nature chemistry

SUPPLEMENTARY INFORMATION

DOI: 10.1038/NCHEM.512

Solution structure of a DNA double helix with consecutive metal-mediated base pairs

Silke Johannsen,[a] *Nicole Megger,*[b] *Dominik Böhme,*[b] *Roland K. O. Sigel,*[a,*] *Jens Müller* [b,*]

a) Institute of Inorganic Chemistry, University of Zürich, Winterthurerstr. 190, 8057 Zürich, Switzerland

b) Westfälische Wilhelms-Universität Münster, Institut für Anorganische und Analytische Chemie, Corrensstr. 28/30, 48149 Münster, Germany

roland.sigel@aci.uzh.ch, phone: +41 44 635 4652, fax: +41 44 635 6802

mueller.j@uni-muenster.de, phone: +49 251 833 6006, fax: +49 251 833 6007

Supplementary Results

The modified DNA adopts a hairpin in solution

For our structural investigations we used a self-complementary sequence with three consecutive artificial imidazole nucleosides in its centre**.** We could previously show that such a type of sequence adopts a hairpin structure with the artificial nucleosides being placed in the loop in the absence of transition metal ions. A rearrangement to form a regular double helix with neighbouring metal-mediated base pairs takes place in the presence of silver(I) ions[1]. Based on this observation, which is also known from other nucleic acids, other nucleobases and other metal ions[2], we substituted the 1,2,4-triazole nucleosides used in those previous experiments by imidazole nucleosides due to their superior silver(I)-binding properties[3]. Moreover, the presence of three hydrogen atoms in imidazole as compared with two hydrogen atoms in triazole yields more ^{1}H NMR resonances, thereby increasing the number of constraints as necessary for the structure refinement.

The proton resonances of the oligonucleotide in the absence of silver(I) ions were assigned by [^{1}H, ^{1}H]-NOESY spectra on the basis of the sequential walks along the H1' and aromatic protons (Fig. 2a). Three additional regions (H2'/aromatic, H2''/aromatic and aromatic/aromatic) served to cross-validate the assignments. Typical resonances for a B-type helical conformation were observed for nucleotides T1-T7 and A11-A17. The [^{1}H, ^{1}H]-NOESY spectrum acquired in H_2O shows a large overlap of the imino proton resonances. Nevertheless, 13 interstrand correlations (mainly between adenine-H2 and H1' on the opposite strand) prove the formation of seven distinct adenine-thymine Watson-Crick base pairs. Only very weak and broad aromatic resonances could be observed for the three imidazole moieties, and consequently the sequential walk could not be followed throughout the loop. Nevertheless, based on the sugar-sugar region in the NOESY spectrum as well as TOCSY and [^{1}H, ^{13}C]-HSQC spectra, the three artificial nucleotides could be reliably identified and assigned. The [^{31}P]-NMR spectrum shows a set of signals between 0 and –0.5 ppm and a small broad peak at 0.45 ppm being consistent with a general hairpin structure.

However, it is difficult to differentiate unambiguously between a hairpin structure and a regular double helix by simple [^{1}H]-NMR experiments. We thus determined the hydrodynamic radius r_H of the oligonucleotide by DOSY (diffusion-ordered spectroscopy) spectra as well as DLS (dynamic light scattering) measurements. These two independent methods gave identical values within the error limits of 1.46±0.09 and 1.4±0.2 nm, respectively. A theoretical calculation of r_H using well-established equations[2] and assuming a hairpin structure is in good agreement with the experimental values observed by DOSY and DLS (Supplementary Table S1), thereby providing evidence that indeed a hairpin structure was formed in the absence of silver(I) ions.

The solution structure of this modified DNA hairpin is shown in Supplementary Fig. S1. The structure determination of the oligonucleotide in the absence of silver(I) ions is based on 354

conformationally restrictive NOE distance restraints (Table 1). The DNA adopts a stable hairpin structure with a well-defined helical region closed by an unstructured loop composed of the three imidazole moieties. The overall r.m.s.d. of all heavy atoms from the 20 lowest energy structures is 1.6±0.4 Å (Table 1) and the independent superposition of the helical region results in an r.m.s.d. of 0.9±0.3 Å. The poor definition of the loop region is well in line with the small number of resonances observed and probably due to the semi-protonated artificial nucleotides: The pK_a of the imidazole base is estimated to be around 6.4 (imidazole β-nucleoside pK_a = 6.01±0.05, plus the influence of the phosphate group of 0.39±0.04 log units[3,4]) and consequently close to the pH of 7.2 as used in the measurements. Obviously, no stable hydrogen bonding network is established between the imidazole moieties under these conditions.

Supplementary Methods

Preparations and characterization

Characterization of crude imidazole (intermediate product during the synthesis of **1**)

^{1}H NMR (400 MHz, D_2O, pD = 9.44): δ 7.77 (t, J = 9.5 Hz, 1H, H2), 7.13 (dd, J = 6.7 Hz, 2.0 Hz, 2H, H4, H5); ^{13}C NMR (101 MHz, D_2O, pD = 9.44): δ 136.96 (t, $^1J_{NC}$ = 7.0 Hz, C2), 121.7 (d, $^1J_{NC}$ = 5.6 Hz, C4, C5). LRMS (m/z): [M] calcd. for $C_3H_4{}^{15}N_2$, 70; found, 70.

Characterization of 3,5-Di-*p*-toluoyl-1,2-dideoxy-β-1-(^{15}N-imidazol-1-yl)-D-ribofuranose **1**

TLC (cyclohexane:CH_2Cl_2:Et_3N, 50:30:8): R_f = 0.18; ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.93 (m, 4H, tol), 7.70 (dd, *J* = 11.2 Hz, 7.9 Hz, 1H, H2), 7.26 (m, 4H, tol), 7.07 (m, 2H, H4, H5), 6.14 (t, 1H, H1'), 5.66 (m, 1H, H3'), 4.58 (m, 3H, H4', H5', H5''), 2.68 (m, 2H, H2', H2''), 2.43 (s, 3H, CH_3), 2.41 (s, 3H, CH_3); ^{13}C NMR (101 MHz, $CDCl_3$): δ 166.29 (carbonyl C), 166.01 (carbonyl C), 144.68 (tol), 144.34 (tol), 136.02 (d, $^1J_{NC}$ = 11.0 Hz, C2), 130.45 (tol), 130.42 (tol), 130.39 (tol), 130.37 (tol), 129.90 (tol), 129.77 (tol), 129.47 (tol), 129.44 (tol), 126.71 (d, $^1J_{NC}$ = 29.6 Hz, C4), 116.37 (d, $^1J_{NC}$ = 13.6 Hz, C5), 86.39 (d, $^1J_{NC}$ = 12.5 Hz, C1'), 82.70 (C4'), 75.12 (C3'), 64.24 (C5'), 39.48 (C2'), 21.86 (CH_3), 21.81 (CH_3); ^{15}N NMR (41 MHz, $CDCl_3$) δ 258.2 (N3), 182.3 (N1). Calcd. for $C_{24}H_{24}{}^{15}N_2O_5$: C, 68.24; H, 5.73; N, 7.10. Found: C, 67.96; H, 5.67; N, 7.10; HRMS (m/z): $[MH]^+$ calcd. 423.1699; found 423.1692.

^{15}N imidazole nucleoside **2**

The toluoyl protected nucleoside **1** (993 mg, 2.37 mmol) was dissolved in methanol (50 ml), treated with aqueous NH_3 (50 ml) and stirred at ambient temperature for 4.5 h. The solvent was removed in vacuo, and the crude product was purified by column chromatography (SiO_2, CH_2Cl_2:MeOH 20:3) yielding the free nucleoside (377 mg, 2.02 mmol, 85%).

TLC (CH_2Cl_2:MeOH 20:3): R_f = 0.14; ^{1}H NMR (400 MHz, D_2O, pD =8.60): δ 7.89 (dd, J = 10.3 Hz, 7.6 Hz, 1H, H2), 7.32 (d, J = 2.4 Hz, 1H, H5), 7.07 (dd, J = 9.0 Hz, 3.2 Hz, 1H, H4), 6.18 (t, 1H, H1'), 4.52 (dt, J = 6.6 Hz, 3.9 Hz, 1H, H3'), 4.06 (dt, J = 5.4 Hz, 3.9 Hz, 1H, H4'), 3.72 (m, 2H, H5', H5''), 2.58 (dtd, J = 9.0, 6.6, 2.6 Hz, 1H, H2'), 2.48 (m, 1H, H2''); ^{13}C NMR (101 MHz, D_2O, pD = 8.60): δ 137.01 (d, $^1J_{NC}$ = 11.6 Hz, C), 128.52 (d, $^1J_{NC}$ = 4.8 Hz, C), 117.63 (d, $^1J_{NC}$ = 12.7 Hz, C), 86.73 (C4'), 85.90 (d, $^1J_{NC}$ = 11.6 Hz, C1'), 70.95 (C3'), 61.55 (C5'), 39.76 (C2'); ^{15}N NMR (41 MHz, D_2O, pD =8.60): δ 245.6 (N3), 187.1 (N1). Calcd. for $C_8H_{12}{}^{15}N_2O_3$: C, 51.61; H, 6.50; N, 16.11. Found: C, 49.27; H, 6.48; N, 16.07; HRMS (m/z): $[MH]^+$ calcd. 187.0861; found 187.0855.

DMT-protected ^{15}N imidazole nucleoside **3**

^{15}N imidazole nucleoside **2** (227 mg, 1.22 mmol) was dried by co-evaporation with dry pyridine (2 × 5 ml) and then dissolved in dry pyridine (8 ml). 4,4'-dimethoxytritylchloride (1.2 equiv., 496 mg, 1.46 mmol) was added to the solution as well as catalytic amounts of DMAP. After being stirred for 3 h at room temperature the reaction mixture was diluted with CH_2Cl_2 (50 ml) and washed with H_2O (2 × 20 ml). The organic layer was dried ($MgSO_4$), the solvent was removed in vacuo, and the crude product was purified by column chromatography (SiO_2, CH_2Cl_2:MeOH 100:1 → 95:5) yielding the DTM-protected nucleoside **3** (181 mg, 370 μmol) in 30 % yield.

TLC (CH_2Cl_2:MeOH 100:1): R_f = 0; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.62 (t, J = 8.8 Hz, 1H, H2), 7.40 (m, 2H, DMT), 7.27 (m, 7H, DMT), 7.01 (s, 2H, H4, H5), 6.81 (m, 4H, DMT), 6.01 (t, 1H, H1'), 4.51 (m, 1H, H3'), 4.07 (m, 1H, H4'), 3.78 (s, 3H, O–CH_3), 3.77 (s, 3H, O–CH_3), 3.31 (ddd, J = 22.2 Hz, 10.2 Hz, 4.7 Hz, 2H, H5', H5''), 2.41 (m, 2H, H2', H2''); ^{13}C NMR (101 MHz, $CDCl_3$) δ 158.81 (DMT), 144.71 (DMT), 135.86 (d, $^1J_{NC}$ = 2.6 Hz, C2), 130.25 (DMT), 129.69 (d, $^1J_{NC}$ = 3.03 Hz, C4), 128.34 (DMT), 128.13 (DMT), 127.18 (DMT), 113.44 (DMT), 86.80 (DMT), 86.21 (C4'), 85.12 (d, $^1J_{NC}$ = 2.2 Hz, C1'), 72.52 (C4'), 64.12 (C5'), 55.44 (O–CH_3), 41.71 (C2'); Calcd. for $C_{29}H_{30}{}^{15}N_2O_5$: C, 71.30; H, 6.19; N, 6.14. Found: C, 69.60; H, 6.20; N, 6.15; HRMS (m/z): $[MH]^+$ calcd. 489.2168; found 489.2162.

DMT- and phosphoramidite protected imidazole nucleoside **4**

To a solution of **3** (178 mg, 364 μmol, dried *in vacuo* over night) in dry CH_2Cl_2 (4 ml), DIPEA (4 equiv., 188 mg, 1.46 mmol, 253 μl) was added, followed by the addition of 2-cyanoethyl-diisopropylchlorophosphoramidite (1.5 equiv., 129 mg, 546 μmol, 122 μl) under argon. After stirring at ambient temperature for 30 min the reaction mixture was quenched by the addition of ethyl acetate (18 ml). The organic layer was washed with saturated aqueous $NaHCO_3$ solution (10 ml), dried ($MgSO_4$), and the solvent removed in vacuo. The double protected imidazole nucleoside **4** was used without any further purification for solid phase oligonucleotide synthesis.

^{1}H NMR (200 MHz, $CDCl_3$) δ 7.66 (m, 1H, H2), 7.42 (m, 2H, DMT), 7.28 (m, 9H, DMT), 7.07 (m, 2H, H4, H5), 6.81 (m, 4H, DMT), 6.03 (t, 1H, H1'), 4.60 (m, 1H, H3'), 4.23 (m, 1H, H4'), 3.78 (s, 6H, O–CH_3); ^{31}P NMR (81 MHz, $CDCl_3$) δ 149.82, 149.71; HRMS (m/z): $[MH]^+$ calcd. 689.3247; found 689.3244.

Oligonucleotide Preparation

Phosphoramidites of the natural nucleosides as well as CPGs were purchased from Glen Research. The oligonucleotides were prepared using an Expedite 8909 synthesizer in the DMT-off mode and purified by HPLC with a Nucleogen 60-7 DEAE column. The following linear gradients were used during HPLC purification: 0 min: 100% A, 0-10 min: 100-90% A, 10-40 min: 90-75% A, 40-45 min: 75-100% A (1.5 mL/min, solvent A: 20 mM sodium acetate (adjusted to pH 6 by means of acetic acid) + MeCN (4:1), solvent B: solvent A plus 2 M lithium chloride). After desalting via NAP 10 columns, the identity was confirmed by MALDI TOF and ESI mass spectrometry (unlabeled oligonucleotide: $[MH]^+$ calcd. 4999, found 5002; oligonucleotide with ^{15}N imidazole: $[M-H]^-$ calcd. 5003, found 5005).

The desalted and lyophilized samples were dissolved in 200 µL D_2O containing 120 mM $NaClO_4$. Duplex formation was initiated by the addition of 1.5 equiv. Ag^+ (0.2 M $AgNO_3$). One equivalent is defined as the amount of Ag^+ needed to form the three imidazole–Ag^+–imidazole base pairs in the duplex. The solution was heated to 70 °C and then slowly cooled down to ambient temperature to ensure a homogeneous hybridization. Excess Ag^+ was removed by the addition of a chelating resin Chelex-100 (BIO-RAD, Hercules, USA). DNA concentrations were between 0.5 and 1.4 mM. All samples were lyophilized and resuspended in either H_2O/D_2O (9:1) or 99.999% D_2O. The pH or pD, respectively, was adjusted to 7.2 prior to the acquisition of the NMR spectra.

Structure Calculation

For structure calculation of the DNA hairpin, based on ^{31}P NMR spectra, the α, ζ angles were set to exclude the *trans*-range[5], except at the imidazole moieties, which were left unrestrained. Sugar pucker restraints were included based on TOCSY experiments with a 50 ms mixing time. All nucleotides showed strong H1'-H2' and H1'-H3' cross peaks and were restrained to *S*-type, except of imidazole 8 which was left unrestrained. The other backbone torsion angles (β, γ, ε) were set to standard B-form values in the helical region of the structure (T1-T7, A11-A17). Based on the intra-nucleotide H1'-aromatic NOEs, the torsion angle χ was restrained to –120±20° (*anti*) for all residues, except the imidazole moieties, which were left unconstrained. The structure calculations were performed with DYANA 1.5[6] and Xplor-NIH 2.15.0 using the standard implemented force field parameters[7]. Ag^+ ions were defined with a single positive charge. The other parameters used were sigma = 2.8500 and eps = 0.1600. First, 200 initial structures were calculated with DYANA using

distances restraints and dihedral restraints only. The 10 lowest energy structures were used to calculate 200 starting structures with Xplor-NIH, which were subsequently refined. Weak planarity for all identified base pairs was enforced during refinement and hydrogen bonds of these base pairs were maintained by distance restraints.

Derivation of the base pair parameters

Base pair parameters were derived using the programme 3DNA[8]. In the 20 lowest-energy duplex structures, the imidazole moieties were substituted by guanine to permit the use of the programme. As a result, the programme misleadingly assumed the presence of three central base pairs with *syn*-oriented nucleobases (see Supplementary Table S3 for a compilation of the glycosidic torsion angles χ). Hence, the values obtained for tilt and rise for the base-pair steps involving imidazole nucleotides are meaningless.

Supplementary Table S1: Hydrodynamic radii r_H of the hairpin and the metal-modified duplex.

			Hairpin	Duplex
		numbers of base pairs	8.5	17
		L / nm	2.55	5.44
		$q = L / d$	1.28	2.72
theoretical values	spherical model	r_H / nm	1.28	(2.72)
	cylinder model	r_H / nm	(1.36)	1.81
experimental values	DOSY	10^6 D / cm^2/s	1.36 ± 0.08	1.1 ± 0.1
		r_H / nm	1.46 ± 0.09	1.9 ± 0.3
	DLS	r_H / nm	1.4 ± 0.2	1.8 ± 0.2

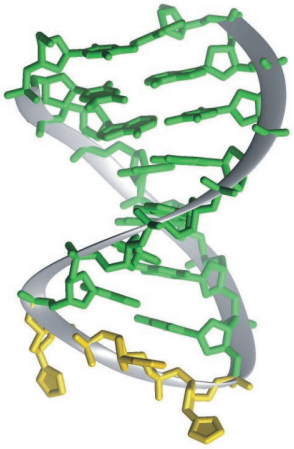
For explanation of abbreviations and further details on the calculation of the theoretical values, see reference 2. For the calculations, a helical rise of 0.34 nm and a diameter of the B-DNA helix of 2.0 nm have been assumed.

Supplementary Table S2: Chemical shift changes upon metalation of the oligonucleotide

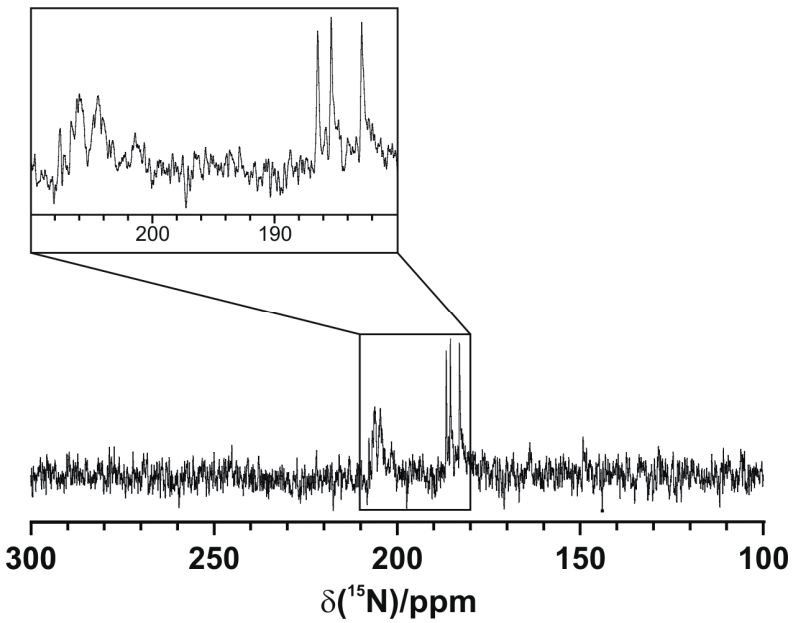
	δ / ppm (without Ag^+)	δ / ppm (with Ag^+)	$\Delta\delta$ / ppm
Im8 H2	8.33	7.25	–1.08
Im8 H4	7.31	6.59	–0.72
Im8 H5	7.50	7.22	–0.28
Im9 H2	7.51	7.11	–0.40
Im9 H4	6.89	6.28	–0.61
Im9 H5	6.96	7.00	0.04
Im10 H2	7.87	6.93	–0.94
Im10 H4	6.82	5.96	–0.86
Im10 H5	7.08	6.78	–0.30
Im8 N1	189.5	186.3	–3.2
Im8 N3	220.5	205.5	–15.0
Im9 N1	188.6	185.3	–3.3
Im9 N3	221.4	206.9	–14.5
Im10 N1	188.1	183.0	–5.1
Im10 N3	221.9	206.0	–15.9

Supplementary Table S3: Glycosidic angle χ / °

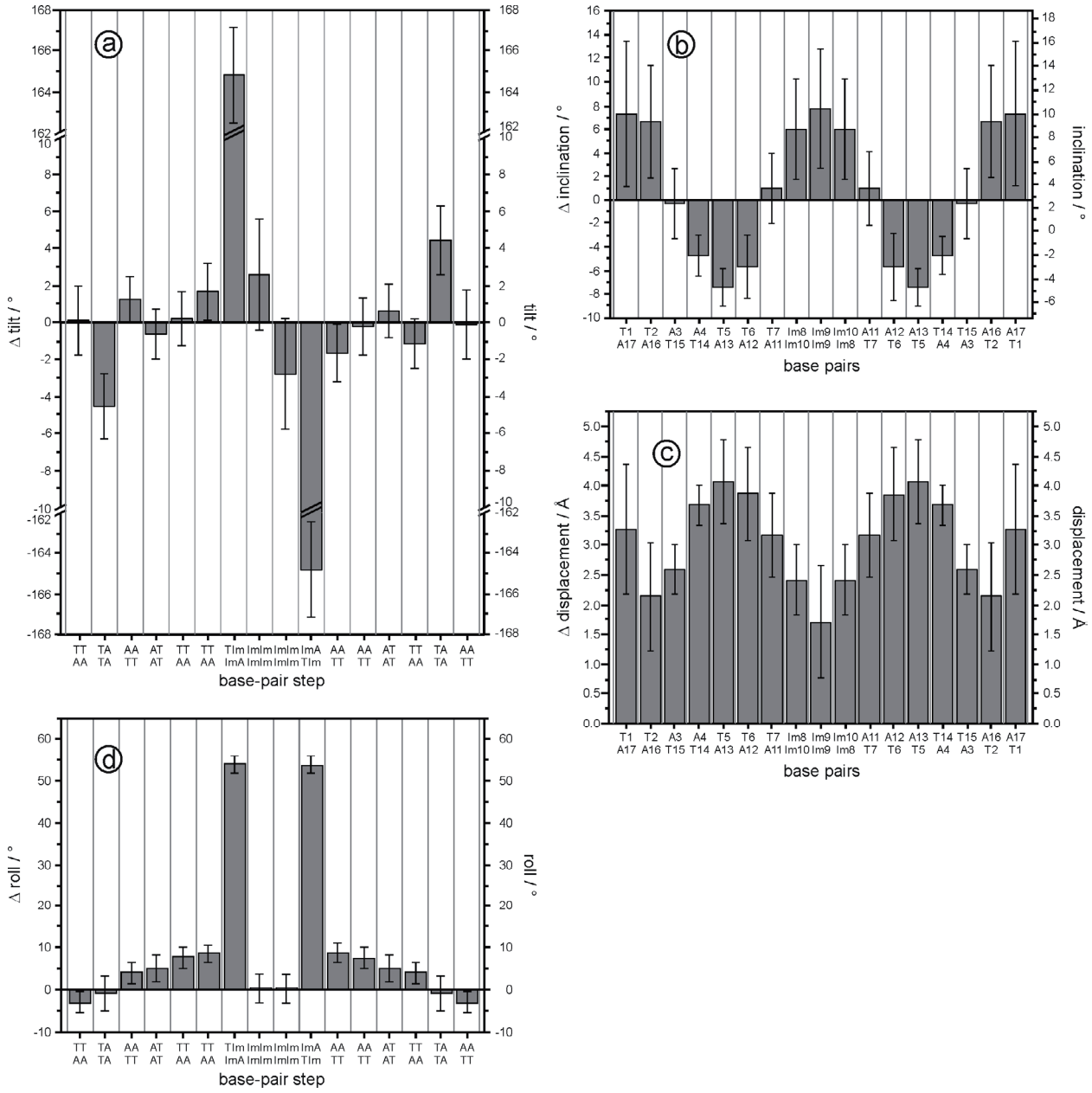
base	strand A	strand B
T1	–121±5	–121±5
T2	–109±3	–109±2
A3	–124±5	–124±4
A4	–111±3	–111±3
T5	–116±4	–116±4
T6	–116±4	–116±5
T7	–109±4	–110±4
Im8	63±5	63±5
Im9	65±3	65±3
Im10	64±7	65±7
A11	–104±4	–104±4
A12	–124±2	–124±2
A13	–112±3	–112±3
T14	–111±5	–111±5
T15	–119±4	–119±4
A16	–116±6	–116±6
A17	–122±3	–122±3



Supplementary Figure S1: Lowest energy solution structure of the hairpin DNA with the three imidazole nucleobases being located in the loop. The B-form helical regions are coloured in green and the imidazole nucleobases in gold. This figure has been prepared with MOLMOL[9].

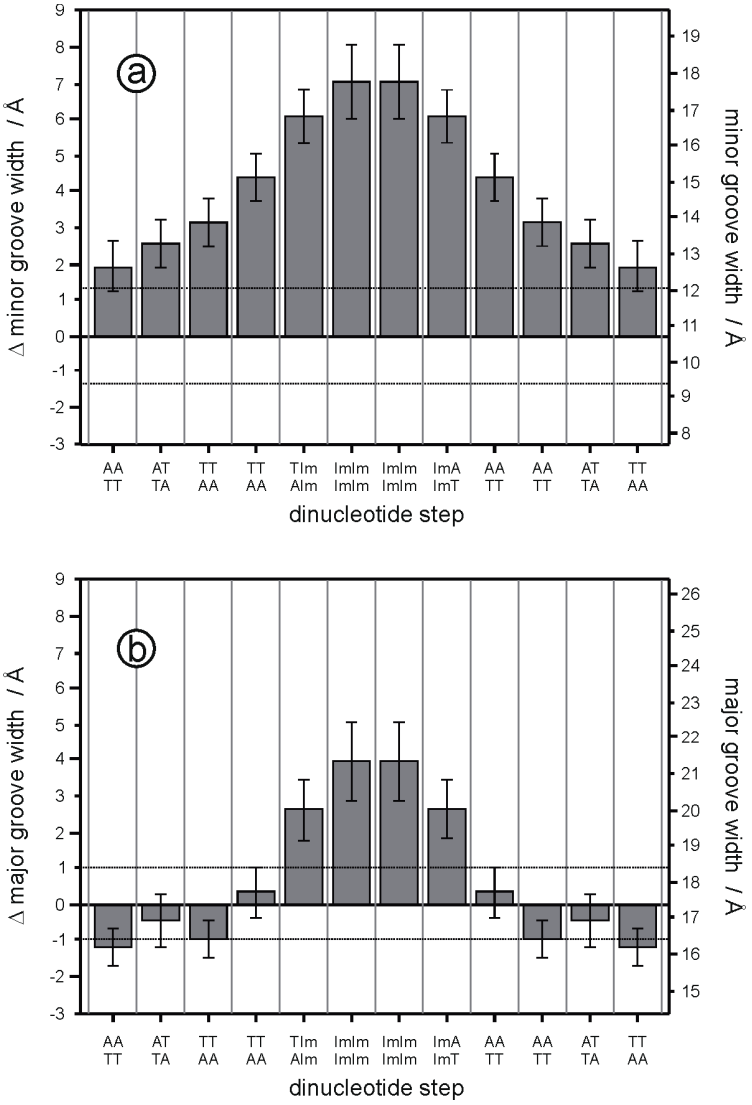


Supplementary Figure S2: [^{15}N]-INEPT spectrum of the Ag^{+}-containing DNA duplex. The lowfield signals of the imidazole N3 are split due to the $^{1}J(^{15}N, ^{107/109}Ag)$ coupling of 86 Hz, but also heavily overlap. This spectrum has been recorded on a Bruker AV 500 MHz machine, equipped with a BBO probe (298K, D_2O, 40k scans, SW = 213 ppm).



Supplementary Figure S3: Selected base pair parameters as derived with the programme 3DNA[8]. Values relative to those of average B-DNA[10] are given on the left scale, absolute values on the right scale. a) Tilt τ, b) inclination η, c) global displacement, and d) roll ρ.

Tilt and roll are local base-pair step parameters, inclination and displacement are global parameters based on C1'-C1' vectors. As a result of how the parameters were derived, the values for tilt and roll for the base-pair steps involving imidazole nucleotides are meaningless (see Methods section in the Supplementary Information for details).



Supplementary Figure S4: Groove widths as derived with the programme 3DNA[8]. Values relative to those of B-DNA[11] are given on the left scale, absolute values on the right scale. The horizontal dotted lines represent the standard deviations of the B-DNA values. a) Minor groove width, b) major groove width.

References

1. Böhme, D., Düpre, N., Megger, D.A. & Müller, J. Conformational Change Induced by Metal-Ion-Binding to DNA Containing the Artificial 1,2,4-Triazole Nucleoside. *Inorg. Chem.* **46**, 10114-10119 (2007).
2. Johannsen, S., Paulus, S., Düpre, N., Müller, J. & Sigel, R.K.O. Using *in vitro* transcription to construct scaffolds for one-dimensional arrays of mercuric ions. *J. Inorg. Biochem.* **102**, 1141-1151 (2008).
3. Müller, J., Böhme, D., Lax, P., Morell Cerdà, M. & Roitzsch, M. Metal Ion Coordination to Azole Nucleosides. *Chem. Eur. J.* **11**, 6246-6253 (2005).
4. Mucha, A., Knobloch, B., Jeżowska-Bojczuk, M., Kozłowski, H. & Sigel, R.K.O. Comparison of the Acid–Base Properties of Ribose and 2'-Deoxyribose Nucleotides. *Chem. Eur. J.* **14**, 6663-6671 (2008).
5. Varani, G., Aboul-ela, F. & Allain, F.H.-T. NMR investigation of RNA structure. *Prog. Nucl. Magn. Reson. Spectrosc.* **29**, 51-127 (1996).
6. Güntert, P., Mumenthaler, C. & Wüthrich, K. Torsion Angle Dynamics for NMR Structure Calculation with the New Program DYANA. *J. Mol. Biol.* **273**, 283-298 (1997).
7. Schwieters, C.D., Kuszewski, J.J., Tjandra, N. & Clore, G.M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65-73 (2003).
8. Lu, X.-J. & Olson, W.K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* **31**, 5108-5121 (2003).
9. Koradi, R., Billeter, M. & Wüthrich, K. MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graphics* **14**, 51-55 (1996).
10. Chandrasekaran, R. & Arnott, S. The structure of B-DNA in oriented fibers. *J. Biomol. Struct. Dyn.* **13**, 1015-1027 (1996).
11. El Hassan, M.A. & Calladine, C.R. Two Distinct Modes of Protein-induced Bending in DNA. *J. Mol. Biol.* **282**, 331-343 (1998).