A novel naphthyridine tetramer that recognizes tandem G–G mismatches by the formation of an interhelical complex

Yihuan Lu, Chikara Dohno* and Kazuhiko Nakatani*

Supplementary data

Contents

1. Experimental

General

Reagents and solvents were purchased from standard suppliers and used without further purification. Oligodeoxynucleotides (ODNs) were purchased from Eurofins Genomics. ^{1}H-NMR and ^{13}C-NMR spectra were measured using JEOL JNM-ECA600. Chemical shifts were reported in ppm relative to the residual solvent peak as an internal standard; 7.26 ppm (^{1}H NMR) in chloroform-d, 77.16 ppm (^{13}C NMR) in chloroform-d, and 39.52 ppm (^{13}C NMR) in DMSO-d_6.

Polyacrylamide gel electrophoresis analysis

Ligand was added to annealed solution of ODN (1 μM or less) in 10 mM Na cacodylate buffer (pH 7.0) containing 100 mM NaCl, 0.1% Tween20, and 8% glycerol. The samples were incubated at 37°C for 3 h and were electrophoresed on native polyacrylamide gel (16%, 19:1) in ice bath at 200 V. Gels were stained with SYBR Gold except for when FAM-labeled ODNs were used.

Chemical probing of flipped out cytosine in dCGGG/dCGGG

Samples (150 μL) containing **ODN1-a1**/**ODN1-a2** (10 μM/10 μM), 100 mM NaCl, 2.8 M NH_2OH (pH 6.0) in the absence/presence of ligand (0, 80 μM NCD or 40 μM *p*-NCTB) was cooled to 0 °C and kept for 2 d. The reactions were monitored by reversed phase HPLC, and mixtures of hydroxylamine–ODN adducts were collected by HPLC. After lyophilization, the resulting samples were dissolved in 10% piperidine (150 μL) and incubated at 90 °C for 30 min. After removal of solvent and three times co-evaporation using MilliQ water, 50 μL alkaline phosphatase (AP, NEB) was added and the samples were incubated at 37 °C for 2 h. The reactions were monitored by reversed phase HPLC. Co-injection with standard **ODN1–a3**, **a4**, **a5**, and **a6** (**ODN1-a3**, 5'-d(GCAA)-3'; **ODN1-a4**, 5'-d(GGGAAGC)-3'; **ODN1-a5**, 5'-d(GCTT)-3'; **ODN1-a6**, 5'-d(GGGTTGC)-3') was performed to identify the products.

HPLC analyses were performed with JASCO HPLC system equipped with CHEMCOBOND 5-ODS-H column (4.6 x 150 mm, Chemco). Solvent conditions: CH_3CN was increased from 5% to 15% in 35 min while 0.1 M TEAA was decreased from 95% to 85% at 40 °C at a flow rate of 1.0 mL/min. The eluates were detected at 254 nm.

CSI-tof-MS analysis

Cold-spray ionization time-of-flight mass (CSI-TOF-MS) measurement was performed on JEOL AccuTOF JMS-T100N mass spectrometer in the negative mode (orifice 1 voltage = –60 V). Samples containing 10 μM ODN with and without ligand in 100 mM NH_4OAc and 50% MeOH were injected at a flow rate of 20 μL/min. Spray temperature was fixed at –10 °C. Nitrogen gas was used as a desolvation gas as well as a nebulizer.

Estimation of binding stoichiometry by HPLC analysis of isolated ligand–DNA complex

Sample solution (100 μL) containing 10 μM ODN **ODN1** and 40 μM *p*-NCTB was prepared according to the previous method and was subjected to native PAGE (19:1, 16%) in ice bath at 200 V for 90 min. Band corresponding to the complex

was cut from the gel, and ODN and ligand in the band were extracted twice using TE buffer. The resulting mixture was analyzed by reversed phased HPLC. Mixture of 5 μM ODN and 5 μM ligand was used as a standard for determination of molar ratio of ODN: ligand. Equations used for the calculations are as following.

$$\frac{C_{ODN}}{C_{Ligand}} = \frac{A_{ODN}}{A_{Ligand}} \times \frac{C^S_{ODN}}{C^S_{Ligand}} \times \frac{A^S_{Ligand}}{A^S_{ODN}}$$

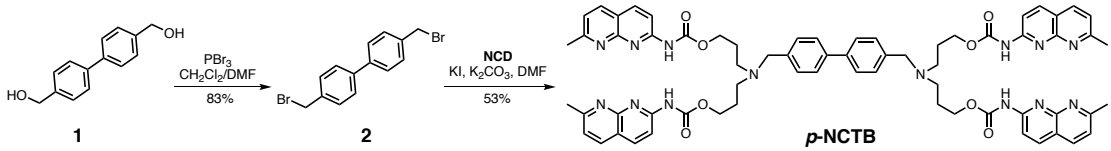
Where $\frac{C_{ODN}}{C_{Ligand}}$ and $\frac{A_{ODN}}{A_{Ligand}}$ are concentration ratio and peak area ratio of ODN and ligand in the tested sample; $\frac{C^S_{ODN}}{C^S_{Ligand}}$ and $\frac{A^S_{Ligand}}{A^S_{ODN}}$ are concentration ratio and peak area ratio of ODN and ligand in the standard sample.

HPLC conditions: Samples were eluted at a flow rate of 1.0 mL/min at 40 °C through CHEMCOBOND 5-ODS-H column (4.6 × 150 mm). 20 min linear gradient from 0 to 25% acetonitrile in 0.1 M TEAA, followed by isocratic wash with distilled water for 10 min, then 20 min linear gradient from 0 to 40% acetonitrile in 0.1% TFA. ODN was detected at 254 nm and ligand was detected at 320 nm.

Circular dichroism (CD) measurements

CD melting analysis of DNAs was performed on J-1500 Circular Dichroism spectrometer (JASCO). Microcal Origin 6.0 was used to perform derivative calculation. Melting analysis based on UV absorbance was not successful due to UV absorption change of *p*-NCTB upon increasing temperature. Each sample (350 μL) contained 3 μM dCGGG/dCGGG (i.e. 3 μM for **ODN1**, 1.5 μM for **ODN1D**) and *p*-NCTB (0, 12 μM) in 10 mM Na cacodylate buffer (pH 7.0), 100 mM NaCl and 0.1% Tween20. CD was measured from 20 to 80 °C with an increase of 1 °C/min.

2. Synthesis of *p*-NCTB



Scheme S1 Synthesis scheme of *p*-NCTB

4,4'-bis(bromomethyl)-1,1'-biphenyl (2). *p*-biphenyldimethanol **1** (0.642 g, 3 mmol) was dissolved in CH_2Cl_2/DMF (5 mL/ 5 mL) and kept at 0 °C. PBr_3 (0.569 mL, 6 mmol) was slowly added and the mixture was stirred at 0 °C for 0.5 h followed by another 0.5 h at ambient temperature. The mixture was diluted with saturated aqueous $NaHCO_3$ and extracted using $CHCl_3$. The compound **2** was purified using silica gel column chromatography (eluent: CH_2Cl_2) as white solid (0.852 g, 83%): ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.56 (d, J = 8.2 Hz, 4H), 7.47 (d, J = 8.2 Hz, 4H), 4.55 (s, 4H); ^{13}C NMR (151 MHz, DMSO-d_6) δ 139.4, 137.5, 130.0, 127.0, 34.3; EI-MS: M^+ 338, 340, 342.

***p*-NCTB**. To a solution of NCD (0.201 g, 0.4 mmol) in 8 mL DMF was added compound **2** (0.061 g, 0.18 mmol), KI (0.066 g, 0.4 mmol) and K_2CO_3 (0.055 g, 0.4 mmol). The mixture was stirred overnight at ambient temperature. The solution was diluted with saturated aqueous NaCl and extracted using $CHCl_3$. The desired compound *p*-NCTB (0.113 g, 53%) was obtained as pale yellow solid after purification with silica gel column chromatography (eluent: $CHCl_3$:MeOH= 20:1). Purity of *p*-NCTB was confirmed by NMR and HPLC with CHEMCOBOND 5-ODS-H column (4.6 x 150 mm, Chemco) (Fig. S9a). Solvent conditions: 20 min linear gradient from 0 to 40% acetonitrile in 0.1% TFA at 40 °C at a flow rate of 1.0 mL/min: ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.27 (d, J = 8.2 Hz, 4H), 8.07 (dd, J = 8.2, 2.7 Hz, 4H), 7.93-7.91 (m, 8H), 7.37 (dd, J = 7.9, 1.7 Hz, 4H), 7.21 (dd, J = 8.5, 2.7 Hz, 4H), 7.15-7.14 (m, 4H), 4.27-4.25 (m, 8H), 3.44 (s, 4H), 2.72 (d, J = 2.7 Hz, 12H), 2.46-2.44 (m, 8H), 1.84-1.82 (m, 8H); ^{13}C NMR (151 MHz, $CDCl_3$) δ 163.0, 154.7, 153.6, 153.3, 139.5, 139.1, 138.1, 136.4, 129.1, 126.8, 121.2, 118.0, 112.8, 64.0, 58.4, 49.8, 26.7, 25.7; HRMS (ESI) *m/e* calcd for $C_{66}H_{69}N_{14}O_8^+$ $[M+H]^+$ 1185.5417, found 1185.5427.

3. ODN sequences used in this study

Table S1 ODN sequences used in this study

ODNs	Sequence (5' to 3')
ODN1	GAGTCGGGACTCTTTTGAGTCGGGACTC
ODN2	TCAACGGTTGA
ODN3	CACTGAGTCGGGACTCACTGTTTTCAGTGAGTCGGGACTCAGTG
ODN-c1	CGGCGGCGGCGGTTTCCGCCGCCGCCG
ODN-c2	GAGTCTTGACTCTTTTGAGTCAAGACTC
ODN1-a1	GCAACGGGAAGC
ODN1-a2	GCTTCGGGTTGC
ODN1-a3	GCAA
ODN1-a4	GGGAAGC
ODN1-a5	GCTT
ODN1-a6	GGGTTGC
ODN1D	GAGTCGGGACTCTTTTGAGTCGGGACTCTTTTTTGCAACGGGAA GCTTTTGCTTCGGGTTGC
ODN1-FAM	GAGTCGGGACTCTTTTGAGTCGGGACTC-FAM
ODN1D-FAM	GAGTCGGGACTCTTTTGAGTCGGGACTCTTTTTTGCAACGGGAA GCTTTTGCTTCGGGTTGC-FAM
ODN1-1T	GAGTTGGGACTCTTTTGAGTCGGGACTC
ODN1-5T	GAGTCGGGACTCTTTTGAGTTGGGACTC
ODN1-1T5T	CGAGTTGGGATCTCTTTTGAGATTGGGACTCG
ODN1-4T	GAGTCGGTACTCTTTTGAGTCGGGACTC
ODN1-8T	GAGTCGGGACTCTTTTGAGTCGGTACTC
ODN1-2T	GAGTCTGGACTCTTTTGAGTCGGGACTC
ODN1-6T	GAGTCGGGACTCTTTTGAGTCTGGACTC
ODN1-3T	GAGTCGTGACTCTTTTGAGTCGGGACTC
ODN1-7T	GAGTCGGGACTCTTTTGAGTCGTGACTC
ODN1-2A	GAGTCAGGACTCTTTTGAGTCGGGACTC
ODN1-6A	GAGTCGGGACTCTTTTGAGTCAGGACTC
ODN1-4A	GAGTCGGAACTCTTTTGAGTCGGGACTC
ODN1-8A	GAGTCGGGACTCTTTTGAGTCGGAACTC
ODN1D-2T	GAGTCGGGACTCTTTTGAGTCGGGACTCTTTTTTGCAACTGGAA GCTTTTGCTTCGGGTTGC
ODN1D-3T	GAGTCGGGACTCTTTTGAGTCGGGACTCTTTTTTGCAACGTGAA GCTTTTGCTTCGGGTTGC
ODN1D-4T	GAGTCGGGACTCTTTTGAGTCGGGACTCTTTTTTGCAACGGTAA GCTTTTGCTTCGGGTTGC
ODN1D-6T	GAGTCGGGACTCTTTTGAGTCGGGACTCTTTTTTGCAACGGGAA GCTTTTGCTTCTGGTTGC
ODN1D-7T	GAGTCGGGACTCTTTTGAGTCGGGACTCTTTTTTGCAACGGGAA GCTTTTGCTTCGTGTTGC
ODN1D-8T	GAGTCGGGACTCTTTTGAGTCGGGACTCTTTTTTGCAACGGGAA GCTTTTGCTTCGGTTTGC
ODN1D-2A2	GAGTCAGGACTCTTTTGAGTCGGGACTCTTTTTTGCAACAGGAA GCTTTTGCTTCGGGTTGC

ODN1D-3A2	GAGTCGAGACTCTTTTGAGTCGGGACTCTTTTTTGCAACGAGAA GCTTTTGCTTCGGGTTGC
ODN1D-4A2	GAGTCGGAACTCTTTTGAGTCGGGACTCTTTTTTGCAACGGAAA GCTTTTGCTTCGGGTTGC
ODN1D-6A2	GAGTCGGGACTCTTTTGAGTCAGGACTCTTTTTTGCAACGGGAA GCTTTTGCTTCAGGTTGC
ODN1D-7A2	GAGTCGGGACTCTTTTGAGTCGAGACTCTTTTTTGCAACGGGAA GCTTTTGCTTCGAGTTGC
ODN1D-8A2	GAGTCGGGACTCTTTTGAGTCGGAACTCTTTTTTGCAACGGGAA GCTTTTGCTTCGGATTGC
ODN1D-AT	GAGTCGGAACTCTTTTGAGTTGGGACTCTTTTTTGCAACGGAAA GCTTTTGCTTTGGGTTGC
ODN1D-TA	GAGTCGGTACTCTTTTGAGTAGGGACTCTTTTTTGCAACGGTAAGCTT TTGCTTAGGGTTGC
ODN1D-CG	GAGTCGGCACTCTTTTGAGTGGGGACTCTTTTTTGCAACGGCAAGCTT TTGCTTGGGGTTGC
ODN-D-CGG	GAGTCGGACTCTTTTGAGTCGGACTCTTTTTTGCAACGGAAGCTTTTG CTTCGGTTGC
ODN-CGG	GAGTCGGACTCTTTTGAGTCGGACTC

4. Hydrogen bonding pattern between guanine and NCD molecules

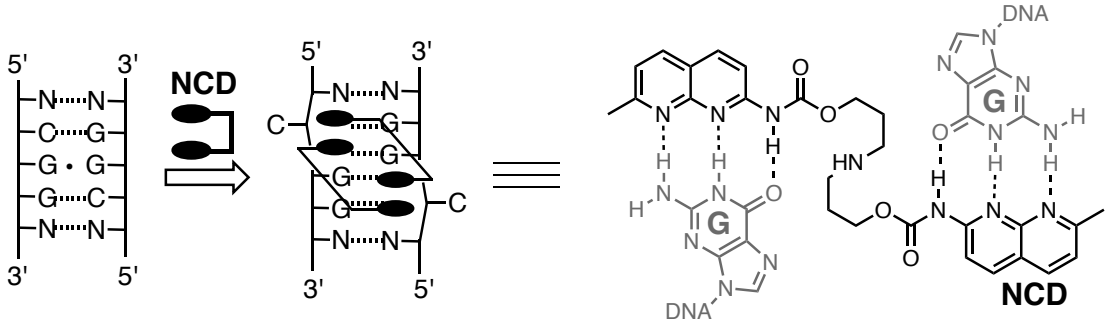


Fig. S1 Hydrogen bonding pattern between guanine and NCD

5. Putative secondary structures of DNA repeat sequences

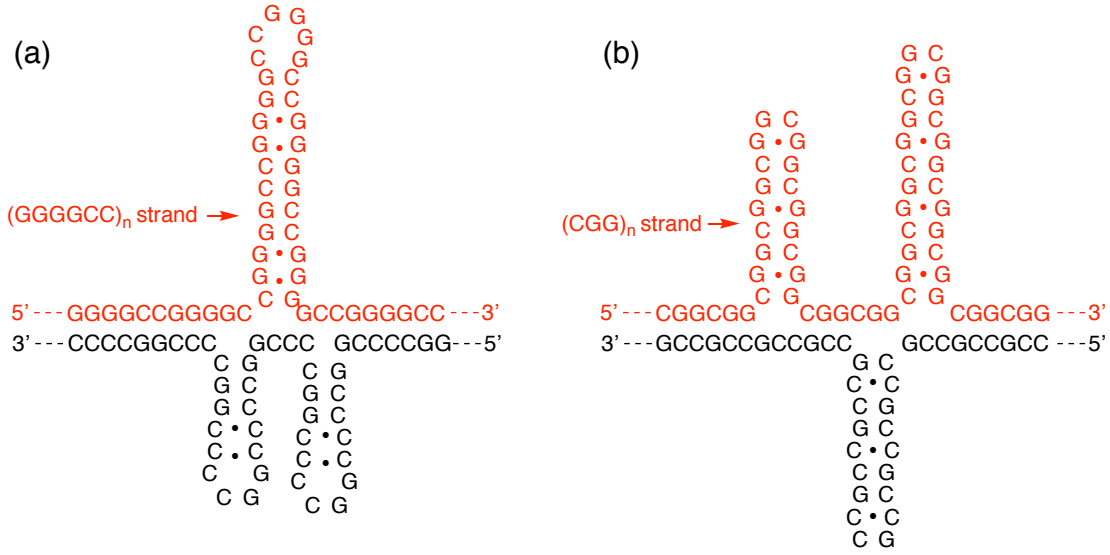


Fig. S2 Putative secondary structures of DNA GGGGCC repeat (a) and CGG repeat (b) sequences bearing tandem and separated G–G mismatches. Formation of the tandem short slip out structures has been previously suggested for CAG/CTG repeat.[9c]

6. Linker difference in *Z*-NCTS and *p*-NCTB

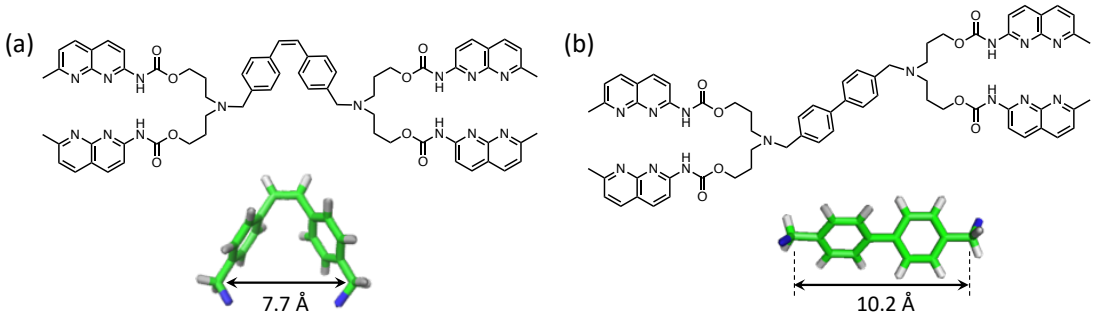


Fig. S3 Chemical structures of *Z*-NCTS (a) and *p*-NCTB (b), and comparison of linker length (distance between two methylene groups). Length of *Z*-stilbene was obtained from reported binding complex structure; *p*-biphenyl linker was optimized with AMBER* force field in water using Maestro 11.5.011 (Schrödinger software) before length measurement.

7. Binding assay of tested ligands to control ODNs

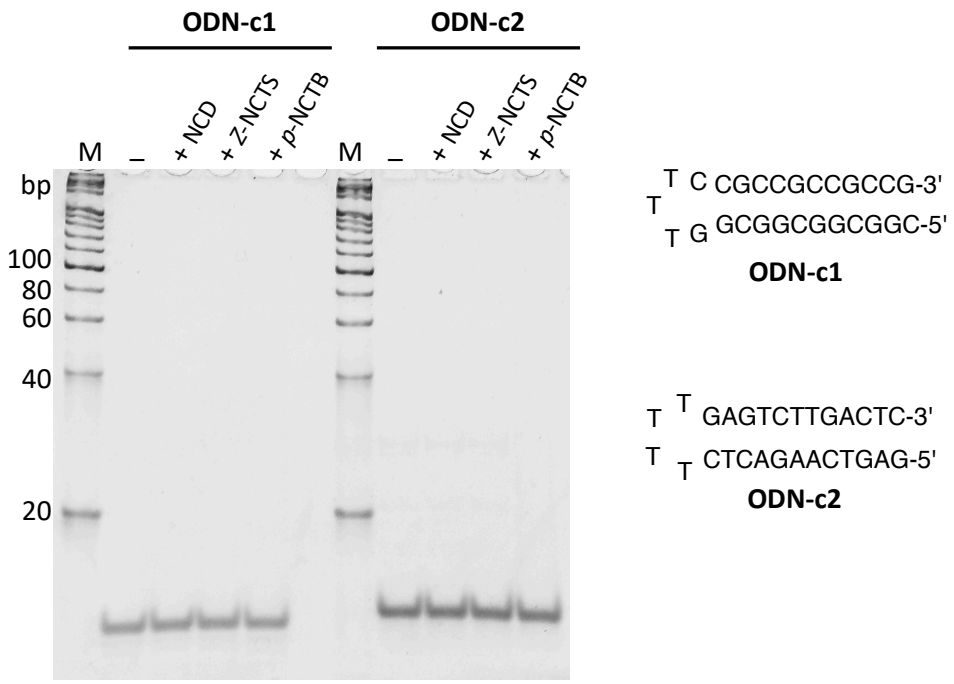


Fig. S4 Native PAGE analysis of tested ligands (NCD: 8 μM, *Z*-NCTS and *p*-NCTB: 4 μM) binding to two control sequences (1 μM). One is full complementary hairpin ODN containing continuous CGG and CCG tracts (**ODN-c1**), and another is full complementary hairpin ODN containing random sequence (**ODN-c2**). Lane M: DNA marker (20 bp). ODNs in the absence and presence of ligand were electrophoresed on native polyacrylamide gel (16%, 19:1) in ice bath at 200 V for 90 min.

8. Melting profiles of ODNs containing dCGGG/dCGGG in the absence/presence of ligand

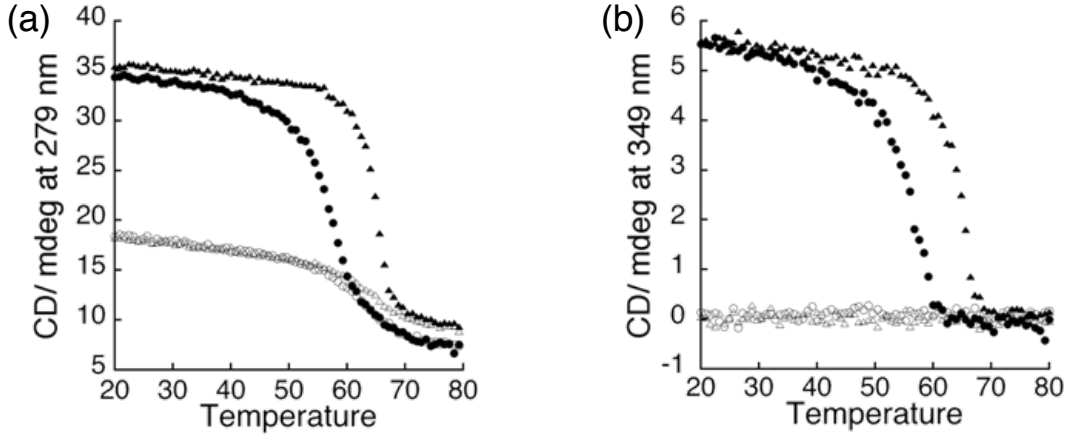
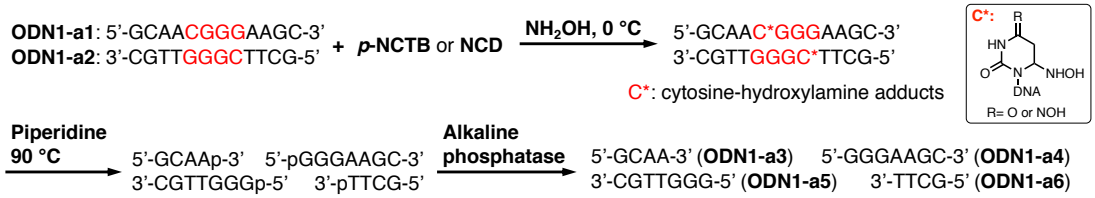


Fig. S5 Thermal melting curve of **ODN1** (circle) and **ODN1D** (triangle) in the absence (open) and presence of *p*-NCTB (solid) using CD values at 279 nm (a) and 349 nm (b). CD was measured from 20 to 80 °C with an increase of 1 °C/min.

9. Chemical probing of flipped-out cytosine in dCGGG/dCGGG



Scheme S2. Hydroxylamine-induced cleavage at cytosine in dCGGG/dCGGG in the presence of *p*-NCTB or NCD.

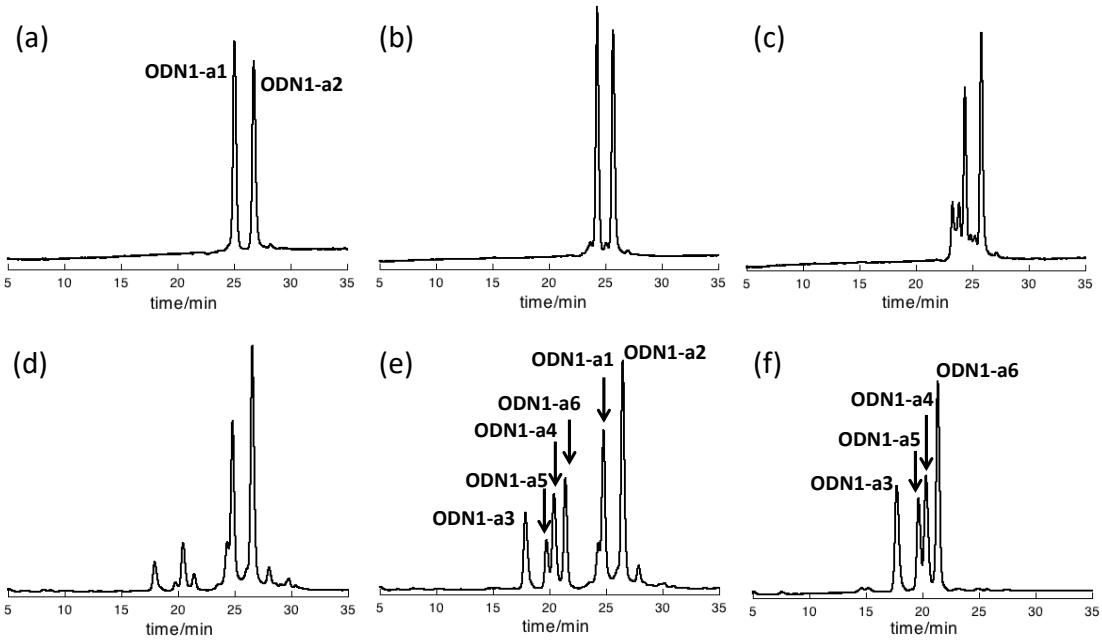


Fig. S6 Reaction of dCGGG/dCGGG with NH_2OH in the absence/presence of NCD. (a) HPLC profile of **ODN1-a1/ODN1-a2** before reaction. (b, c) **ODN1-a1/ODN1-a2** treated with NH_2OH in the absence (b) and presence of NCD (c) for 2 d at 0 °C. (d) ODN mixture obtained after treatment of (c) with piperidine at 90 °C for 30 min followed by alkaline phosphatase (AP). (e) Co-injection of standard ODNs with (d). f) Standard **ODN1-a3–a6.**

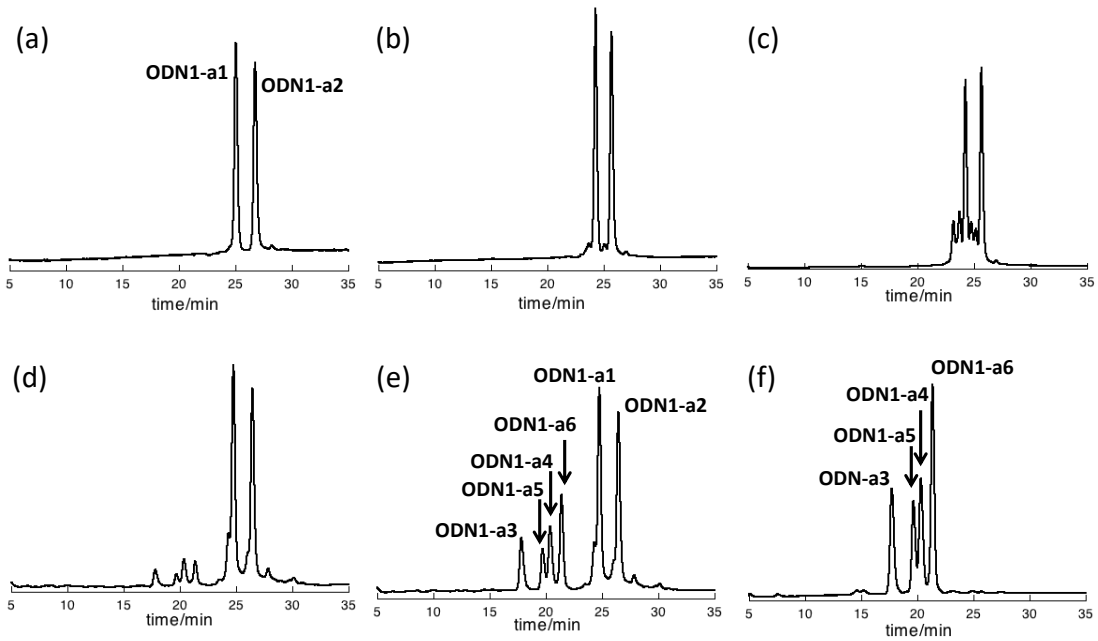


Fig. S7 Reaction of dCGGG/dCGGG with NH_2OH in the absence/presence of *p*-NCTB. (a) HPLC profile of **ODN1-a1/ODN1-a2** before reaction. (b, c) **ODN1-a1/ODN1-a2** treated with NH_2OH in the absence (b) and presence of *p*-NCTB (c) for 2 d at 0 °C. (d) ODN mixture obtained after treatment of (c) with piperidine at 90 °C for 30 min followed by alkaline phosphatase (AP). (e) Co-injection of standard ODNs with mixture of (d). (f) Standard **ODN1-a3–a6.**

10. Estimation of binding stoichiometry

(10-1) CSI-TOF MS measurements

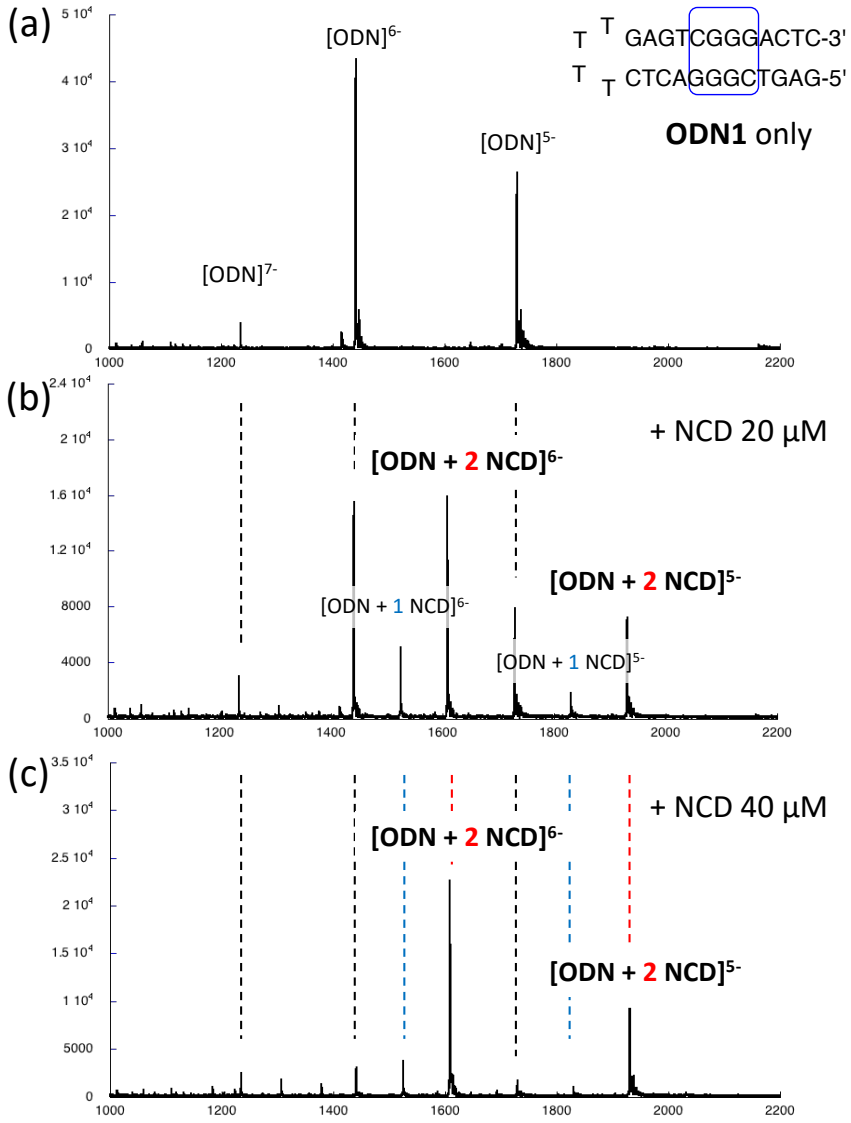


Fig. S8 CSI-MS spectra of **ODN1** in 50% aqueous methanol and 100 mM ammonium acetate in the absence (a) and presence of NCD (b: 20 µM NCD; c: 40 µM NCD). The sample solution was cooled to –10 °C during injection.

(10-2) Isolation and analysis of the complex by HPLC

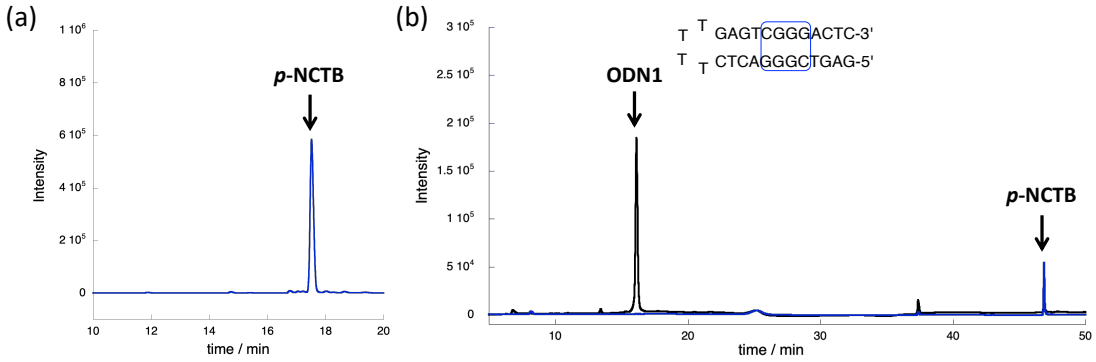


Fig. S9 (a) HPLC chart of *p*-NCTB detected at 320 nm. (b) HPLC chart for estimation of molar ratio of *p*-NCTB/**ODN1** in the binding complex. **ODN1** (10 μM) and *p*-NCTB (40 μM) were mixed to produce the complex. The desired binding complex was isolated from the native PAGE, which was subjected to HPLC analysis. Absorptions at 254 nm (black) and 320 nm (blue) were used for detection of ODN and *p*-NCTB, respectively.

Table S2 Stoichiometry estimated by HPLC and CSI-TOF MS

	HPLC method	CSI-TOF MS method
ODN1 + *p*-NCTB	0.92	-[a]
ODN1 + NCD	2.09	2

[a] Ions attributed to the complexes were not observed under the same conditions.

11. Sequence requirements for interstrand complex formation between *p*-NCTB and dCGGG/dCGGG

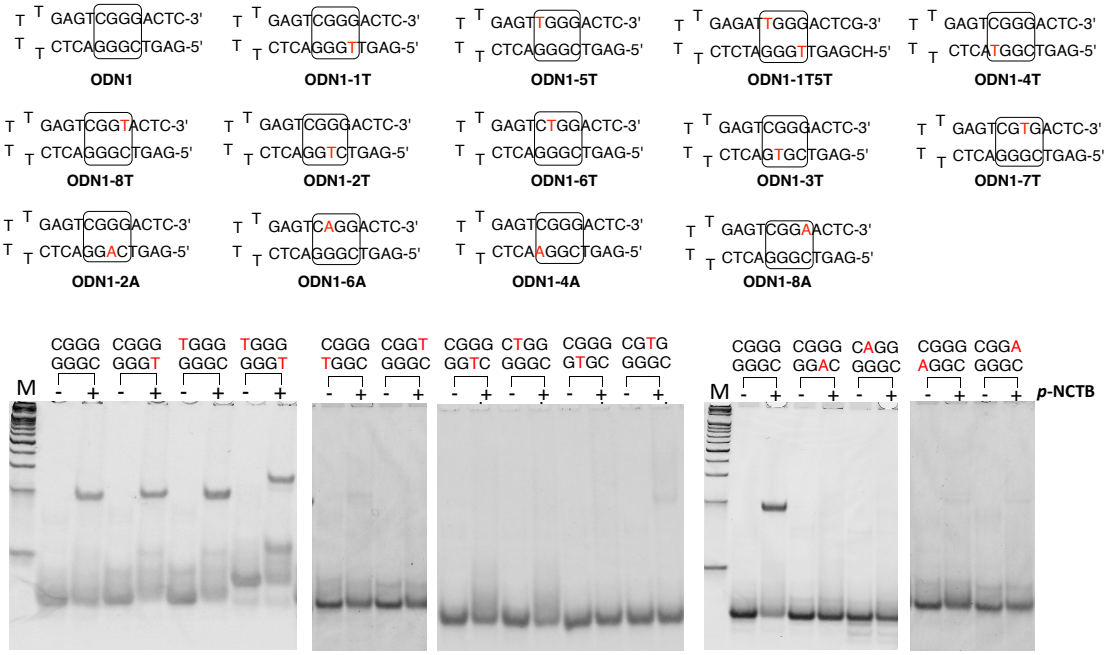


Fig. S10 Native PAGE analysis to monitor interstrand binding (I) of *p*-NCTB to a series of hairpin ODNs to explore the necessity of each base in dCGGG/dCGGG for the *p*-NCTB binding. The expected binding sites are marked with rectangular box, and changed bases from **ODN1** are red-colored. ODN: 1 μM; *p*-NCTB: 4 μM; lane M: DNA marker (20 bp ladder).

12. Sequence requirements for intrastrand complex formation between *p*-NCTB and dCGGG/dCGGG

(12-1) Single base alterations in a dCGGG/dCGGG site in the tandem hairpin ODN

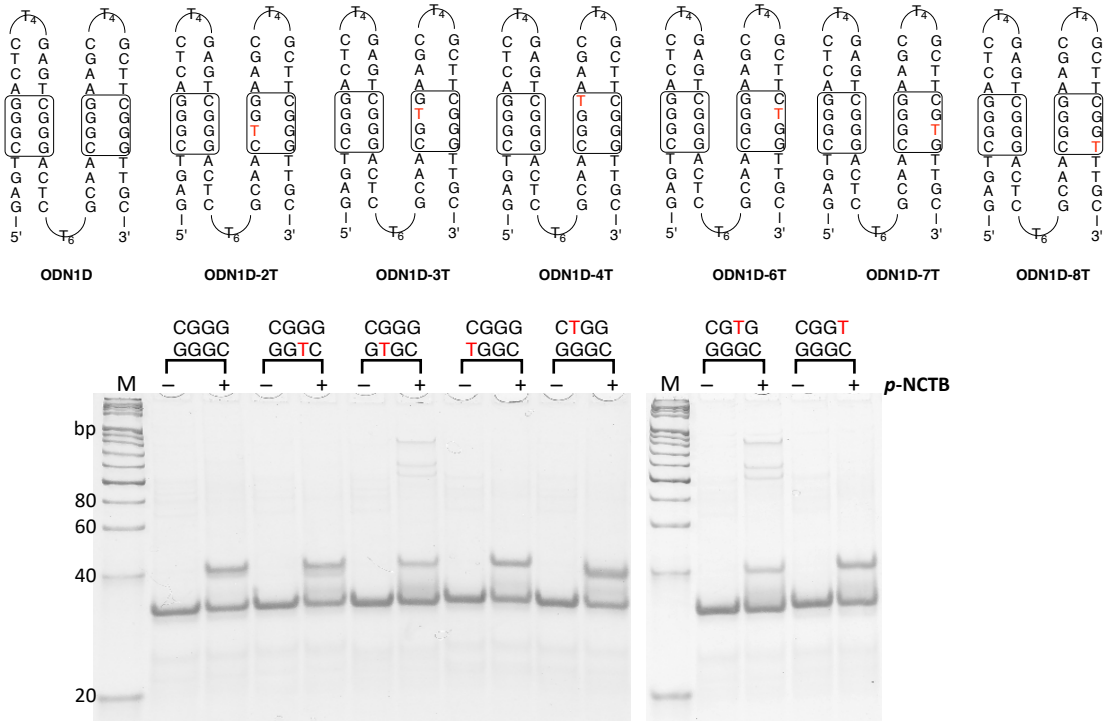


Fig. S11 Native PAGE analysis to monitor intrastrand binding (II) of *p*-NCTB to a series of tandem hairpin ODNs to explore the necessity of each base in dCGGG/dCGGG for the *p*-NCTB binding. The expected binding sites are marked with rectangular box, and the base alterations are red-colored. ODN: 1 μM; *p*-NCTB: 2 μM; lane M: DNA marker (20 bp ladder).

(12-2) Two base alterations in each dCGGG/dCGGG site in the tandem hairpin ODN

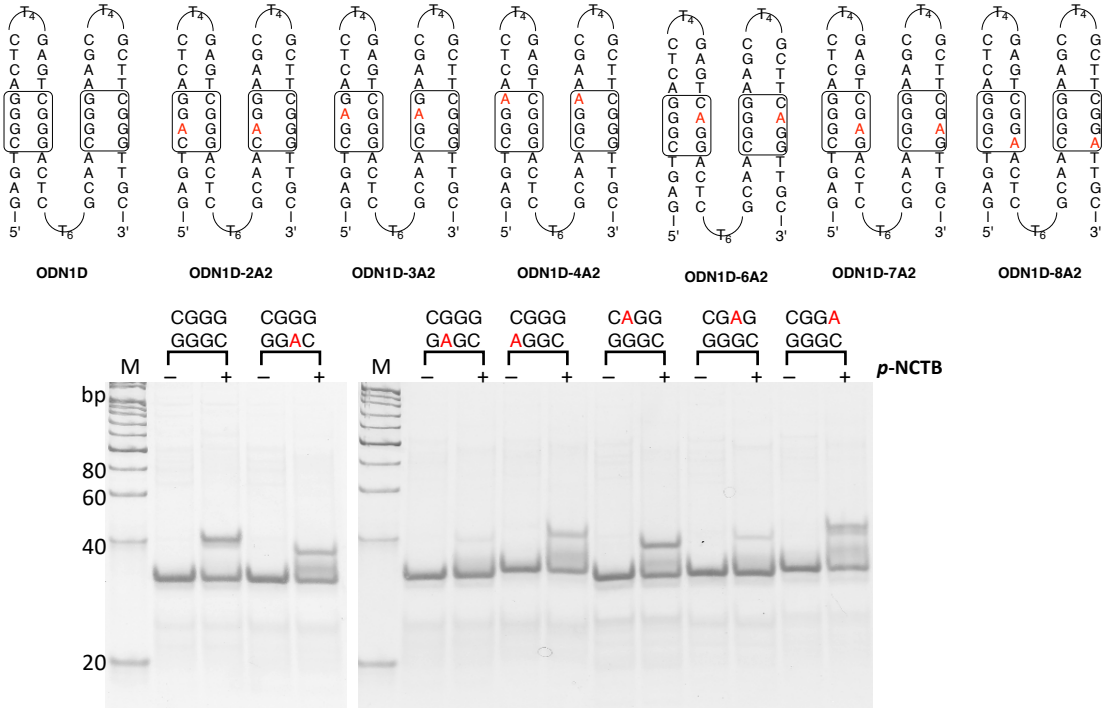


Fig. S12 Native PAGE analysis to monitor intrastrand binding (II) of *p*-NCTB to a series of tandem hairpin ODNs to explore the necessity of each base in dCGGG/dCGGG. Two dCGGG/dCGGG sites in the tandem hairpins carry the same nucleotide change from ODN1D. The expected binding sites are marked with rectangular box, and the base alterations are red-colored. ODN: 1 μM; *p*-NCTB: 2 μM; lane M: DNA marker (20 bp ladder).

(12-3) dCGG/dGGG: change of a terminal G–C base pair in each dCGGG/dCGGG site in the tandem hairpin ODN

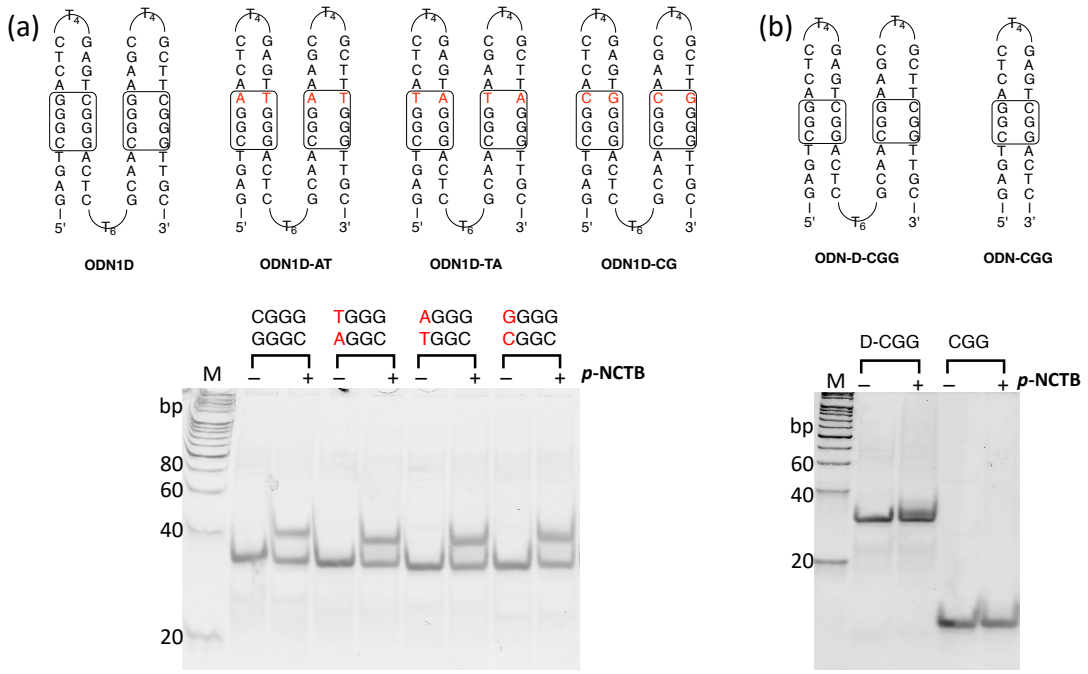


Fig. S13 (a)Native PAGE analysis to monitor intrastrand binding (II) of *p*-NCTB to a series of tandem hairpin ODNs to explore the necessity of each base in dCGGG/dCGGG. One of a terminal G–C pair in a dCGGG/dCGGG site was changed in the designed sequences. The expected binding sites are marked with rectangular box, and the base alterations are red-colored. (b) Native PAGE analysis of *p*-NCTB binding to hairpin ODNs bearing single and two dCGG/dCGG sites. ODN: 1 μM; *p*-NCTB: 2 μM; lane M: DNA marker (20 bp ladder).

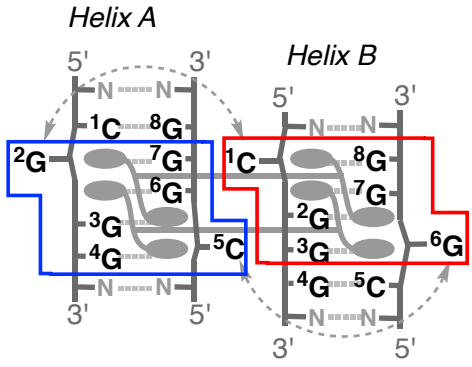


Fig. S14 One of the possible binding complexes between *p*-NCTB and two dCGGG/dCGGG sites (^{2}G^{3}G^{4}G/^{5}C^{6}G^{7}G: blue box; ^{1}C^{2}G^{3}G/^{6}G^{7}G^{8}G: red box)

The binding analysis using a series of tandem hairpin ODNs in Figure S10–12 demonstrated that intrastrand binding (II) of *p*-NCTB had a weaker sequence requirement than interstrand binding (I) (Fig. S9). Base alterations at G3 and G7 had a relatively large effect on reducing the binding of *p*-NCTB (see significant effects on ODN1D-3A2 and ODN1D-7A2 in Figure S11, and ODN1D-3T and ODN1D-7T in Figure S10). The importance of G3 and G7 could be rationalized by considering that G3 and G7 are always involved in the recognition by naphthyridine moieties in all possible binding complex i-iv in Figure S15 (see also Fig. S14). Changes of terminal G–C base pair did not abolish intrastrand binding (Fig. S13), which supported that a dCGG/dGGG was binding site of *p*-NCTB in dCGGG/dCGGG.

13. Possible binding complexes between *p*-NCTB and dCGGG/dCGGG sequence

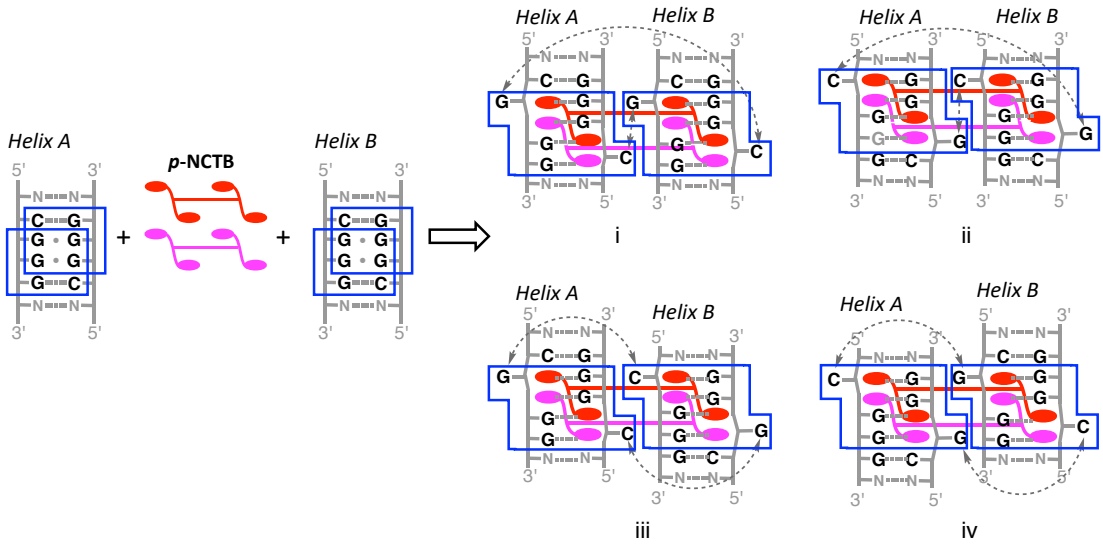


Fig. S15 Possible binding complexes between *p*-NCTB and dCGGG/dCGGG sequence. Two *p*-NCTB molecules are shown in red and magenta. The dCGGG/dCGGG contains two overlapped dCGG/dGGG sites as depicted in blue boxes, each of which can be recognized by NCD moieties. Depending on a combination of the binding sites used, there are four kinds of interhelical complexes (i-iv). Orientation of each helix in the complex are not considered in these complexes. Interhelical G–C interactions shown by dotted arrows might stabilize the complexes.

14. Molecular modeling of the complex between *p*-NCTB and dCGGG/dCGGG

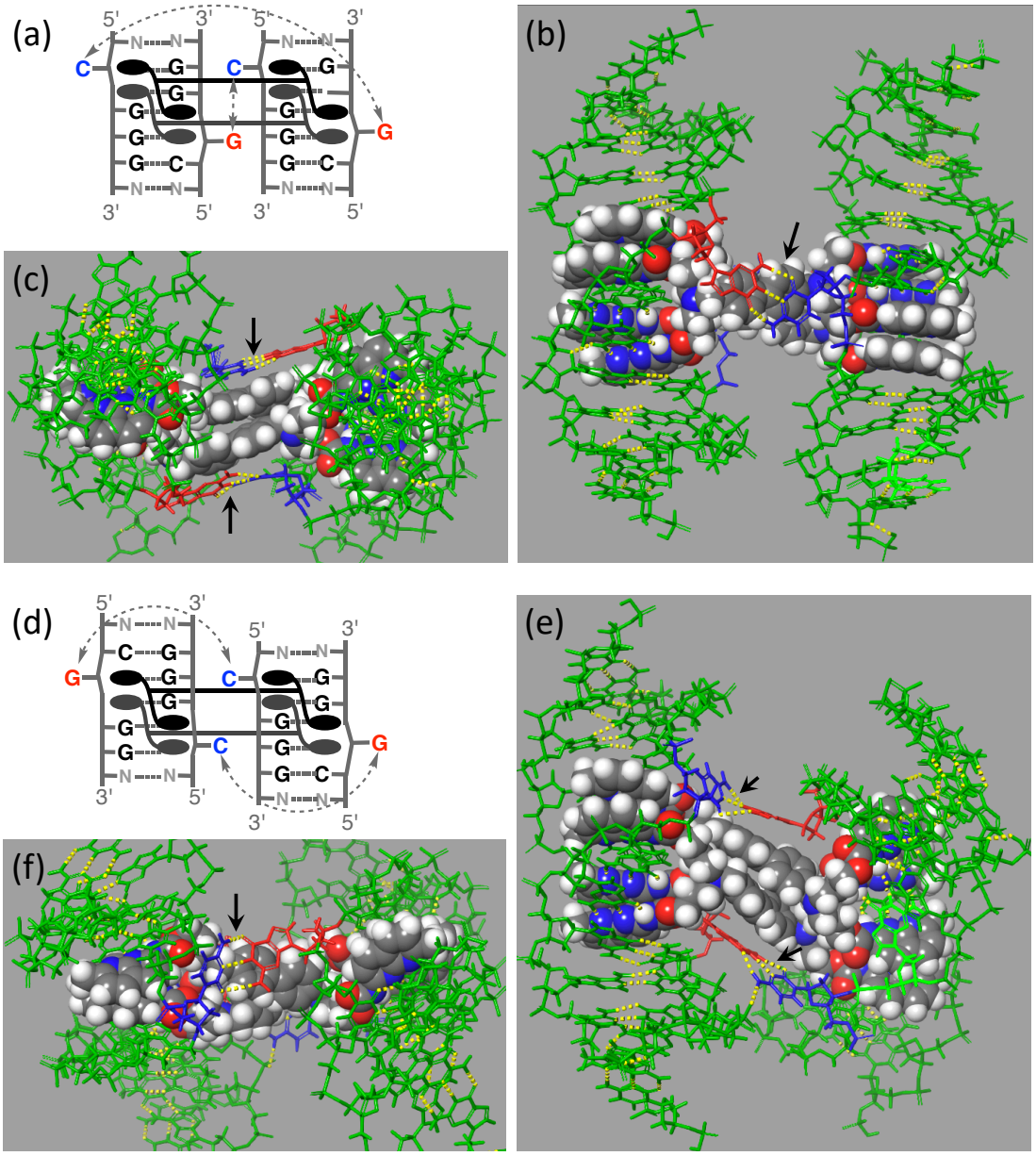


Fig. S16 Molecular modeling simulation of *p*-NCTB–dCGGG/dCGGG complex. Two possible binding patterns were shown (a and d). (a–c) *p*-NCTB–dCGGG/dCGGG complex corresponding to i or ii in Figure S15. (d–e) *p*-NCTB–dCGGG/dCGGG complex corresponding to iii or iv in Figure S15. (b, e) Structures from front view, and (c, f) structures from top view. Self-complementary d(5'-GAGT**CGGG**ACTC-3') oligomer was used for the simulation, and *p*-NCTB is represented in sphere shape. In the binding models, four guanine bases in a dCGGG/dCGGG are recognized by four 2-acylamino-1,8-naphthyridine moieties, and the other two guanines are either paired with cytosine or flipped out of helices. Cytosines in a dCGGG/dCGGG are also either paired with guanine or flipped out of helices. The flipping-out guanines (red) and cytosines (blue) from different helices may form hydrogen bondings (indicated by arrow, and three hydrogen bondings were formed from each flipped-out G–C base pair). The structure was optimized with AMBER* force

field in water using Maestro 11.5.011 (Schrödinger software). Energy minimization was performed by the Polak–Ribier Conjugate Gradient (PRCG) method with a convergence threshold of 0.05. Initial structures were constructed manually referring to the NMR–structure that we determined previously.