated to 100 °C overnight. The next morning, the solution was cooled, all solvent removed, the resulting oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via flash chromatography over silica gel starting with 100:1 EtOAc:Et_3N, then eluting with

50:1:0.01 EtOAc:MeOH:Et_3N to afford compound **3** as a light yellow solid (1.70 g, 60%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ (ppm) 7.87 (br, 1H), 7.52 (s, 1H), 7.41 (m, 2H), 7.29 (m, 6H), 7.21 (m, 1H), 6.83 (m, 4H), 5.64 (br, 1H), 4.24 (dd, *J* = 13.8, 2.5 Hz, 1H), 3.78 (s, 6H), 3.73 (dd, *J* = 14.0, 6.6 Hz, 1H), 3.23 (dd, *J* = 9.7, 5.1 Hz, 1H), 3.02 (dd, *J* = 9.7, 6.4 Hz, 1H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ (ppm) 162.8, 158.7, 156.8, 147.4, 144.8, 135.9, 135.8, 130.0, 128.10, 128.03, 127.1, 113.4, 86.8, 86.6, 69.5, 64.5, 55.4, 54.3. IR (solid) ν (cm^{-1}) = 3442 (br), 2989, 2836, 1679, 1656, 1597, 1444, 1348, 1299, 1246, 1175, 1071, 1031, 948, 827, 727, 701, 622, 582, 417. HRMS calcd for $C_{28}H_{28}N_3O_5BrNa$ $(M+Na)^+$ 588.1105, found $(M+Na)^+$ 588.1115.

Compound 4. To a solution of compound **3** (435 mg, 0.768 mmol) in MeOH (2.5 mL) was added dimethylformamide-dimethylacetal (0.36 mL, 2.7 mmol) and heated to 50 °C for one hour. After cooling and removal of the MeOH, the residue was redissolved in methylene chloride, washed once with water, dried over Na_2SO_4, and finally concentrated. The product was purified via flash chromatography over silica gel starting with 100:1 EtOAc:Et_3N, then eluting with 20:1:0.01 EtOAc:MeOH:Et_3N to afford compound **4** as a white foam (470 mg, 98%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ (ppm) 8.72 (s, 1H), 7.64 (s, 1H), 7.42 (m, 2H), 7.29 (m, 6H), 7.21 (m, 1H), 6.83 (m, 4H), 4.29 (dd, *J* = 13.9, 2.4 Hz, 1H), 4.23 (d, *J* = 4.8 Hz, 1H), 4.13 (br, 1H), 3.89-3.76 (m, 7H), 3.28 (dd, *J* = 9.7, 5.3 Hz, 1H), 3.21 (s, 3H), 3.18 (s, 3H), 2.98 (dd, *J* = 9.6, 6.9 Hz, 1H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ (ppm) 168.1, 159.0, 158.7, 157.7, 147.6, 144.8, 136.0, 135.8, 130.0, 128.09, 128.01, 127.0, 113.4, 96.8, 86.6, 70.0, 64.4, 55.3, 54.6, 41.6, 35.6. IR (solid) ν (cm^{-1}) = 3269 (br), 2930, 2835, 1651, 1607, 1586, 1484, 1447, 1383, 1347, 1299, 1246, 1173, 1114, 1070, 1030, 982, 905, 827, 777, 725, 701, 643, 582. HRMS calcd for $C_{31}H_{34}N_4O_5Br$ $(M+H)^+$ 621.1707, found $(M+H)^+$ 621.1701.

Phosphoramidite 1. To a solution of compound **4** (450 mg, 0.724 mmol) in anhydrous methylene chloride (3.6 mL) under nitrogen was added a 1 M solution of 4,5-dicyanoimidazole (0.51 mL in

anhydrous acetonitrile). This was followed by the dropwise addition of 2-cyanoethyl *N*,*N*,*N'*,*N'*-tetraisopropylphosphordiamidite (0.24 mL, 0.76 mmol) and the solution stirred at room temperature. After two hours, the reaction mixture was diluted with methylene chloride, washed twice with saturated aqueous $NaHCO_3$, dried over Na_2SO_4, and then concentrated. The product was purified via flash chromatography over silica gel starting with 3:2:0.01 hexanes:acetone:Et_3N, then eluting with 1:1:0.01 hexanes:acetone:Et_3N to afford the phosphoramidite **1** as a white foam (460 mg, 77%). ^{31}P-NMR (162 MHz, $CDCl_3$) δ (ppm) 150.1, 149.9. HRMS calcd for $C_{40}H_{51}N_6O_6BrP$ $(M+H)^+$ 821.2786, found $(M+H)^+$ 821.2814.

2.) GNA Oligonucleotide Synthesis and Purification

The GNA oligonucleotide 3'-G^{Br}CGCGC-2' was prepared on an ABI 394 DNA/RNA Synthesizer on a one micromole scale. GNA phosphoramidites were used at a concentration of 100 mM with a standard protocol for 2-cyanoethyl phosphoramidites, except that the coupling was extended to 3 minutes. After the trityl-on synthesis, the resin was incubated with concentrated aqueous ammonia at room temperature overnight in the dark. The entire solution was then applied directly to a Sep-Pak Classic reverse phase column (Waters, 360 mg) and washed sequentially with 3% NH_4OH (15 mL), water (10 mL), 1.5% aqueous TFA (10 mL), water (10 mL), and the oligo finally eluted with 20% aqueous acetonitrile. The oligo was further purified by HPLC eluting over a Waters XTerra column (MS C_{18}, 4.6 x 50 mm, 2.5 µM particle size) at 60 °C with a linear gradient (flow = 1.0 mL/min) of acetonitrile and aqueous triethylammonium acetate buffer (50 mM, pH = 7.0). Identity was confirmed by MALDI-TOF MS.

3.) CD-Melting Experiments

Melting curves of GNA duplexes with the self-complementary strands 3'-G^{Br}CGCGC-2' and 3'-GCGCGC-2' were measured by CD spectroscopy at 273 nm with a JASCO J-810 spectropolarimeter at an oligonucleotide concentration of 20 µM in 10 mM sodium phosphate buffer (pH = 7.0) containing 100 mM NaCl.

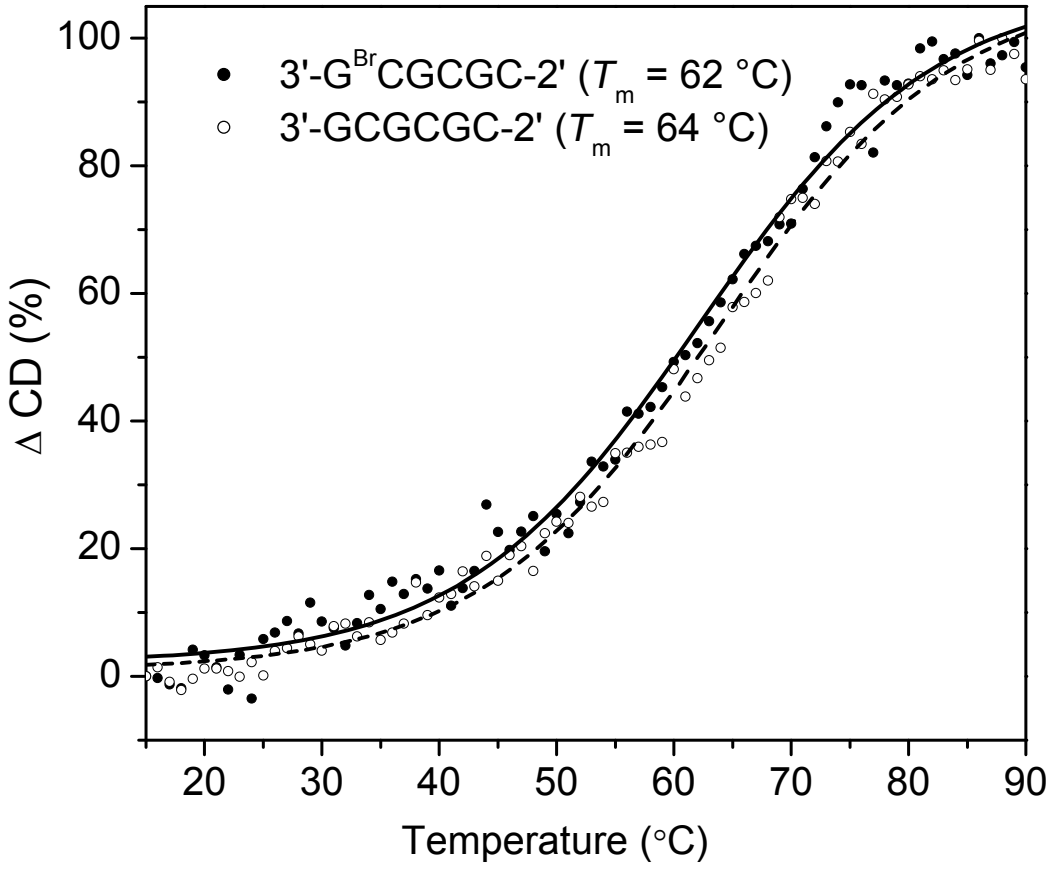


Figure S1. CD-melting curves measured at 273 nm with 20 µM oligonucleotides in 10 mM sodium phosphate buffer (pH = 7.0), 100 mM NaCl. The bromine modification decreases the T_m value only slightly (ΔT_m = -2 °C).

4.) Crystallization, Data Collection, and Structure Solution

Crystals of self-complementary duplex GNA 3'-G^{Br}CGCGC-2' were grown using the sitting drop vapor diffusion method with buffers from the Nucleic Acid Mini Screen (Hampton Research). Crystallization conditions consisted of 2 mM duplex GNA (2 µL) and buffer (4 µL) against a reservoir of 35% MPD in water (1 mL). The duplex crystallized as rectangular rods which appeared after 1-2 weeks at 4 °C in a buffer consisting of 10% 2-methyl-2,4-pentanediol, 40 mM sodium cacodylate (pH = 5.5), 20 mM cobalt hexamine, 80 mM sodium chloride, and 20 mM magnesium chloride. Crystals were cryoprotected by raising the concentration of MPD in the buffer to 30% and subsequently picked from the drop with nylon loops and frozen in liquid N_2. Data was recorded from a single crystal at beamline ID29, ESRF in Grenoble. After performing a fluorescence scan to determine the precise bromine absorption edge in the crystal, SAD data were collected with separate scans at high and low resolution for the brominated derivative. All data were integrated and merged using XDS and XSCALE[2] and phased by SHELXC/D/E[3] and SHARP.[4] The initial map provided density that was unambiguous for all the bases and phosphates of the duplex. Automated and manual refinements were performed using REFMAC5[5] and COOT.[6] The coordinates and structure factors have been deposited into the RCSB Protein Data Bank under accession code 2WNA.

Table S1. Data collection, phasing, and refinement statistics of 3'-G^{Br}CGCGC-2'.

Data collection

Dataset (wavelength Å)	Cell (Å)	Resol. (Å)	Measured, unique reflections	R_{merge}[a]	$I/\sigma(I)$[b]	B_{Wilson} (Å^2)	Completeness
Peak (0.91826 Å)	a=20.42, b=42.00, c=28.67	21.0 - 0.965	30224, 7576	0.059 (0.132)	17.7 (7.2)	4.62	0.982 (0.914)

SAD Phasing in SHELX

Resolution (Å)	21-2.14	-1.68	-1.47	-1.33	-1.24	-1.16	-1.11	-1.06	-1.02	-0.97
CC (map)	0.903	0.958	0.937	0.934	0.945	0.953	0.951	0.941	0.956	0.942
Total FOM	0.850									

Refinement

Resolution range (Å)	8.00-0.965
Reflections (work, test)	7025, 533
R-factor / R_{free}[c]	0.105 / 0.127
Average *B*-factor	6.69
Bases, phosphates, ions	6, 5, 3
Water molecules	58
R.m.s. deviation bonds (Å)	0.006
R.m.s. deviation angles (°)	1.622

[a] $R_{merge} = \sum\sum|I_j(h)-\langle I(h)\rangle|/\sum\sum I_j(h)$; values in parentheses correspond to highest resolution shell. [b] As calculated with the program SCALA. [c] $R = \sum||\boldsymbol{F}_o|-k|\boldsymbol{F}_c||/\sum|\boldsymbol{F}_o|$ with k as scaling factor; R_{free} calculated with test set.

5.) Backbone Torsional Angles and Phosphate Bond Angles

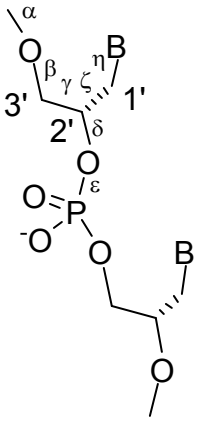
Table S2. Backbone torsional angles.[a]

Nucleotide	α	β	γ	δ	ε	ζ	η
G-1			-63	-97	-72	-65	-89
^{Br}C-2	-171	-151	-178	-92	-68	-56	-76
G-3	149	152	-68	-103	-95	-67	-91
C-4	-170	-119	-173	-109	-44	-62	-82
G-5	127	146	-68	-89	-76	-63	-87
C-6	-112	178	171			-59	-86

[a] Measured in the 3'→2' direction

Table S3. Phosphate bond angles.[a]

Linkage	O3G-P-O2G	O3G-P-O1P	O3G-P-O2P	O1P-P-O2P	O2G-P-O1P	O2G-P-O2P
G-1 – ^{Br}C-2	99.1 (99.6)	109.5 (107.0)	111.7 (112.1)	114.4 (116.3)	107.5 (108.8)	113.5 (111.7)
^{Br}C-2 – G-3	100.3 (100.4)	111.6 (105.8)	110.3 (110.5)	116.8 (117.1)	105.1 (111.5)	111.4 (110.2)
G-3 – C-4	99.9 (100.0)	110.8 (107.5)	110.9 (111.4)	114.1 (116.7)	107.8 (109.9)	112.4 (110.1)
C-4 – G-5	103.0 (103.6)	105.3 (100.6)	108.7 (108.0)	129.3 (126.1)	99.3 (107.0)	108.3 (109.2)
G-5 – C-6	104.7 (104.0)	111.6 (107.8)	107.4 (110.1)	113.7 (114.0)	106.5 (111.1)	112.6 (109.4)



(*S*)-**GNA**

[a] Determined by refining the structure with no restraints on bond angles. Bond angles after refinement using the new stereochemical library shown in parentheses.

6.) Bromine-Bromine Packing Contacts

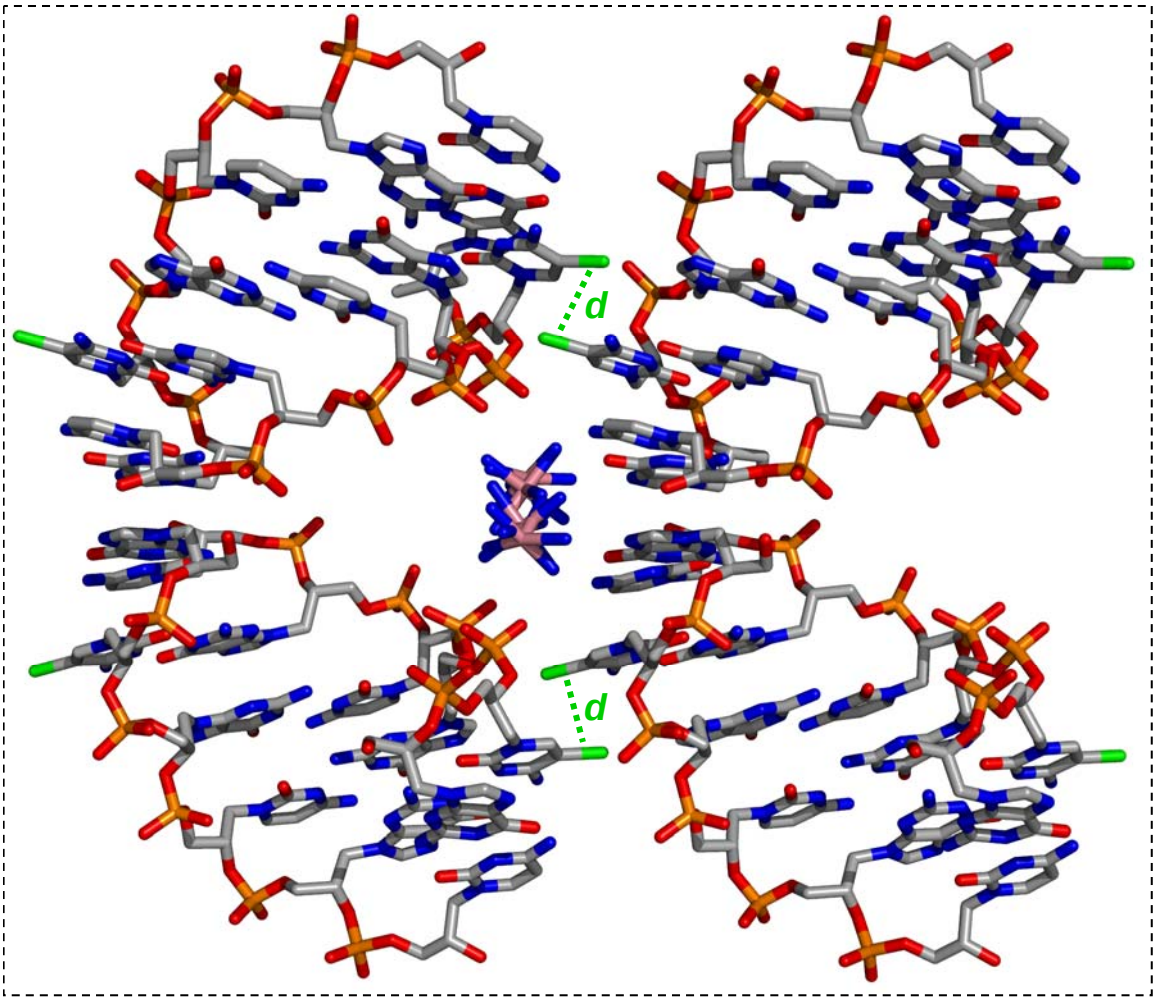


Figure S2. Packing contacts between duplexes in the crystal structure of 3'-G^{Br}CGCGC-2'. Bromine atoms are highlighted in green. A single cobalt (pink) hexamine molecule, adopting two different positions in the crystal, is shown in the middle. d = 3.96 Å.

7.) Analysis of GNA Duplex Parameters

Although the GNA duplexes form roughly a quasi-continuous helix throughout the crystal lattice (Figure S3), the distances of the free 3' and 2' hydroxyls of two adjacent duplexes exceeded with 4.8 Å the range (for comparison: 2.54 Å intra-strand distance), where an analysis of overall duplex parameters by CURVES is feasible. We therefore constructed a continuous model for a Type N helix by repetitively superimposing the ends of the GNA duplexes onto each other: A5-A6/A1*-A2* onto A1-A2/A5*-A6*. As a consequence the P-P distance along the groove of the GNA duplex decreased from 11.5 Å to 8 Å and the turn height from 12 base pairs per turn to 10 (Figure S3A,B). Given the intrinsic conformational flexibility of GNA as compared with its Type M conformation, it may be suggested that the extrapolated Type N conformation, which was then used for the CURVES analysis, adopts a somehow squeezed conformation. Nevertheless, the circular cross section of this band-like GNA duplex is likewise apparent as in the crystal lattice (Figure S3C).

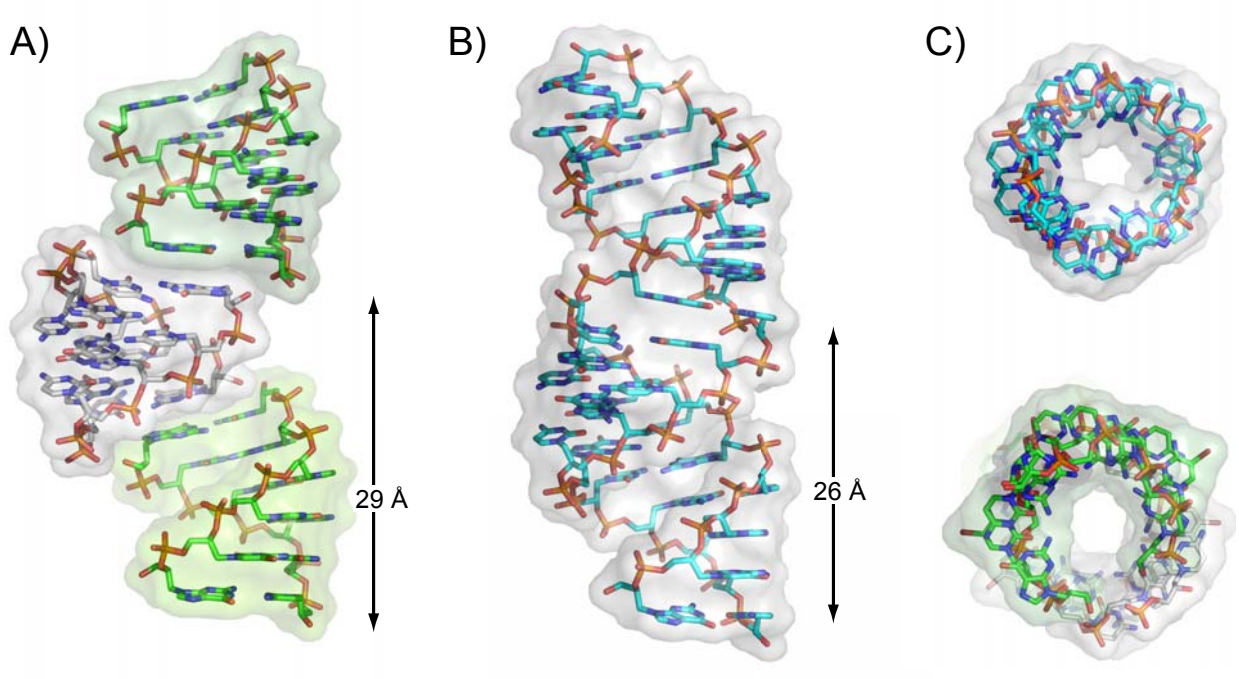


Figure S3. A) Three stacked GNA duplexes are shown as found in the crystal lattice. B) The continuous GNA duplex with 10 base pairs per turn was obtained by extrapolating its conformation from the 6mer structure. C) View along the helix axis of the experimental (top) and modeled (bottom) Type N GNA structures.

References:

1. Zhang, L.; Peritz, A. E.; Carroll, P. J.; Meggers, E. *Synthesis* **2006**, 645.
2. Kabsch, W., *J. Appl. Crystallog*. **1993**, *26*, 795-800.
3. Schneider, T. R.; Sheldrick, G. M., *Acta Crystallog. D* **2002**, *58*, 1772-1779.
4. Bricogne, G.; Vonrhein, C.; Flensburg, C.; Schiltz, M.; Paciorek, W., *Acta Crystallog. D* **2003**, *59*, 2023-2030.
5. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J., *Acta Crystallog. Sect. D* **1997**, *53*, 240-255.
6. Emsley, P.; Cowtan, K., *Acta Crystallog. Sect. D* **2004,** *60*, 2126-2132.