View Article Online / Journal Homepage / Table of Contents for this issue

Atomic resolution duplex structure of the simplified nucleic acid GNA†

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

Received (in Cambridge, UK) 14th August 2009, Accepted 25th November 2009
First published as an Advance Article on the web 21st December 2009
DOI: 10.1039/b916851f

Published on 21 December 2009. Downloaded by Christian Albrechts Universitat zu Kiel on 23/10/2014 14:30:16.

Double helix variations of glycol nucleic acids (GNA) are revealed by the atomic resolution crystal structure of a 6mer GNA duplex containing solely Watson–Crick type hydrogen-bonded base pairs.

We recently discovered the remarkable duplex formation properties of glycol nucleic acids (GNA).[1–5] Bearing an acyclic backbone of propylene glycol nucleosides that are connected by phosphodiester bonds, GNA constitutes the most simplified solution for a phosphodiester-bond-containing nucleic acid (Fig. 1). Nevertheless, the thermal stabilities of GNA duplexes significantly surpass the stabilities of analogous duplexes of DNA and RNA. The ability to form antiparallel Watson–Crick type double helices was confirmed by an X-ray crystal structure of a (*S*)-GNA duplex from a self-complementary strand 3′-CGHATHCG-2′, in which H denotes an artificial hydroxypyridone nucleobase which forms a copper(II)-mediated homobase pair.[3,4] This metallo-base pair was utilized as a necessary handle to site-selectively introduce two heavy atoms per duplex for phasing the crystallographic data. However, it is uncertain how representative such a metallo-GNA duplex is, since it is unknown to what extent the artificial metallo-base pair influences the global GNA duplex structure.

Here we report the first crystal structure of a (*S*)-GNA duplex composed entirely of base pairs with Watson–Crick hydrogen bonding. We have determined the atomic resolution structure of a (*S*)-GNA duplex formed from the self-complementary strand 3′-G^{Br}CGCGC-2′ by X-ray crystallography. The GNA nucleotide ^{Br}C is composed of the 5-bromocytosine nucleobase which establishes a hydrogen-bonded base pair with guanine in the same fashion as a cytosine nucleotide with little or no structural pertubation.[6]

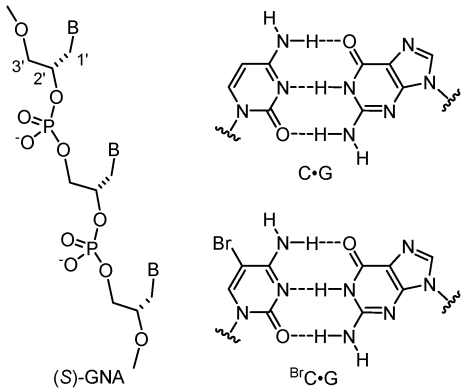
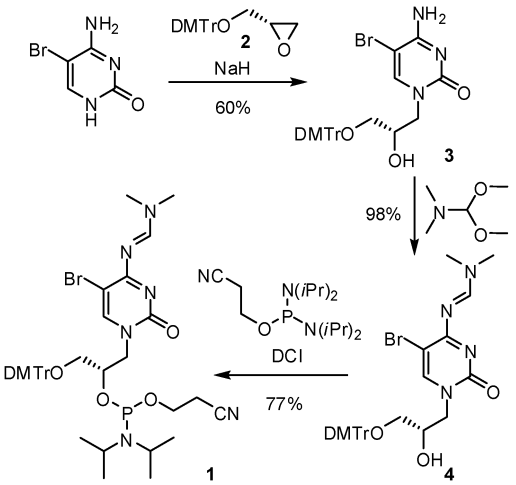


Fig. 1 Constitution of (*S*)-GNA oligonucleotides and the base pairs used in this study.

The phosphoramidite **1** was synthesized from 5-bromocytosine in just three steps with an overall yield of 45% (Scheme 1). For this, 5-bromocytosine was reacted with (*S*)-glycidyl 4,4′-dimethoxytrityl ether **2** in the presence of 0.2 equivalents of NaH in DMF to afford the nucleoside **3** in yields of 60%. The exocyclic amino group was subsequently protected as a dimethylformamidine group by reacting **3** with dimethylformamide dimethylacetal in MeOH at 50 °C to afford **4** in nearly quantitative yield. Conversion to the phosphoramidite **1** was then accomplished in 77% yield by reaction with 2-cyanoethyl *N*,*N*,*N*′,*N*′-tetraisopropyl phosphordiamidite and substoichiometric amounts of 4,5-dicyanoimidazole in CH_2Cl_2.[7] The self-complementary 6mer (*S*)-GNA strand 3′-G^{Br}CGCGC-2′ was synthesized and purified according to published standard procedures.[2,7]

Crystals diffracting up to 0.965 Å resolution were obtained from a sample of the purified oligonucleotide using the sitting drop vapor diffusion method and 2-methyl-2,4-pentanediol as precipitant. The asymmetric unit of the refined structure 3′-G^{Br}CGCGC-2′ contains one single strand, 58 H_2O, one Mg^{2+}, and two $[Co(NH_3)_6]^{3+}$ cations. The atomic resolution of this structure allows us for the first time to determine



Scheme 1 Synthesis of the phosphoramidite **1** for automated solid phase synthesis of ^{Br}C-containing GNA strands. DCI = 4,5-dicyanoimidazole.

Fachbereich Chemie, Philipps-Universität Marburg, Hans-Meerwein-Strasse, 35032 Marburg, Germany. E-mail: essen@chemie.uni-marburg.de, meggers@chemie.uni-marburg.de

† Electronic supplementary information (ESI) available: Synthetic procedures with analytical data, experimental details regarding GNA crystallography, backbone torsional angles and phosphate bond angles, analysis of GNA duplex parameters. See DOI: 10.1039/b916851f

 This journal is © The Royal Society of Chemistry 2010

View Article Online

Published on 21 December 2009. Downloaded by Christian Albrechts Universitat zu Kiel on 23/10/2014 14:30:16.

Table 1 Crystallographic data and refinement statistics

Data collection statistics	
Wavelength/Å	0.9183
Resolution/Å	21.0–0.965
Space group	$C222_1$
Completeness	0.982 (0.914)
Redundancy	4.0
R_{merge} [a,b]	0.059 (0.132)
$I/\sigma(I)$ [a,c]	17.7 (7.2)
B_{wilson}/Å^2	4.6
Refinement statistics	
Resolution/Å	8.00–0.965
R-factor/R_{free} [d]	0.105/0.127
Average B-factor/Å^2	6.69
Water molecules	58
Rmsd bonds/Å	0.006
Rmsd angles/Å	1.622

[a] Values in parentheses correspond to the highest resolution shell. [b] $R_{merge} = \sum\sum|I_j(h) - \langle I(h)\rangle|/\sum\sum I_j(h)$. [c] As calculated with the program SCALA. [d] $R = \sum||F_o| - k|F_c||/\sum|F_o|$ with k as scaling factor; R_{free} calculated with test set (7% of all data). The coordinates and structure factors are deposited in the RCSB under accession code 2WNA.

Fig. 2 SIGMAA-weighted $2F_{obs} - F_{calc}$ electron density of the 3′-terminal glycol nucleotides contoured at 1.5σ.

accurate bond lengths, torsional angles, and phosphate bond angles (see ESI†). Crystallographic data and refinement statistics are given in Table 1. Fig. 2 visualizes the atomic resolution with the electron density of a dinucleotide unit contoured at 1.5σ.

The self-complementary (*S*)-GNA strand forms an anti-parallel duplex with Watson–Crick base pairing and coaxial stacking of individual duplexes, resulting in a quasi-continuous helix within the crystal lattice (Fig. 3, see also ESI†). All base pairs form standard Watson–Crick hydrogen bonding patterns with the 5-bromocytosine nucleotide appearing to have little, if any, distorting effect (Fig. 4).[8] The distances between the C1′ carbons of opposing nucleotides range from 10.71 to 10.85 Å which is in agreement with average values found in DNA duplexes (10.85 Å). A striking common feature of (*S*)-GNA and the previous metallo-(*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 in this new duplex from −46° to −53° (Table 2).[9] This results in extensive 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. Another common aspect with the metallo-(*S*)-GNA is the formation of a helical ribbon by the (*S*)-GNA that is loosely wrapped around the central helix axis, displaying just one large groove (Fig. 3 and 4).

Unlike the previously reported metallo-(*S*)-GNA double helix (from now on referred to as Type M),[3] the structure of the continuous right-handed (*S*)-GNA double helix (referred to as Type N), that can be derived from its 6mer structure by repetitive superimposition, shows two significant differences: first of all, a comparison of the two (*S*)-GNA helices in Fig. 3 demonstrates that the new Type N helix is considerably compressed along the *z*-axis relative to the Type M helix. Although this compression is accompanied by very little change in the helix diameter, it causes the Type N helix to adopt a significantly shorter helical pitch of 26 Å with 10 residues per turn, in comparison to 60 Å with 16 residues

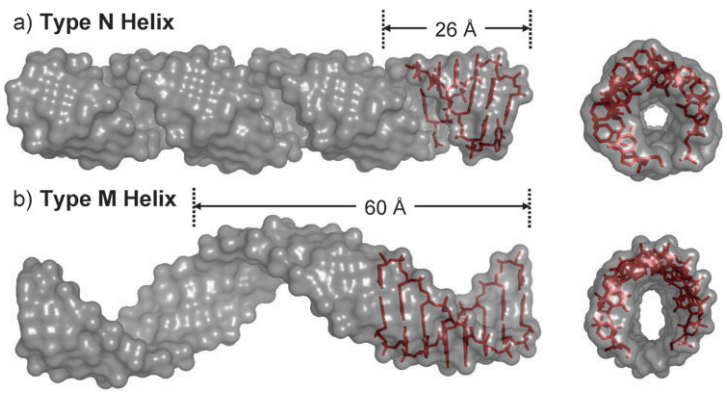


Fig. 3 Comparison of the (*S*)-GNA overall duplex structures of 3′-G^{Br}CGCGC-2′ (Type N) and the metallo-GNA duplex 3′-CGHATHCG-2′ (Type M). The continuous model of the Type N helix was derived by repetitively superimposing the ends of individual 6mer GNA duplexes onto each other and it differs slightly from the coaxially stacked 6mer duplexes in the crystal lattice (see ESI† for more details). Single duplexes are shown in red sticks. Right: view along the helix axis. Images generated using PyMOL.

View Article Online

Published on 21 December 2009. Downloaded by Christian Albrechts Universitat zu Kiel on 23/10/2014 14:30:16.

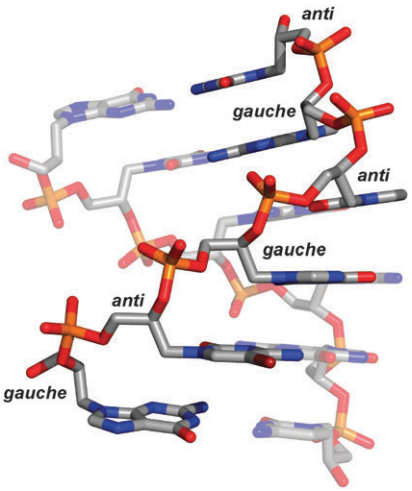


Fig. 4 Structure of a single (*S*)-GNA duplex from 3′-G^{Br}CGCGC-2′. *Gauche* and *anti* refer to the torsional angles between the vicinal C–O bonds C2′–O and C3′–O. Generated using PyMOL.

Table 2 Comparison of average helical parameters for (*S*)-GNA, B-DNA, and A-DNA

	(*S*)-GNA Type M[a]	(*S*)-GNA Type N[a]	B-DNA[a]	A-DNA[a]
Helical sense	Right	Right	Right	Right
Residues per turn	16	10	10	12
Helical pitch/Å	60	26	34	34
Helical rise/Å	3.8	2.6	3.4	2.9
x-Displacement/Å	−7.0	−6.0	0.1	−4.2
Tilt[b]/°	0.0	0.5	0.1	−0.1
Roll[b]/°	−2.7	6.4	0.6	8.0
Twist[b]/°	23.5	35.7	36.0	31.0
Slide[b]/°	−3.5	−3.4	0.2	−1.5
P–P distance[c]/Å	5.4	5.4	7.0	5.9

[a] Data for GNA were calculated using the program CURVES (ref. 12). Data for B-DNA and A-DNA were taken from ref. 10 and 11. [b] Local base pair step parameters. [c] Intrastrand P–P distances.

per turn for the Type M helix (Table 2).[10–12] This shorter helical pitch results from the much stronger twist in the new Type N helix which subsequently brings the phosphate groups of opposing strands in closer contact. The base pairs of the Type N helix are displaced from the helix axis (*x*-displacement) by 5.4 to 6.8 Å resulting in a large hollow core, similar to the previous metallo-GNA duplex (Type M). However, the hollow core is circular rather than elliptical as observed in the Type M helix for which there is a greater variance in *x*-displacement values from the central helix axis (5.1–8.6 Å).

The second difference compared to Type M helices concerns the conformation of the propylene glycol backbone, especially the torsional angles between the vicinal C3′–O and C2′–O bonds. In the new Type N structure the nucleotides adopt strictly alternating *gauche* and *anti* conformations with average torsional angles γ of −66° and −174°, respectively (Fig. 4). Apparently, the acyclic backbone has the flexibility to adopt these two distinct conformations and can vary them within a single duplex.

Interestingly, the crystal structure of a DNA-fragment with the same sequence, 5′-CGCGCG-3′, demonstrated the existence of the Z-DNA.[13] Like the latter, the repeat unit of Type N GNA-helices also comprises two nucleotides with alternating backbone conformations suggesting an influence of GC-rich sequences on GNA backbone conformations. However, unlike the Z-DNA, the interconversion between Type M and N helices of GNA should not require extensive conformational changes of the Watson–Crick-like base pairs.

In conclusion, we have presented the first atomic resolution crystal structure of a (*S*)-GNA duplex containing solely Watson–Crick type hydrogen-bonded base pairs. This new GNA double helix structure reveals common features but also significant differences compared to the previously reported metallo-GNA duplex structure and gives insight into the degree of variation within the family of GNA duplex structures. The main contributor of this variation is the ability of the propylene glycol nucleotides to choose between a *gauche* or *anti* conformation around the vicinal C–O bonds and it can be hypothesized that this may vary with the duplex sequence. The atomic resolution of this structure allows us to accurately determine bond and angle parameters and will serve as the basis for solving and refining additional GNA duplex structures of different sequences and length in order to gain further insight into the exceptional duplex formation abilities of this minimal nucleic acid backbone.

Notes and references

1 L. Zhang, A. Peritz and E. Meggers, *J. Am. Chem. Soc.*, 2005, **127**, 4174.
2 M. K. Schlegel, A. E. Peritz, K. Kittigowittana, L. Zhang and E. Meggers, *ChemBioChem*, 2007, **8**, 927.
3 M. K. Schlegel, L.-O. Essen and E. Meggers, *J. Am. Chem. Soc.*, 2008, **130**, 8158.
4 M. K. Schlegel, L. Zhang, N. Pagano and E. Meggers, *Org. Biomol. Chem.*, 2009, **7**, 476.
5 M. Schlegel, X. Xie, L. Zhang and E. Meggers, *Angew. Chem., Int. Ed.*, 2009, **48**, 960.
6 For the use of 5-bromocytosine in nucleic acid crystallography, see for example: D. B. Tippin and M. Sundaralingam, *J. Mol. Biol.*, 1997, **267**, 1171; M. Ortiz-Lombardia, A. Gonzalez, R. Eritja, J. Aymami, F. Azorin and M. Coll, *Nat. Struct. Biol.*, 1999, **6**, 913.
7 For DMF-protected GNA phosphoramidites, see also: M. K. Schlegel and E. Meggers, *J. Org. Chem.*, 2009, **74**, 4615.
8 Short intermolecular Br···Br contacts at 3.96 Å suggest that the bromine substituents appear to have an effect on the packing of the individual duplexes in the crystal. See ESI†.
9 Such a strong backbone–base inclination has also been observed for unnatural hexose nucleic acids. See: R. Declercq, A. Van Aerschot, R. J. Read, P. Herdewijn and L. Van Meervelt, *J. Am. Chem. Soc.*, 2002, **124**, 928; M. Egli, P. S. Pallan, R. Pattanayek, C. J. Wilds, P. Lubini, G. Minasov, M. Dobler, C. J. Leumann and A. Eschenmoser, *J. Am. Chem. Soc.*, 2006, **128**, 10847.
10 R. E. Dickerson, *Methods Enzymol.*, 1992, **211**, 67.
11 W. K. Olson, M. Bansal, S. K. Burley, R. E. Dickerson, M. Gerstein, S. C. Harvey, U. Heinemann, X.-J. Lu, S. Neidle, Z. Shakked, H. Sklenar, M. Suzuki, C.-S. Tung, E. Westhof, C. Wolberger and H. M. Berman, *J. Mol. Biol.*, 2001, **313**, 229.
12 R. Lavery and H. Sklenar, *J. Biomol. Struct. Dyn.*, 1988, **6**, 63; R. Lavery and H. Sklenar, *J. Biomol. Struct. Dyn.*, 1989, **6**, 655.
13 A. H.-J. Wang, G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel and A. Rich, *Nature*, 1979, **282**, 680.

 This journal is © The Royal Society of Chemistry 2010