Electronic supplementary information

Mechanistic Understanding of Multistep Assembly of DNA with Carbazole Ligand by Simple Adjustment of Host-Guest Concentrations

Norie Inukai, Tsuyoshi Kawai*, and Junpei Yuasa*

Graduate School of Materials Science, Nara Institute of Science and Technology, 8916-5 Takayamacho, Ikoma, Nara 630-0192, Japan,

* To whom correspondence should be addressed.

E-mail: yuasaj@ms.naist.jp; tkawai@ms.naist.jp

Experimental Section

Calf thymus DNA (approximate average size: >13 kb) was obtained from Aldrich. The synthesis of 9-ethyl-3,6-bis[(1-methyl-1H-imidazol-2-yl)ethynyl]-9-carbazole [Im_2Cz] was reported previously.[1] 2,2'-(9-ethyl-9H-carbazole-3,6-diyl)bis(ethyne-2,1-diyl)bis(1,3-dimethyl-1H-imidazol-3-ium) [$(Im^+)_2Cz$] was synthesized by methylation of Im_2Cz with methyl iodide in acetonitrile. A 15 ml flask was charged with Im_2Cz (515 mg, 0.75 mmol) and methyl iodide (500 μl, 8.0 mmol) in acetonitrile. The mixture was stirred for 24 h and then filtered. This methylation reaction occurred stoichiometrically and the crude product was recrystallized in methanol to provide yellow solid. ^{1}H NMR (300 MHz, CD_3OD): δ 8.73 (s, 2H), 7.81 (d, J = 8.7 Hz, 2H), 7.67 (d, J = 8.7 Hz, 2H), 7.58 (s, 4H), 4.47 (q, J = 7.2 Hz, 2H), 3.96 (s, 12H), 1.38 (t, J = 7.2 Hz, 3H). HRMS (ESI): m/z calcd. for $C_{28}H_{27}IN_5$ ([M – I]$^+$), 560.1311; found 560.1311. Emission quantum yields of $(Im^+)_2Cz$ (2.5×10^{-6} M) in the presence of ct-DNA (2.8×10^{-6} M and 9.6×10^{-5} M) were measured using a calibrated integrating sphere system in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA.

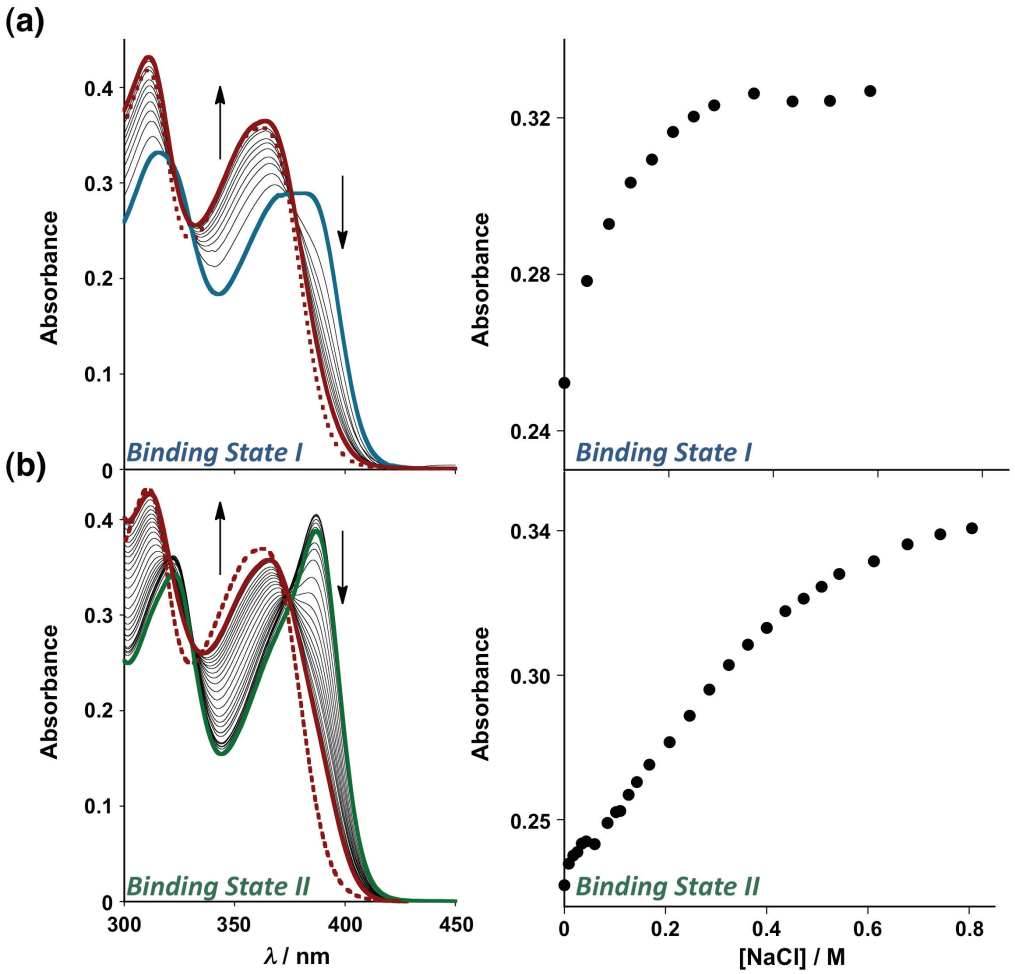


Fig. S1 UV-vis absorption spectra of (a) $(Im^+)_2Cz$ (9.8×10^{-6} M) in the presence of ct-DNA (base pair concentration: 9.8×10^{-6} M) and NaCl [0 M (blue line) – 0.59 M (red solid line)] (b) $(Im^+)_2Cz$ (1.0×10^{-5} M) in the presence of ct-DNA (base pair concentration: 1.0×10^{-4} M) and NaCl [0 M (green line) – 0.81M (red solid line)] in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA at 298 K. The red dashed lines show UV-vis absorption spectra of free $(Im^+)_2Cz$. The right Figures show plots of absorbance (A) at $\lambda = 360$ nm versus [NaCl].

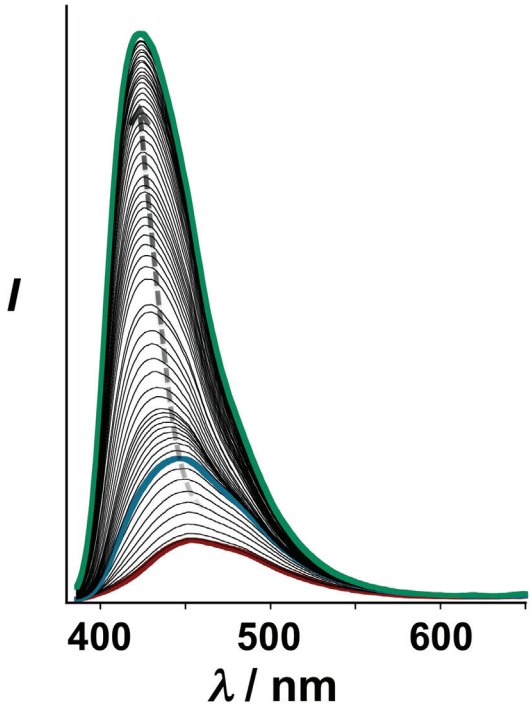
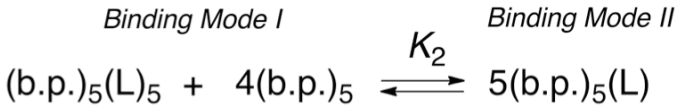


Fig. S2 Fluorescence spectra of $(Im^+)_2Cz$ ($[L]_0 = 1.2 \times 10^{-5}$ M) in the presence of ct-DNA (base pair concentration, [b.p.] = 0 M: red line – 1.5×10^{-6} M: blue line – 2.1×10^{-4} M: green line) in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA at 298 K. Excitation at $\lambda = 364$ nm.



The equilibrium constant (K_2) is expressed by eqn (1), where $[(b.p.)_5]_f$ denotes the concentration of unbound sequences of 5 base pairs in ct-DNA.

$$K_2 = [(b.p.)_5(L)]^5/([(b.p.)_5(L)_5][(b.p.)_5]_f^4) \quad (1)$$

Under the present conditions, almost all carbazole ligands (L) are converted to $(b.p.)_5(L)_5$ [binding state I] or $(b.p.)_5(L)$ [binding state II] in the presence of high concentrations of ct-DNA. In such a case, the initial concentration of carbazole ligands ($[L]_0$) can be explained by eqn (2).

$$[L]_0 = 5[(b.p.)_5(L)_5] + [(b.p.)_5(L)] \quad (2)$$

The conversion ratio of $(b.p.)_5(L)_5$ to $(b.p.)_5(L)$ [$\alpha = (A - A_0)(A_\infty - A_0)^{-1}$ or $(\Delta\varepsilon - \Delta\varepsilon_0)(\Delta\varepsilon_\infty - \Delta\varepsilon_0)^{-1}$] is expressed by eqn (3), where A_0 and A_∞ are absorbance at $\lambda = 370$ nm, $\Delta\varepsilon_0$ and $\Delta\varepsilon_\infty$ are the molar circular dichroism at $\lambda = 330$ nm of $(b.p.)_5(L)_5$ and $(b.p.)_5(L)$, respectively.

$$\alpha = [(b.p.)_5(L)]/[L]_0 \quad (3)$$

Then, eqn 4 is derived from eqn (2) and (3).

$$[(b.p.)_5(L)_5] = [L]_0(1 - \alpha)/5 \quad (4)$$

On the other hand, the total concentration of sequences of 5 base pairs in ct-DNA ($[(b.p.)_5]_T$) can be given by eqn (5).

$$[(b.p.)_5]_T = [(b.p.)_5(L)_5] + [(b.p.)_5(L)] + [(b.p.)_5]_f \quad (5)$$

From eqn (3)-(5), $[(b.p.)_5]_f$ can be explained by eqn (6).

$$[(b.p.)_5]_f = [(b.p.)_5]_T - [L]_0(1 + 4\alpha)/5 \quad (6)$$

Then, eqn (7) is derived from eqn (1) and (6).

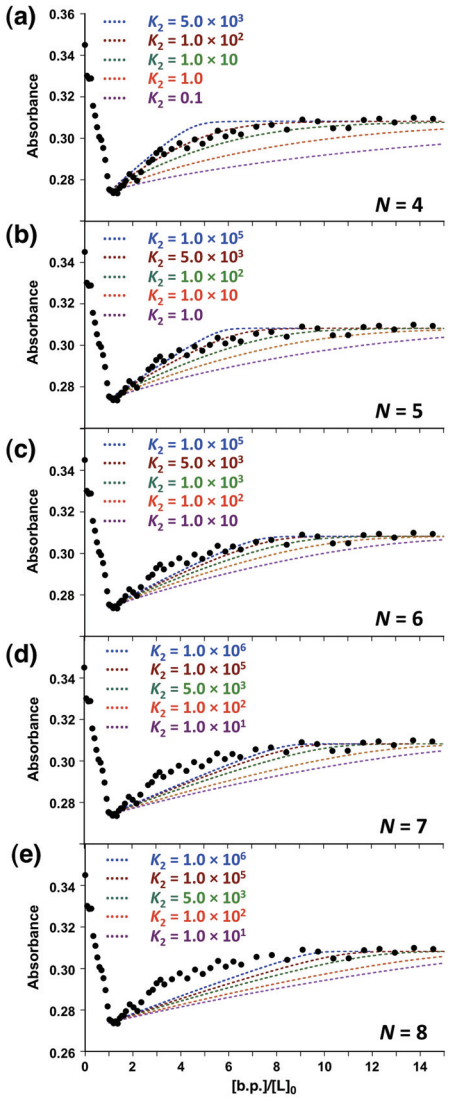
$$K_2\{[(b.p.)_5]_T - [L]_0(1 + 4\alpha)/5\}^4 = 5[L]_0^4\alpha^5(1 - \alpha)^{-1} \quad (7)$$

Since $[(b.p.)_5]_T = [b.p.]/5$, eqn (7) can be rearranged to eqn (8), where [b.p.] denotes the total concentration of base pairs introduced into the solution.

$$K_2\{[b.p.] - [L]_0(1+ 4\alpha)\}^4 = (5\alpha)^5(1 - \alpha)^{-1}[L]_0^4 \quad (8)$$

Fitting equation for $N = 4$, 6, 7, and 8 can be derived by the same procedure.

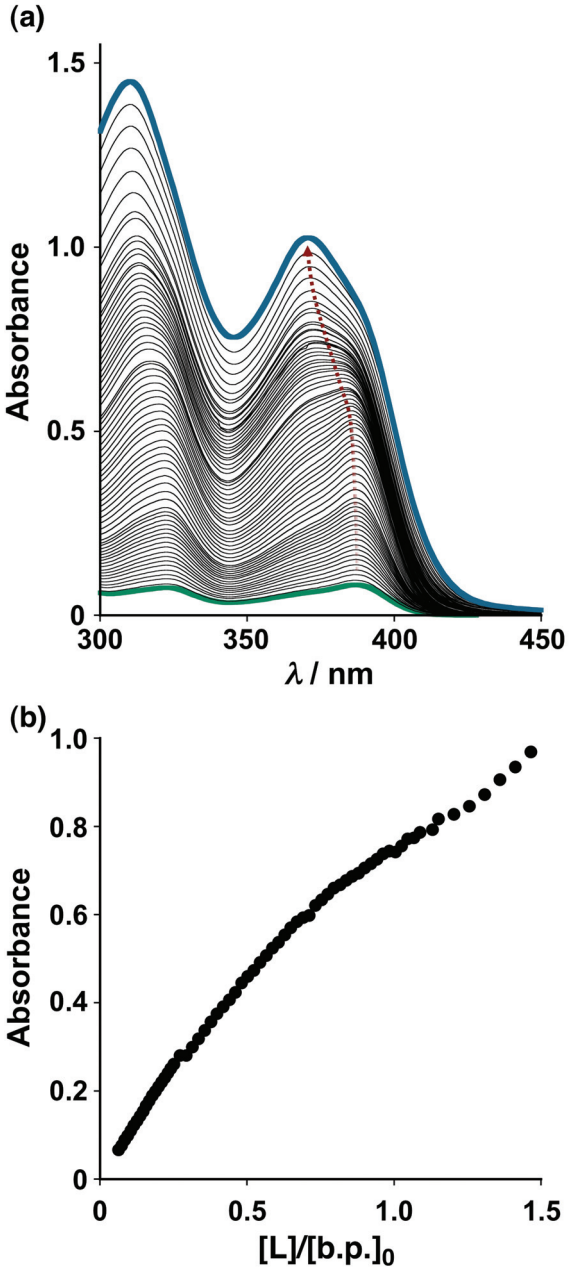
$$K_2\{[b.p.] - [L]_0(1+ (N - 1)\alpha)\}^{(N-1)} = (N\alpha)^N(1 - \alpha)^{-1}[L]_0^{(N-1)} \quad (9)$$



[b.p.] / M	Abs	[b.p.] / M	Abs
0	0.34505	6.1×10^{-5}	0.30092
8.9×10^{-7}	0.33017	6.4×10^{-5}	0.30336
1.8×10^{-6}	0.32875	6.7×10^{-5}	0.30187
2.7×10^{-6}	0.32886	7.4×10^{-5}	0.30561
3.6×10^{-6}	0.31566	8.0×10^{-5}	0.30647
4.4×10^{-6}	0.31096	8.7×10^{-5}	0.30419
5.3×10^{-6}	0.30534	9.3×10^{-5}	0.30903
6.2×10^{-6}	0.30067	1.0×10^{-4}	0.30816
7.1×10^{-6}	0.29922	1.1×10^{-4}	0.30483
8.0×10^{-6}	0.29536	1.1×10^{-4}	0.30493
8.9×10^{-6}	0.28958	1.2×10^{-4}	0.30885
1.0×10^{-5}	0.28177	1.3×10^{-4}	0.30943
1.1×10^{-5}	0.27534	1.3×10^{-4}	0.30764
1.2×10^{-5}	0.27462	1.4×10^{-4}	0.30994
1.2×10^{-5}	0.27356	1.5×10^{-4}	0.3094
1.3×10^{-5}	0.27451		
1.4×10^{-5}	0.27351		
1.5×10^{-5}	0.27601		
1.6×10^{-5}	0.27715		
1.7×10^{-5}	0.27728		
1.8×10^{-5}	0.27939		
1.9×10^{-5}	0.28274		
2.1×10^{-5}	0.28125		
2.3×10^{-5}	0.27955		
2.4×10^{-5}	0.28373		
2.8×10^{-5}	0.28838		
2.9×10^{-5}	0.28972		
3.1×10^{-5}	0.29288		
3.3×10^{-5}	0.29458		
3.4×10^{-5}	0.29233		
3.8×10^{-5}	0.2948		
4.1×10^{-5}	0.29768		
4.4×10^{-5}	0.29522		
4.8×10^{-5}	0.29946		
5.1×10^{-5}	0.29737		
5.4×10^{-5}	0.3002		
5.7×10^{-5}	0.30363		

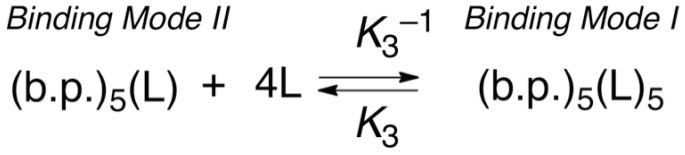
Fig. S4 Fitted lines calculated from eqn (1) using values for (a) $N = 4$, (b) $N = 5$, (c) $N = 6$, (d) $N = 7$, and (e) $N = 8$ with various K_2 values are drawn in plot of absorbance at $\lambda = 370$ nm *vs*. [b.p.]/[L]$_0$ for titration of $(Im^+)_2Cz$ ([L]$_0$ = 1.0×10^{-5} M) by ct-DNA in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA at 298 K.

* Calculated curves obtained from eqn (1) using values for $N = 4$ with $K_2 = 1.0 \times 10^2$ [red line in (a)] and for $N = 5$ with $K_2 = 5.0 \times 10^3$ [red line in (b)] fit to the experimental data. According to the self-complementary oligomer ([d-(GCGC)]$_2$) result, DNA sequence with length shorter than 5 base pairs is not support binding state II (see ESI† S8). Thus, we postulated that the $(Im^+)_2Cz$ ligand is bound per 5 DNA base pairs ($N = 5$) in binding state II.



[L] / M	ε / M⁻¹ cm⁻¹	[L] / M	ε / M⁻¹ cm⁻¹
1.75×10^{-6}	38076	2.28×10^{-5}	30104
2.05×10^{-6}	37478	2.34×10^{-5}	29687
2.34×10^{-6}	38241	2.39×10^{-5}	29494
2.63×10^{-6}	37695	2.45×10^{-5}	29217
2.92×10^{-6}	37920	2.50×10^{-5}	28974
3.21×10^{-6}	38185	2.56×10^{-5}	28836
3.50×10^{-6}	37877	2.62×10^{-5}	28465
3.79×10^{-6}	37894	2.67×10^{-5}	27789
4.08×10^{-6}	37789	2.73×10^{-5}	27732
4.37×10^{-6}	38130	2.78×10^{-5}	27769
4.66×10^{-6}	38121	2.84×10^{-5}	27309
4.95×10^{-6}	38340	2.89×10^{-5}	27223
5.24×10^{-6}	38105	2.95×10^{-5}	26919
5.53×10^{-6}	37922	3.05×10^{-5}	26754
5.82×10^{-6}	37839	3.11×10^{-5}	26626
6.11×10^{-6}	37688	3.25×10^{-5}	26068
6.40×10^{-6}	37668	3.38×10^{-5}	25809
6.69×10^{-6}	37506	3.52×10^{-5}	25771
6.98×10^{-6}	37420	3.65×10^{-5}	25605
7.55×10^{-6}	37207	3.79×10^{-5}	25607
8.13×10^{-6}	36876	3.92×10^{-5}	25663
8.70×10^{-6}	36618		
9.28×10^{-6}	36396		
9.85×10^{-6}	36234		
1.04×10^{-5}	36020		
1.10×10^{-5}	35602		
1.16×10^{-5}	35220		
1.21×10^{-5}	34934		
1.27×10^{-5}	35057		
1.33×10^{-5}	34674		
1.38×10^{-5}	34224		
1.44×10^{-5}	34136		
1.50×10^{-5}	33901		
1.55×10^{-5}	33732		
1.61×10^{-5}	33364		
1.67×10^{-5}	33234		
1.72×10^{-5}	33110		
1.78×10^{-5}	32822		
1.84×10^{-5}	32322		
1.89×10^{-5}	31615		
1.95×10^{-5}	31893		
2.00×10^{-5}	31629		
2.06×10^{-5}	31396		
2.12×10^{-5}	31212		
2.17×10^{-5}	30740		
2.23×10^{-5}	30429		

Fig. S5 (a) UV-visible absorption spectra of $(Im^+)_2Cz$ ([L] = 1.8×10^{-6} M: green line – 3.9×10^{-5} M: blue line) in the presence of ct-DNA (base pair concentration, $[b.p.]_0 = 2.8 \times 10^{-5}$ M) in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA at 298 K. (b) Plot of absorbance (A) at λ = 375 nm versus the ratio of $(Im^+)_2Cz$ and base pair ($[L]/[b.p.]_0$) for the titration of ct-DNA by $(Im^+)_2Cz$.



The equilibrium constant (K_3) is expressed by eqn (1), where $[L]_f$ denotes the concentration of unbound (free) carbazole ligands.

$$K_3 = [(b.p.)_5(L)][L]_f^4/[(b.p.)_5(L)_5] \quad (1)$$

Under the present conditions, almost all base pairs are converted to $(b.p.)_5(L)_5$ [binding state I] or $(b.p.)_5(L)$ [binding state II] in the presence of high concentrations of the carbazole ligand. In such a case, the initial concentration of base pair ($[b.p.]_0$) can be explained by eqn (2).

$$[b.p.]_0 = 5[(b.p.)_5(L)_5] + 5[(b.p.)_5(L)] \quad (2)$$

The conversion ratio (β) of $(b.p.)_5(L)$ to $(b.p.)_5(L)_5$ is expressed by eqn (3).

$$\beta = 5[(b.p.)_5(L)_5]/[b.p.]_0 \quad (3)$$

Then, eqn (4) is derived from eqn (2) and (3), and eqn (4) is rearranged to eqn (5).

$$[b.p.]_0 = \beta[b.p.]_0 + K_3\beta[L]_f^{-4}[b.p.]_0 \quad (4)$$

$$K_3\beta(1 - \beta)^{-1} = [L]_f^4 \quad (5)$$

On the other hand, eqn (6) represents the total concentration of the carbazole ligand ([L]) introduced into the solution.

$$[L] = 5[(b.p.)_5(L)_5] + [(b.p.)_5(L)] + [L]_f \quad (6)$$

Then, eqn (7) is derived from eqn (1), (3), (5), and (6).

$$[L] = \beta[b.p.]_0 + (1 - \beta)[b.p.]_0/5 + \{K_3\beta(1 - \beta)^{-1}\}^{1/4} \quad (7)$$

Thus, eqn (7) can be rearranged to eqn (8).

$$\{[L] - [b.p.]_0(1 + 4\beta)/5\}^4 = K_3\beta(1 - \beta)^{-1} \quad (8)$$

$$5L + (b.p.)_5 \overset{K_1}{\rightleftharpoons} (b.p.)_5(L)_5$$

$$(b.p.)_5(L)_5 + 4(b.p.)_5 \overset{K_2}{\rightleftharpoons} 5(b.p.)_5(L)$$

$$(b.p.)_5(L)_5 \overset{K_3}{\rightleftharpoons} (b.p.)_5(L) + 4L$$

The equilibrium constants K_1, K_2, and K_3 are expressed by eqn (1), (2), and (3), respectively. In these equations, $[L]_f$: the concentration of unbound (free) carbazole ligands; $[(b.p.)_5]_f$: the concentration of unbound sequences of 5 base pairs in ct-DNA.

$$K_1 = [(b.p.)_5(L)_5]/([L]_f^5[(b.p.)_5]_f) \quad (1)$$

$$K_2 = [(b.p.)_5(L)]^5/([(b.p.)_5(L)_5][(b.p.)_5]_f^4) \quad (2)$$

$$K_3 = [(b.p.)_5(L)][L]_f^4/[(b.p.)_5(L)_5] \quad (3)$$

Then, K_1K_2 and K_1K_3 values are derived from eqn (1)-(3).

$$K_1K_2 = \{[(b.p.)_5(L)]/([L]_f[(b.p.)_5]_f)\}^5 \quad (4)$$

$$K_1K_3 = [(b.p.)_5(L)]/([L]_f[(b.p.)_5]_f) \quad (5)$$

Thus, eqn (6) can be derived from eqn (4)-(5).

$$K_1K_2 = (K_1K_3)^5 \quad (6)$$

Thus, eqn (7) can be obtained from eqn (6).

$$K_2 = K_1^4K_3^5 \quad (7)$$

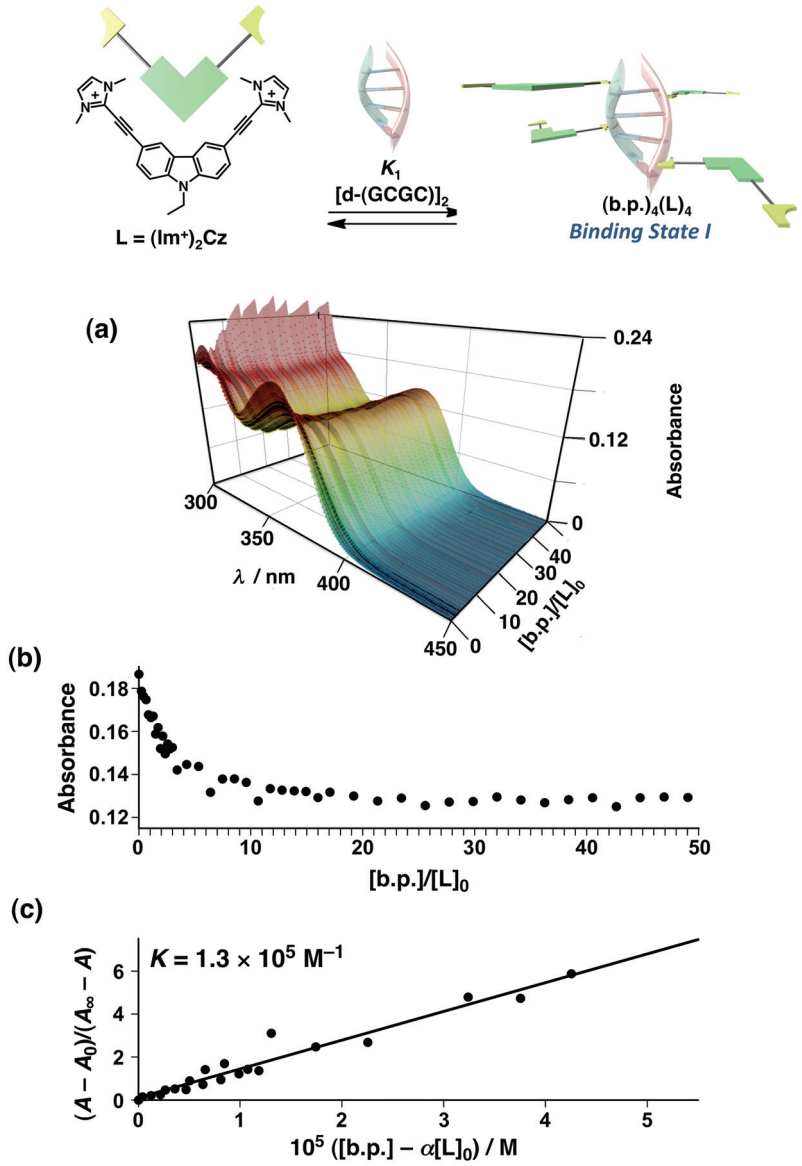


Fig. S8 (a) UV-vis absorption spectra of $(Im^+)_2Cz$ ($[L]_0 = 4.9 \times 10^{-6}$ M) in the presence of $[d\text{-}(GCGC)]_2$ (base pair concentration, [b.p.] = 0 M – 2.5×10^{-4} M) in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA at 298 K. Plots of (b) absorbance (A) at $\lambda = 360$ nm versus $[b.p.]/[L]_0$ and (c) $(A - A_0)(A_\infty - A)^{-1}$ versus $([b.p.] - \alpha[L]_0)$, where $\alpha = (A - A_0)(A_\infty - A_0)^{-1}$ at $\lambda = 360$ nm. Equation for the simple 1:1 binding equilibrium: the absorbance change due to binding between $(Im^+)_2Cz$ and $[d\text{-}(GCGC)]_2$ can be expressed by $(A - A_0)/(A_\infty - A) = K([b.p.] - \alpha[L]_0)$.

Reference

(1) N. Inukai, J. Yuasa and T. Kawai, *Chem. Commun*. 2010, **46**, 3929.