Electronic Supplementary Material (ESI) for ChemComm.

X-ray Structure of a Lectin B bound DNA Duplex Containing an Unnatural Phenanthrenyl Pair

Pascal Roethlisberger, Andrei Istrate, Maria J. Marcaida Lopez, Achim Stocker, Jean-Louis Reymond and Christian J. Leumann

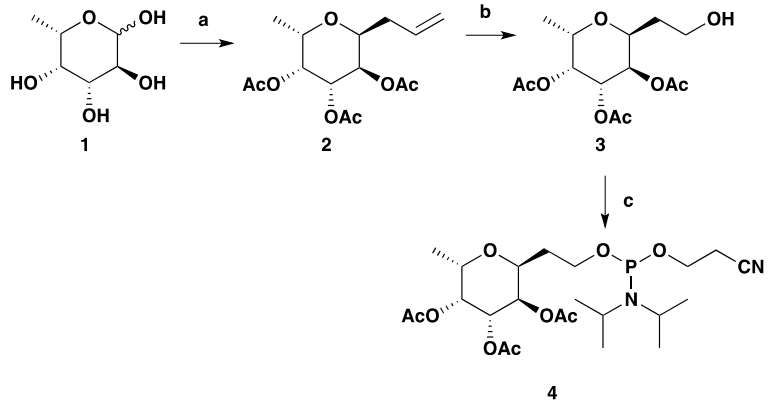
Electronic Supplementary Information

Index

General

All reactions were performed under argon in dried glassware. Anhydrous solvents were obtained by filtration through activated aluminum oxide. Unless stated otherwise all chemicals and solvents were purchased from Sigma. Solvents for extractions and flash chromatography were distilled before use. Flash column chromatography (FC) was performed using silica gel (230–400 mesh) from Silicycle. Thin layer chromatography was carried out on glass-backed plates precoated with silica gel (0.25mm, UV_{254}) from Macherey-Nagel. ^{1}H NMR was recorded at 300MHz or 400MHz on a Bruker AC-300 or a Bruker DRX-400. ^{13}C NMR (75 MHz) was recorded on a Bruker AC-300. ^{31}P NMR spectra were recorded at 121.4 MHz. All spectra were referenced to the signals of the corresponding solvent. Chemical shifts are given in ppm (δ scale) and coupling constants (*J*) in Hz. High resolution nanospray ionization (NSI) mass spectra were recorded on a Thermo Scientific LTQ Orbitrap XL instrument.

Synthesis of FUC analog



Scheme S1. Conditions and reactants: a) i) Triethylamine, DMAP, DCM, Ac_2O, 0°C to RT for 20 h; a) ii) allyltrimethylsilane, TMSOTf, $MeNO_2$, -10°C, 74 %; b) O_3, MeOH, methyl sulfide, $NaBH_4$, 96 %; c) (iPr$_2$N)(NCCH$_2$CH$_2$O)PCl, iPr$_2$NEt, THF , RT, 1.5 h, 70 %.

C-nucleosidic fucose building block **4**, compatible with standard solid phase oligonucleotide synthesis, was prepared starting form commercially available α-(L)-fucose (**1**). Peracetylation, followed by a stereo selective allylation at low temperature, favoring the formation of the α-anomer, yielded compound **2**. Ozonolysis of **2** was achieved following the protocol developed by Uchiyam*a et al*.[1] Subsequent reduction of the carbonyl function led to alcohol **3**. Phosphitylation yielded the final phosphoramidite **4** in good yield (70 %). This route is very efficient with respect to the stereo selectivity for the substitution step with an overall yield of 50 % over 3 steps.

(2*S*,3*R*,4*R*,5*R*,6*S*)-2-(2-(((2-cyanoethoxy)(diisopropylamino)phosphanyl)oxy)ethyl)-6-methyltetrahydro-2*H*-pyran-3,4,5-triyl triacetate

The alcohol **3** (0.2 g, 0.63 mmol, 1 eq.) was dissolved in dry THF (12 mL) at RT. To this solution *N,N*-diisopropylethylamine (0.33 mL, 1.89 mmol, 3 eq.) was added followed by CEP-Cl (0.21 mL, 0.9 mmol, 1.5 eq.). The reaction mixture was stirred for 1.5 h at RT, quenched with $NaHCO_3$ sat. (10 mL). The aqueous phase was extracted three times with EtOAc (3x 20mL), the combined organic phases were dried over $MgSO_4$ and concentrated under reduced pressure. The crude product was purified by FC (hex/EtOAc 8/2 to 1/1) to yield 230 mg (70 %) of the title compound (**4**) as a clear oil.

R_f (HEx/EtOAc 2.5/7.5) = 0.84

MS *m/z* calculated for $C_{23}H_{39}N_2O_9P+H^+$: 519.2466, found 519.2466.

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.31 (ddd, *J* = 9.8, 5.6, 3.2 Hz, 1H, H-C2), 5.28 – 5.25 (m, 1H, H-C4), 5.20 – 5.15 (m, 1H, H-C3), 4.41 – 4.32 (m, 1H, H-C1), 4.01 – 3.92 (m, 1H, H-C5), 3.90 – 3.77 (m, 2H, CH_2-O), 3.71 (m, 2H, CH_2-CN), 3.64 – 3.53 (m, 2H, H-C8), 2.63 (t, *J* = 6.4 Hz, 2H, CH-N), 2.15 (s, 3H, Ac), 2.06 (s, 3H, Ac), 2.08 – 1.97 (m, 1H, H-C7), 2.01 (s, 3H, Ac), 1.87 – 1.76 (m, 1H, H-C7), 1.22 – 1.09 (m, 15H, CH_3-iPr, H-C6').

^{13}C NMR (101 MHz, $CDCl_3$) δ 170.86, 170.28 , 170.02, 169.98, 117.73, 70.73, 70.70, 69.38, 69.26, 68.78, 68.76, 68.27, 68.23, 66.05, 66.00, 60.13, 59.95, 59.89, 59.72, 58.62, 58.55, 58.44, 58.35, 43.29, 43.16, 27.43, 27.37, 24.81, 24.78, 24.74, 24.70, 22.66, 22.62, 21.39, 21.37, 20.88, 20.84, 20.57, 20.50, 19.72, 19.63, 16.08, 16.06.

^{31}P NMR (162 MHz, $CDCl_3$) δ 147.67, 147.62.

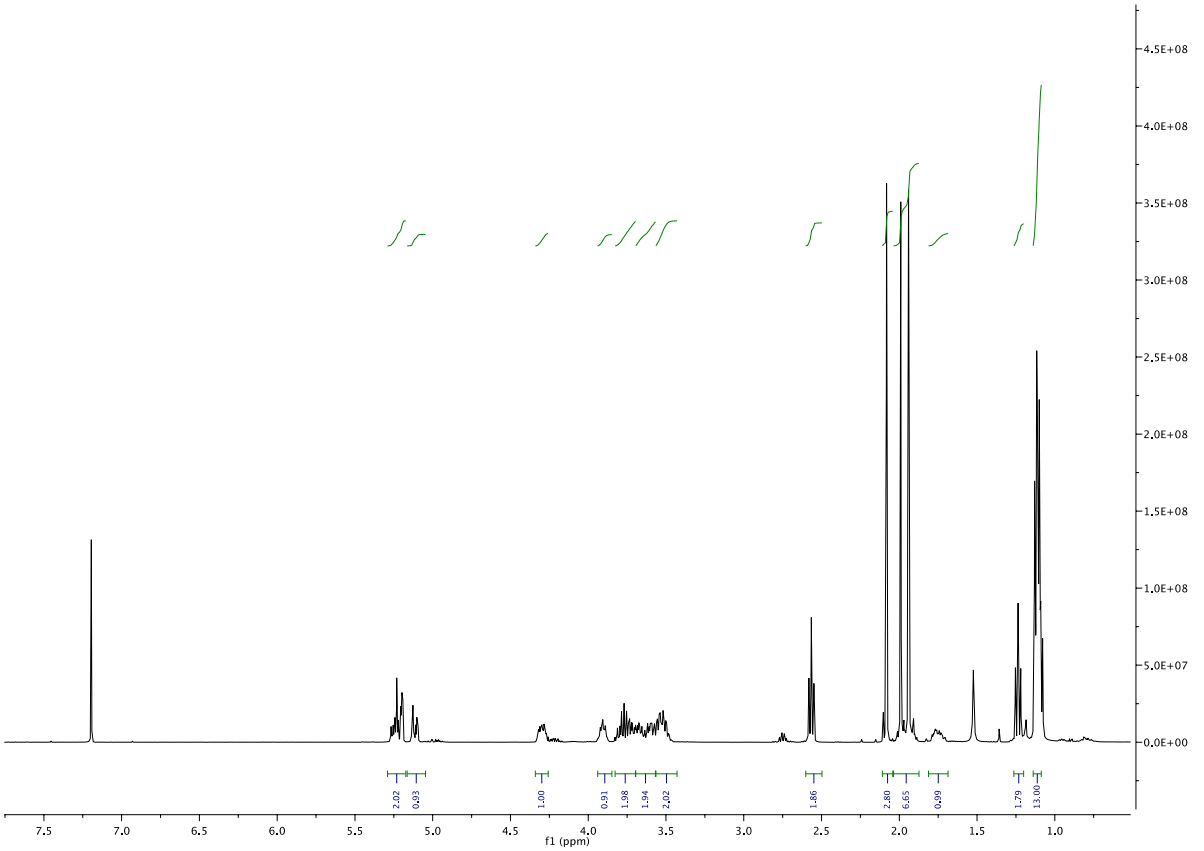


Figure S1. 1H NMR (400 MHz, CDCl3) spectra of **4**

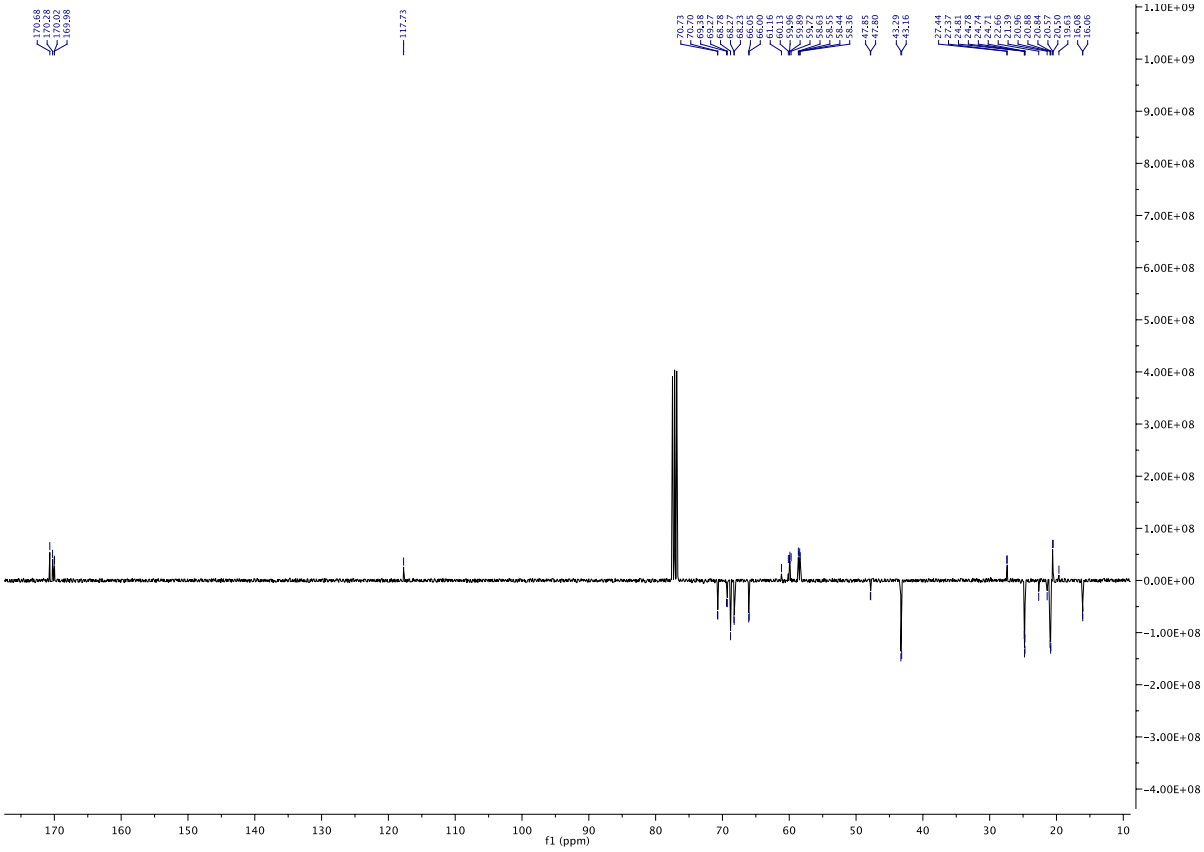


Figure S2. ^{13}C APT NMR (101 MHz, $CDCl_3$) spectra of **4.**

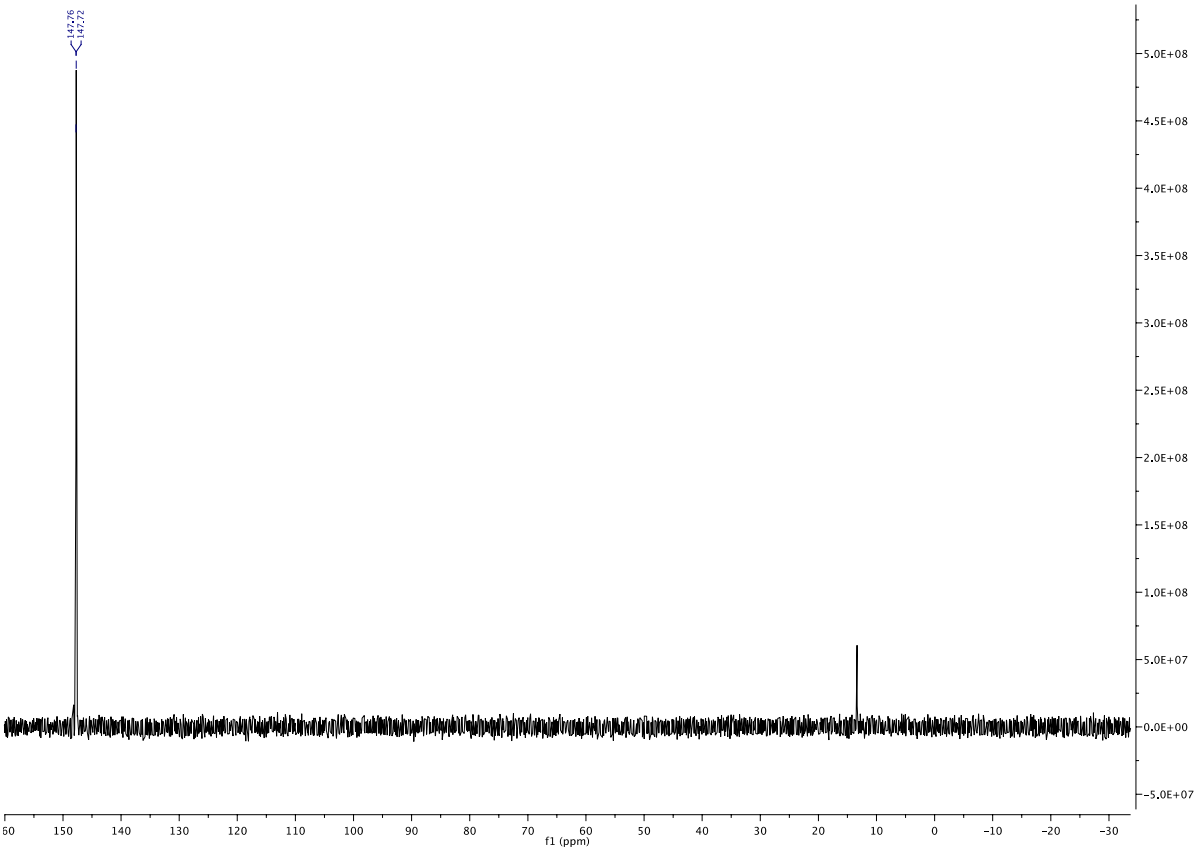


Figure S3. 31P NMR (162 MHz, CDCl3) spectra of **4.**

DNA synthesis

DNA synthesis of was performed on a 1.3 µmole scale on a *Pharmacia Gene Assembler Plus* DNA synthesizer following standard phosphoramidite protocols with 5-(ethylthio)-1H-tetrazole (0.25 M in CH_3CN) as the activator and coupling times of 6 min for the modified building blocks. Cleavage from the solid support and final deprotection was achieved with 30 % NH_4OH solution (55 °C, 12 h). Purification was performed by ion exchange HPLC with a DNAPAC PA200, 4 x 250 nm analytical column (*Dionex*), using

A) 25 mM 2-amino-hydroxymethyl-1,3-propanediol (Trizma) in H_2O, pH 8.0

B) 25 mM Trizma, 1.25 M NaCl in H_2O, pH 8.0

With a flow rate of 1 ml/min and a detection wavelength at 260 nm. For all oligonucleotides a gradient from 0 - 60% B over 40 min was used. The purified sequences were desalted by Sep Pak size exclusion. Sequences were characterized by mass spectrometry.

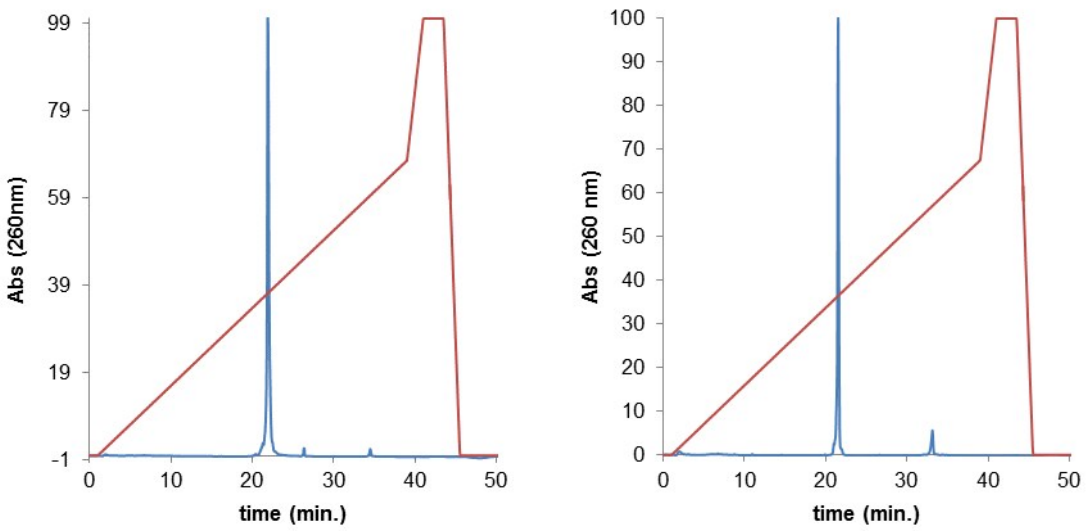


Figure S4. Analytical ion exchange HPLC traces of 3'-CGCTATPhenTTACGC-5' (left) and 3'-GCGTAAPhenATAGCG-FUC (right) for quality control. Conditions applied, 60% B in 40 min with a flow rate of 1 mL/min.

Strand	Sequence	m/z calcd	m/z found
E	3'-GCGTAAPhenATAGCG-FUC	4304.9	4304.8
B	3'-CGCTATPhenTTACGC-5'	3952.7	3952.7

Table S1. Summary of duplexes with their corresponding masses.

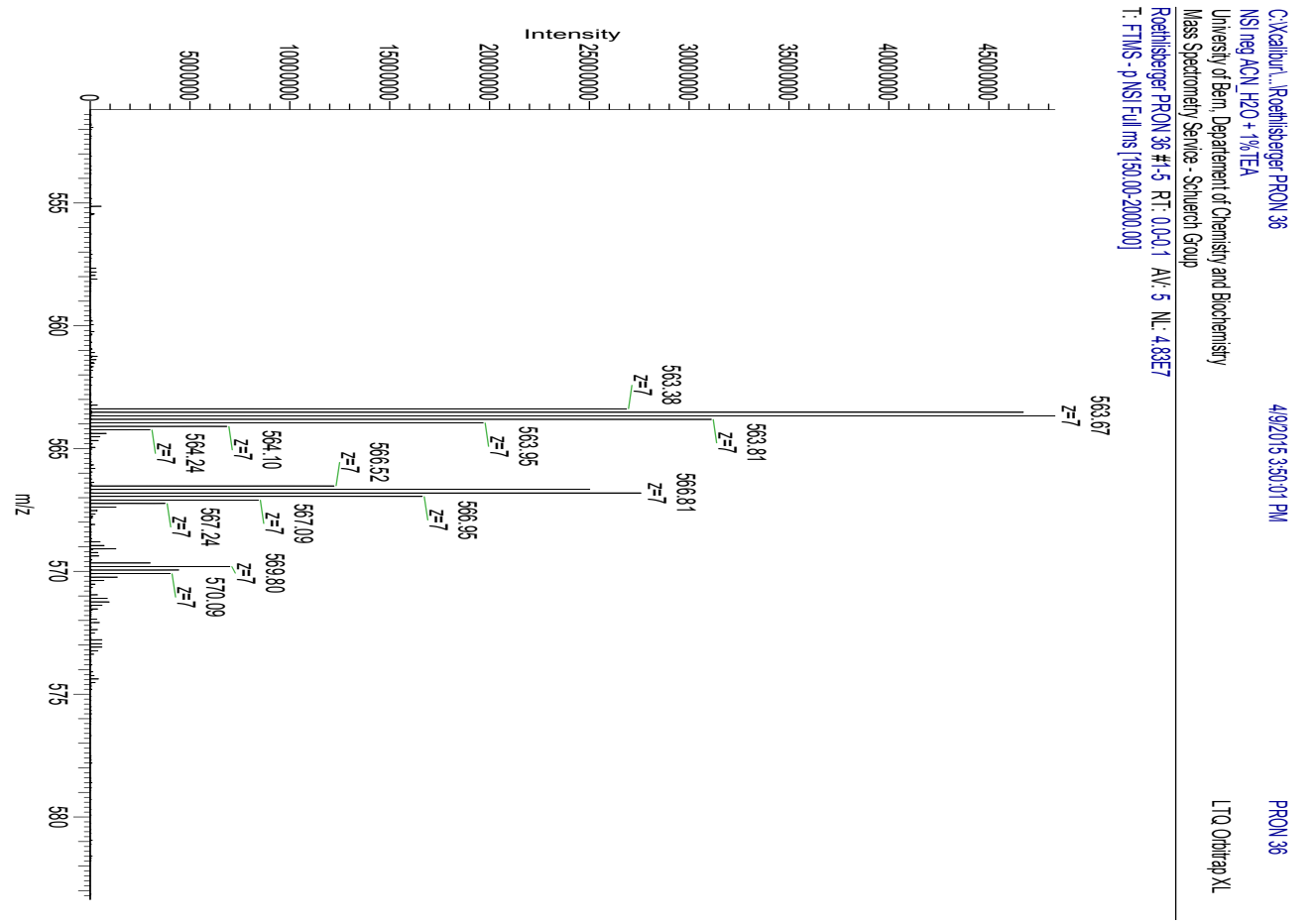


Figure S5. NSI MS of strand E.

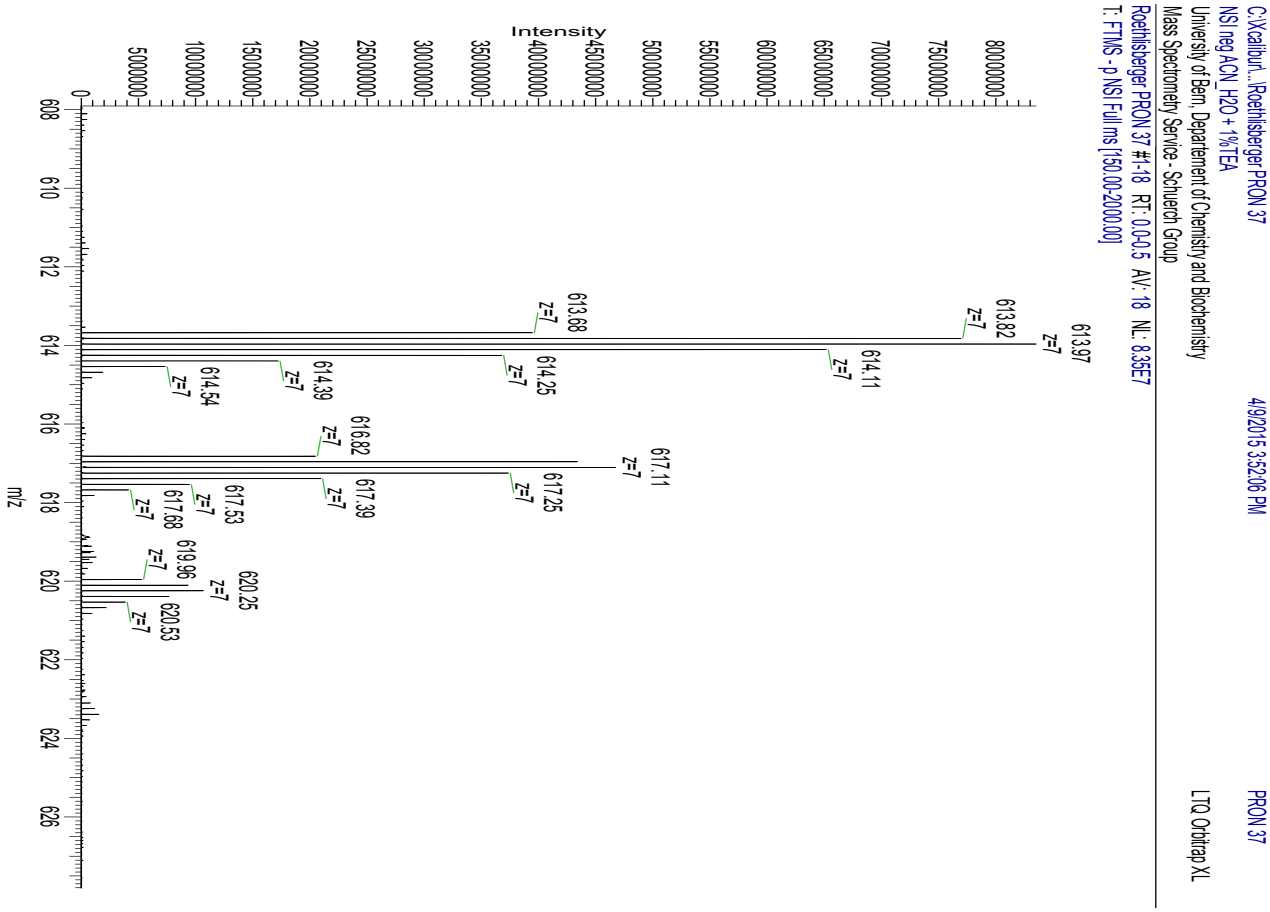


Figure S6. NSI-MS of strand B.

UV melting curves

UV melting curves were recorded on a Varian Cary 100-Bio UV/VIS spectrophotometer (Varian Inc.), equipped with a Peltier element at 260 nm with a rate of 0.5°C/min. A cooling/heating cycle in the temperature range 10–80 °C was applied. T_m values were obtained from the maxima of the first derivatives of the melting curves using WinUV software. To avoid evaporation of the solution, the sample solutions were covered with a layer of dimethylpolysiloxane. All measurements were carried out in a buffer consisting of 10 mM Na_2HPO_4, 150 mM NaCl, pH 7.0 at 4 µM total strand concentration.

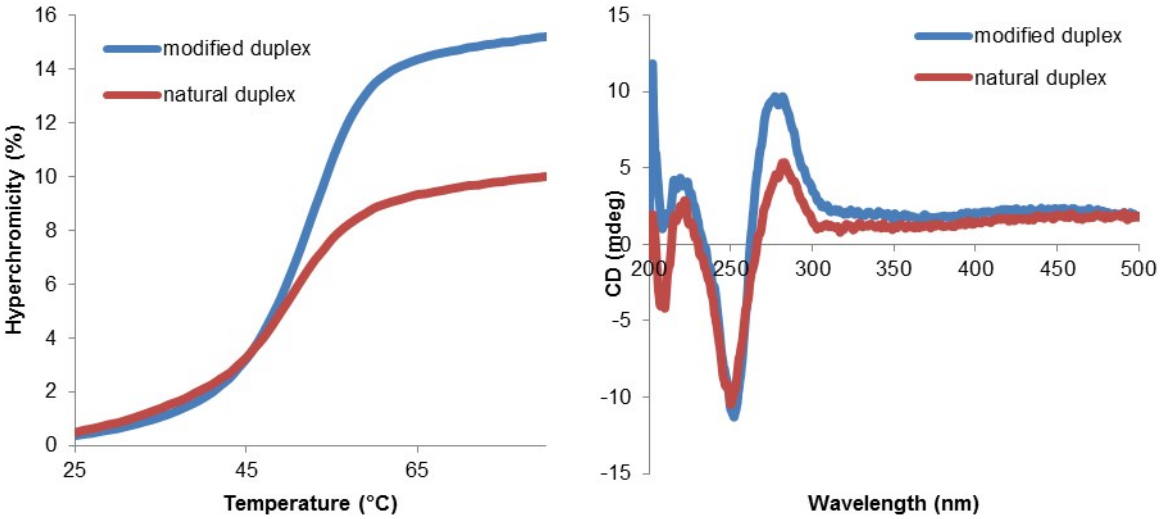


Figure S6. UV melting curves and CD spectra of the modified duplex and the natural DNA for comparison.

Duplex	Sequence	T_m(°C)
modified	^{FUC}U-GCGATAPhenAATGCG-3'	54.2 ± 0.9
	3'-CGCTATPhenTTACGC-5'	
natural	5'-GCGATAAAATGCG-3'	54.3 ± 0.5
	3'-C GCTATTTTACGC-5'	

Table S2. Tm of the modified duplex and the natural DNA for comparison in 10 mM Na_2HPO_4, 150 mM NaCl, pH 7.0 at 4.0 µM strand concentration.

CD spectra

CD spectra were measured on a JASCO J-715 spectropolarimeter at the temperature indicated using quartz cuvettes with a path length of 1 cm. All measurements were carried out in a buffer consisting of 10 mM Na_2HPO_4, 150 mM NaCl, pH 7.0 at 4.0 µM strand concentration.

DNA& LecB crystals

The complementary sequences were mixed in a 1/1 ratio (concentrations of the single strands were determined by OD). Duplexes were lyophilized and dissolved in milli-Q H_2O to reach a duplex concentration of 1.15 mM. Each duplex was annealed by heating to 60°C for 5 min and cooling slowly down overnight. After annealing, the duplex (c = 1.15 mmol/L in

milli-Q H_2O) was mixed in a 2.5 equivalent excess with a LecB monomer (c = 0.57 mmol/L) derived form *pseudomona aeruginosa* (provided by Ricardo Visini).[2]
CrystalQuick 96 well sitting drop vapor diffusion plates was prepared by pipetting 100 µl of the Crystal Screen, Index Screen or SaltRx screen reagents (*Hampton Research*) into the well block reservoirs. 1 microliter of the adjusted DNA/Protein solution was pipetted into each well and mixed with 1 microliter of the crystallization reagent. The well plates were sealed with clear sealing tape and stored at 18 °C. The plate was examined after each day under a microscope (Leica DM1000, 10 to 100x magnification). Conditions that produced single crystals were noted, crystals were fished and cryo-protected by adjusting the ethylene glycol content up to 30% and stored in liquid nitrogen (77K).

X-ray diffraction

X-ray data was collected at the PSI on the Xo6DA (PXIII) or the X10SA (PXII) macromolecular crystallography beamline. The (PXII) beamline receives light from a 2T superbend magnet, with a beam focus of 80 to 45 microns at the sample position. The spectral range of the beam is in-between 6.0 to 17.5 keV with a Flux of $5*10^{11}$ photons/sec/400 mA (at 12.4 keV). The PXII beamline posses a higher beamfocus of 50 to 10 microns with a energy resolution better than 0.02%. The spectral range of the beam is in-between 6.0 to 20 keV with a Flux of $2*10^{12}$ photons/sec/400 mA (at 12.4 keV). The crystals were mounted manually on a multi axis goniometer of either beamline under liquid nitrogen stream. Prior to data collecting, for each crystal a measuring strategy was developed to minimize the radiation damage. Therefore two orthogonal diffraction patterns were recorded and evaluated with labelit from XDS.[3] The recorded data set was integrated and scaled with XDS. Further refinements of the data were carried out using *PHENIX*.[4]

	5HCH
Crystal data	
Space Group	P $6_2$22
Cell constatns	a=60.45 Å b=60.45 Å c=204.83 Å α=90° β=90 ° γ=120°
Data collection	
Beamline	PXII
Wavelength (Å)	1.000
Resolution (Å)	46.62 - 2.90; 46.62 – 2.90
Completeness (%)	98.5 (46.62-2.90); 98.5 (46.62-2.90)
Structure refinement	
Refinement program	PHENIX
I/σ(I)	2.30 (at 2.91 Å)
R_{merge}	0.09
R, R_{free}	0.227, 0.257; 0.235, 0.257
R_{free} test set	538 reflections (11.12 %)
Wison B-factor (Å^2)	62.6
Outliers	0 of 5374 reflections
Total numbers of atoms	1409
Average B, all atoms (Å^2)	91.0

Table S3. Data and refinement statistics validated by the protein data bank evaluation tool.

Crystal structure analysis

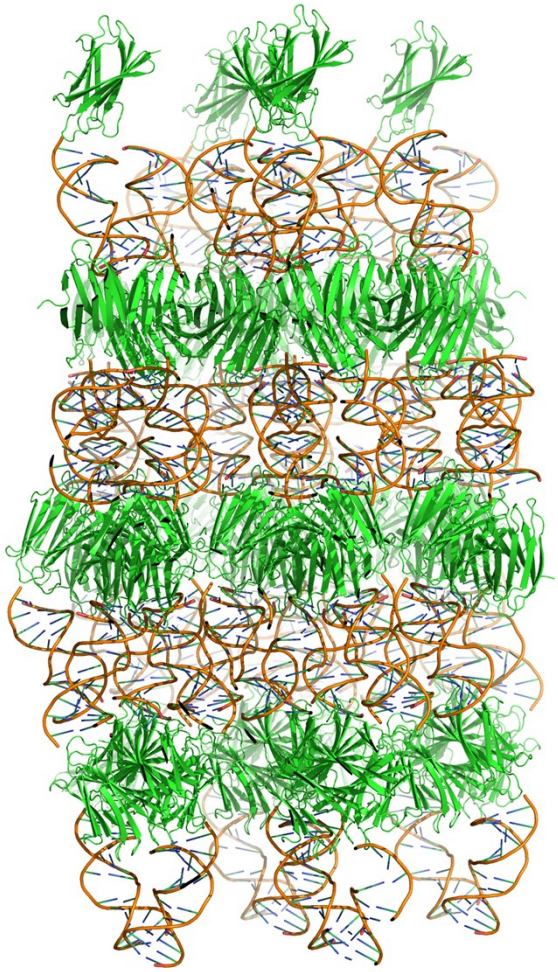


Figure S7. Crystal structure of the DNA-LecB complex represented as cartoon graphics with the neighboring units of the monomer in 50 Å radius

For structural analysis of the DNA from the crystal structure, the duplex was analyzed without the LecB. All structural parameters were extracted with the program x3DNA or determined by measurements performed with coot. The torsion angles were determined by the following atom positions: alpha O3'(i-1)-P-O5'-C5'; beta P-O5'-C5'-C4'; gamma O5'-C5'-C4'-C3'; delta C5'-C4'-C3'-O3'; epsilon C4'-C3'-O3'-P(i+1); zeta C3'-O3'-P(i+1)-O5'(i+1); chi for pyrimidines(Y) O4'-C1'-N1-C2; chi for purines(R) O4'-C1'-N9-C4.

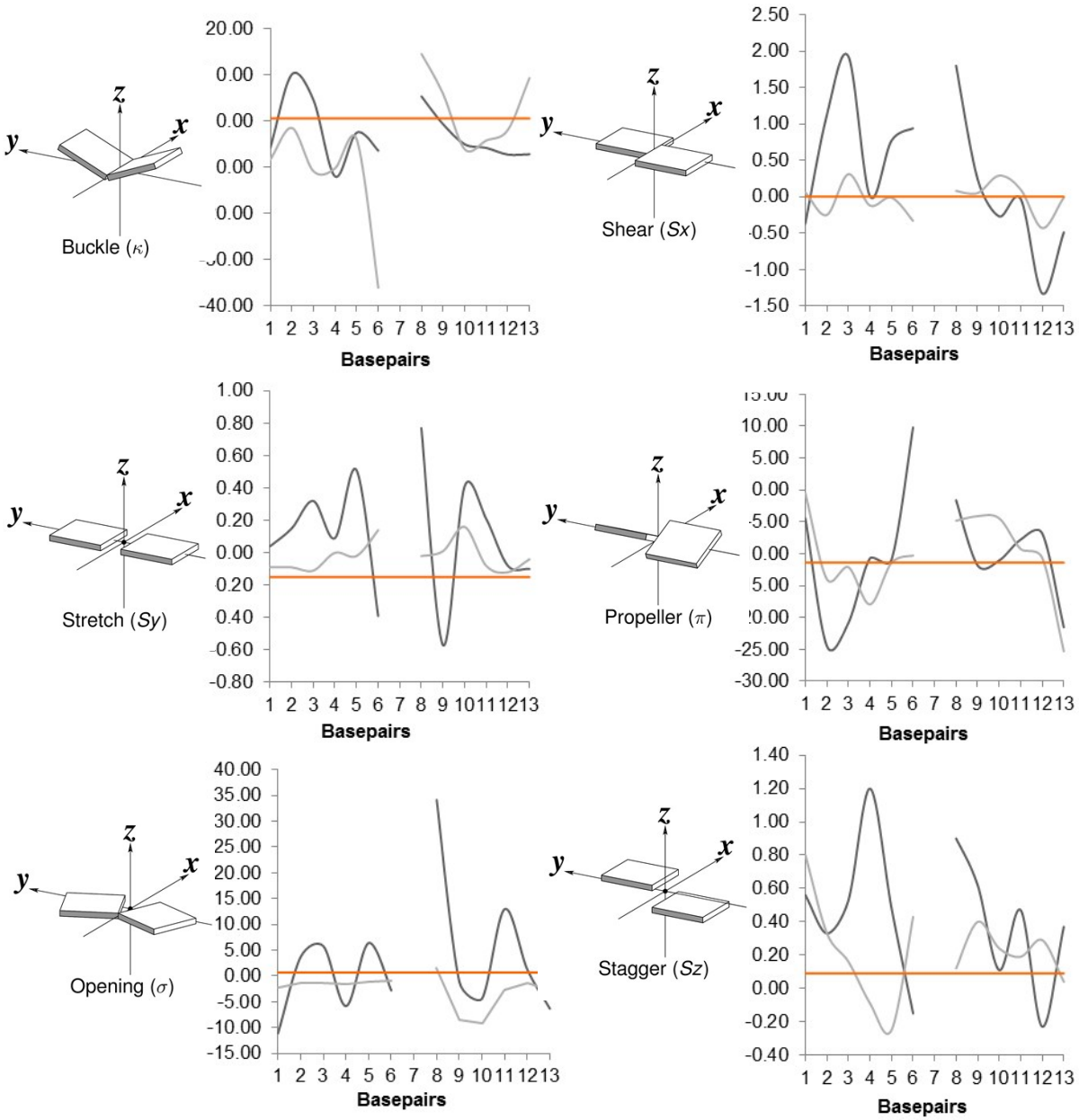


Figure S8. Structural analysis of base orientation within the DNA. The dark grey lines depict the analysis of the crystal structure, the light grey lines the analysis of the energy minimized structure and the straight red lines represent an average value of natural B-DNA.[5]

Strand /Nr	Base	α (°)	β (°)	γ (°)	β (°)	ε (°)	ζ (°)	χ (°)
B/13	C	---	---	38.4	95.4	-179.5	-80.2	-165.8
B/12	G	-171.9	-113.9	129.8	139.4	-155.8	-132.8	-131
B/11	C	-0.9	149.3	-0.2	143.6	-175.3	-178.2	-106.7
B/10	A	148.6	-101.3	163.1	117.3	-160.9	-84.2	-148.6
B/9	T	-60.8	173.3	32.7	84.9	-159.7	-63	-140.1
B/8	T	-77.5	170.8	56.3	88.8	-137.14	-141.34	-133.1
B/7	Phen	96.61	-134.17	156.63	97.4	-130.37	-72.69	-164.37
B/6	T	-70.09	154.8	94.9	151.7	-141.9	-164.5	-108.3
B/5	A	156.7	-150.2	156.5	151.3	-178.2	-100.7	-114.3
B/4	T	-52.6	-176.5	38.4	128.1	-175.2	-112.6	-125
B/3	C	-50.7	164.7	58.4	139.6	-150.4	-121.5	-99.9
B/2	G	-63.7	151.8	53.8	129	-149.1	-121.4	-105
B/1	C	-85.7	162.3	60	121.2	---	---	-131.9
E/26	G	-49.3	134.7	38.3	133.9	---	---	-108.3
E/25	C	-58.6	175.1	51.4	140.1	-124.3	-164.2	-100
E/24	G	29.4	153.1	-13.5	154.6	175.2	-103	-92.2
E/23	T	-57.6	167.6	53.1	135.2	179.4	-132.7	-110
E/22	A	-177.9	-173.1	155.5	111.7	-159.4	-97.1	-144.9
E/21	A	-91.38	170.5	78.5	156	162.7	-94	-100.3
E/20	Phen	34.87	155.7	-48.2	173.32	-137.2	-104.87	-88.64
E/19	A	-64.7	167.8	53.6	142.5	-141.83	173.37	-110
E/18	T	-60.1	169.9	45.5	94.2	-169.8	-92.8	-139.8
E/17	A	160.3	-107.2	152.5	106.8	-166.2	-86.1	-145
E/16	G	-48.8	155.4	71.3	144	-172.6	-153.6	-125.3
E/15	C	103.4	-155.1	-156.6	84.9	-149.3	-82.3	-152.9
E/14	G	---	---	109.6	120.1	-169	-107	-111.2

Table S3. Torsion angles of the DNA duplex

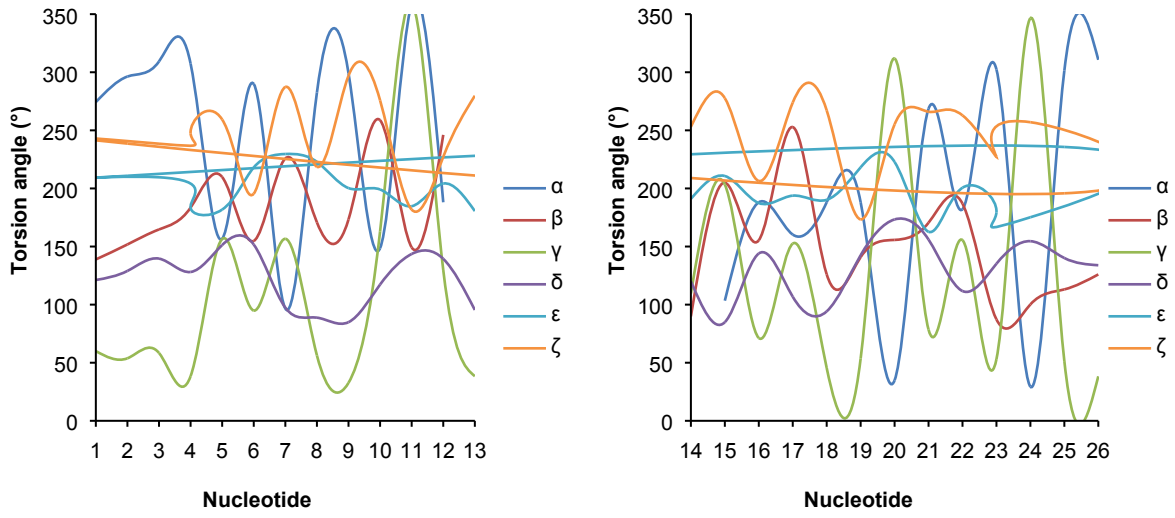


Figure S9. Measured torsion angles of the phosphate backbone of the DNA duplex from the crystal structure. Nucleotides 1 to 13 represent strand B and nucleotides 14 to 26 represent strand E.

The sugar pucker was determined by ν_0 C4'-O4'-C1'-C2'; ν_1 O4'-C1'-C2'-C3'; ν_2 C1'-C2'-C3'-C4'; ν_3 C2'-C3'-C4'-O4'; ν_4 C3'-C4'-O4'-C1'. The pseudorotaion phase angle (P) and the pseudorotation angle amplitude (ν_{max}) were determined after Equation 1 and Equation 2.

$$\tan P = \frac{(\nu_4 + \nu_1) - (\nu_3 + \nu_0)}{2\nu_2(\sin(36°) + \sin(72°))}$$

Equation 1. Definition of the pseudorotation phase angle P.[6]

$$\nu_{max} = \frac{\nu_2}{\cos(P)}$$

Equation 2. Definition of the pseudorotation angle amplitude (ν_{max}).[6]

Strand/Nr	Base	ν_0	ν_1	ν_2	ν_3	ν_4	ν_{max}	P	Puckering
B/13	C	12	-26.8	30.5	-24.4	8.1	30.5	356.2	C2'-exo
B/12	G	-31.8	42.4	-36.3	19	7.7	41.7	150.6	C2'-endo
B/11	C	-23.7	37.3	-36	23.3	0	38.1	161.3	C2'-endo
B/10	A	-32.9	30.5	-17.2	-1.1	21.3	32.7	121.6	C1'-exo
B/9	T	-19.9	-3.6	23.9	-36.1	35.5	37.2	50.1	C4'-exo
B/8	T	2.9	-22.9	33	-32.2	18.7	34	13.9	C3'-endo
B/7	Phen	-50.6	36.2	-10.5	-18	41.1	n.d.	n.d.	n.d.
B/6	T	-19.3	37.2	-39.8	29.8	-6.8	40.4	170.8	C2'-endo
B/5	A	-8.6	27.5	-34.8	30.8	-14.2	34.9	184.8	C3'-exo
B/4	T	-32.3	36.8	-27.3	9.6	14	36.3	138.8	C1'-exo
B/3	C	-21.3	33.2	-31.8	20.4	0.3	33.7	160.7	C2'-endo
B/2	G	-33.5	38.2	-28.3	9.9	14.6	37.6	138.7	C1'-exo
B/1	C	-32.5	32.7	-20.9	2.9	18.5	33.5	128.4	C1'-exo
E/26	G	-30.3	38.2	-31.3	14.8	9.4	37.4	146.8	C2'-endo
E/25	C	-31.2	42.2	-36.5	19.6	7	41.5	151.5	C2'-endo
E/24	G	-5.6	26.1	-35.5	33.2	-17.6	36.1	189.9	C3'-exo
E/23	T	-29.4	38.3	-32.2	16.1	8.1	37.6	149	C2'-endo
E/22	A	-32	25.7	-10.5	-7.4	24.8	31	109.8	C1'-exo
E/21	A	7.4	15	-30.1	35.2	-27.1	35	210.5	C3'-exo
E/20	Phen	2.6	29.4	-48.6	52.9	-33.4	n.d.	n.d.	n.d.
E/19	A	-24.2	37.2	-35.3	22.3	0.9	37.6	159.8	C2'-endo
E/18	T	-30.5	12.7	8.1	-25.8	35.7	35	76.6	O4'-endo
E/17	A	-37.5	28	-9	-12.1	31.1	36.4	104.4	O4'-endo
E/16	G	-25.1	38.8	-37.1	23.5	0.7	39.4	160.2	C2'-endo
E/15	C	1	-23.9	36.2	-36.6	22.6	37.9	17.1	C3'-endo
E/14	G	-33.9	33.5	-20.7	1.9	20	34.7	126.7	C1'-exo

Table S4. Sugar Puckers of the DNA duplex.

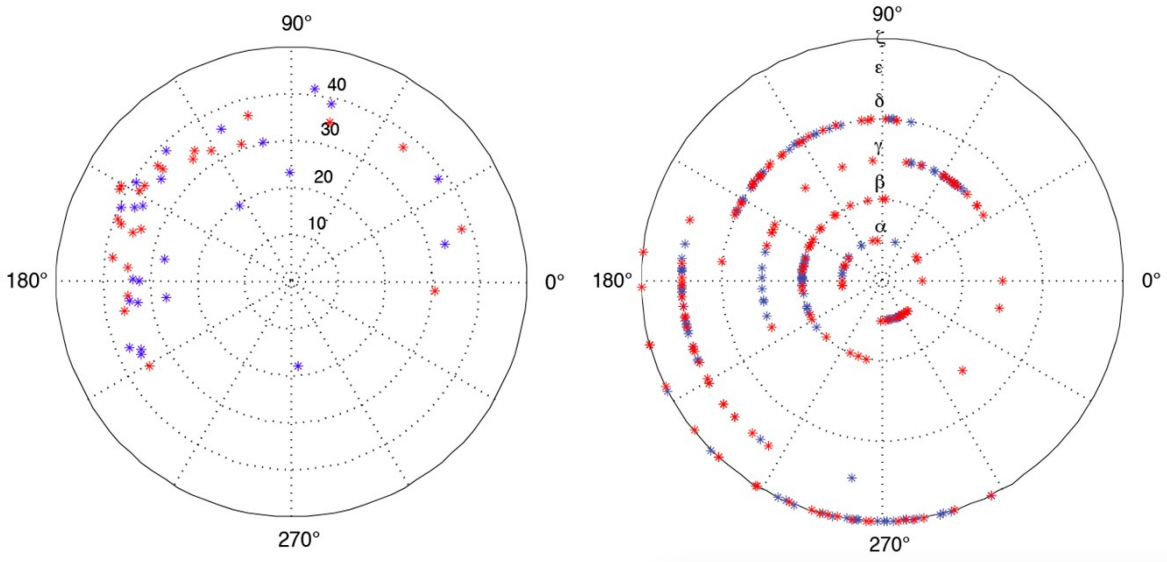


Figure S10. Calculated sugar puckers (left) and torsion angles (right) of the DNA duplex represented in a polar plot. The red crosses (*)represent the angles and sugar puckers of the crystal structure and the blue crosses (*) represent the geometry-optimized angles and sugar puckers.

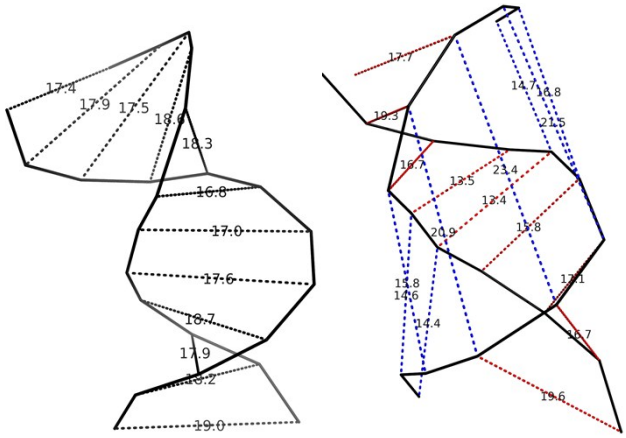


Figure S11. Representation of phosphate distances between the adjacent nucleotides. Measurements of the minor groove width are represented in red and the major groove width in blue.

Molecular modeling

Optimization of the X-ray structure was performed using Gromacs.[7] For optimization only the DNA duplex without the FUC analog was considered. The duplex was placed in a cubic cell of 0.5 x 0.5 x0.5 nm, then 10372 water molecules of SPC/E model were added. The total charge of the system was adjusted to zero by the addition of 24 sodium atoms to counterbalance the negative charges of each phosphate of the duplex backbone. The potential energy was minimized by steepest descent algorithms. Then a MD trajectory of 1 ns with restricted positions of all the nucleotide atoms was calculated (position restrained). The force field AMBER 03 was used.[8] The procedure for the MD trajectory of the unrestricted positions was the same as outlined for the restricted calculation. The length of the trajectory was prolonged to 10 ns in order to collect a more comprehensive dataset.

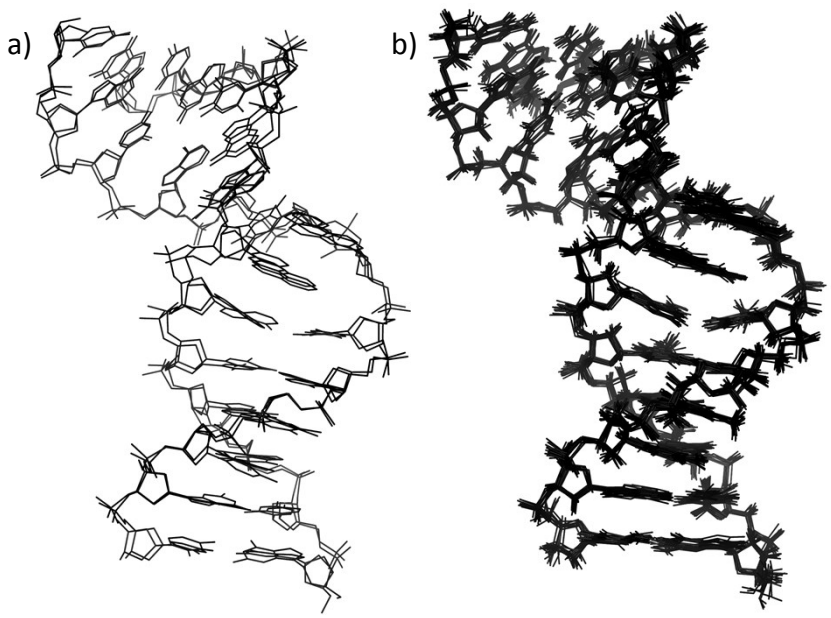


Figure S12. a) Overlay of the crystal structure and geometry optimized structures; b) Overlay of 10 structures of the 1 ns trajectory of the molecular dynamics calcualtion.

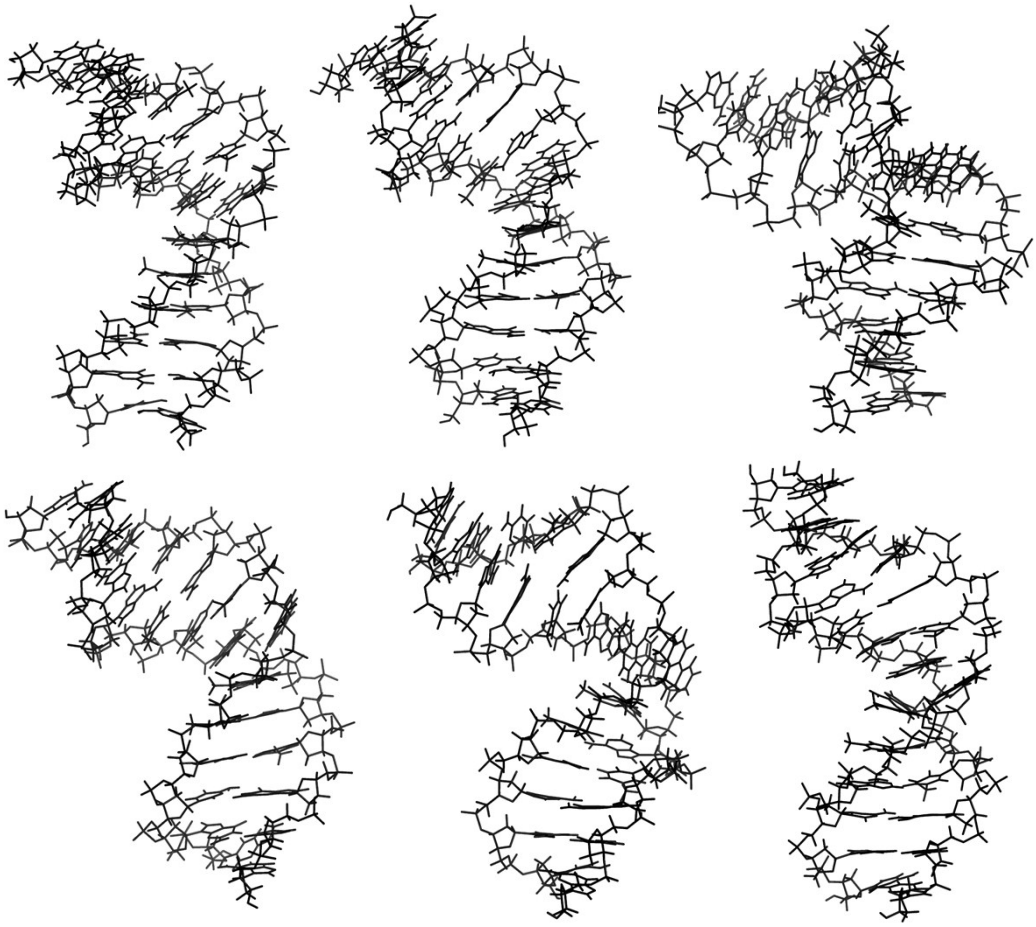


Figure S13. .A set of 6 structures of the 10 ns trajectory obtained by a molecular dynamic simulation.

References

1. T. Uchiyama, T. J. Woltering, W. C. Wong, C. C. Lin, T. Kajimoto, M. Takebayashi, G. WeitzSchmidt, T. Asakura, M. Noda and C. H. Wong, *Bioorganic & Medicinal Chemistry*, 1996, **4**, 1149-1165.
2. H. Funken, K.-M. Bartels, S. Wilhelm, M. Brocker, M. Bott, M. Bains, R. E. W. Hancock, F. Rosenau and K.-E. Jaeger, *PLoS ONE*, 2012, **7**, e46857.
3. W. Kabsch, *Journal of Applied Crystallography*, 1988, **21**, 67-71.
4. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallographica Section D*, 2010, **66**, 213-221.
5. 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**, 229-237.
6. C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, 1972, **94**, 8205-8212.
7. H. J. C. Berendsen, D. van der Spoel and R. van Drunen, *Computer Physics Communications*, 1995, **91**, 43-56.
8. Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. M. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. M. Wang and P. Kollman, *Journal of Computational Chemistry*, 2003, **24**, 1999-2012.