Article

Prototropically Allosteric Probe for Superbly Selective DNA Analysis

Fan Lin, Yufeng Zhou, Qiusha Li, Xiao-Shun Zhou, Yong Shao, Benoit Habermeyer, Hui Wang, Xinghua Shi, and Zhiai Xu

Anal. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.analchem.7b02077 • Publication Date (Web): 25 Jul 2017

Prototropically Allosteric Probe for Superbly Selective DNA Analysis

Fan Lin,[†] Yufeng Zhou,[†] Qiusha Li,[†] Xiaoshun Zhou,[†] Yong Shao,*[,†] Benoit Habermeyer,[‖] Hui Wang,[¶] Xinghua Shi,[¶] and Zhiai Xu*[,‡]

[†]Institute of Physical Chemistry, College of Chemistry and Life Sciences, Zhejiang Normal University, Jinhua 321004, Zhejiang, China

[‖]PorphyChem SAS, Dijon 21000, France

[¶]Chinese Academy of Sciences (CAS) Key Laboratory of Nanosystem and Hierarchy Fabrication, CAS Center for Excellence in Nanoscience, National Center for Nanoscience and Technology, Beijing 100190, China

[‡]School of Chemistry and Molecular Engineering, East China Normal University, Shanghai 200241, China

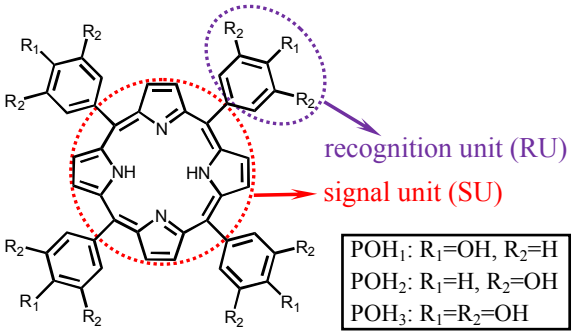
*E-mail: yshao@zjnu.cn (S. Y.). Fax: 86 579 82282595; zaxu@chem.ecnu.edu.cn (Z. X.)

ABSTRACT: Selective nucleotide recognition for biosensor evolution requires rational probe design towards the binding pattern-susceptible readout but without serious poison in selectivity from the context sequences. In this work, we synthesized a dual-function trihydroxyphenyl porphyrin (POH_3) to target abasic site (AP site) in ds-DNA using the trihydroxyphenyl substituent and the tetrapyrrole macrocycle as the recognition unit (RU) and the fluorescent signal unit (SU), respectively. RU and SU are separated each other but are prototropically allosteric. We found that an appropriate pH favors formation of the nonfluorescent quinine/pyrrole (O-NH) conformer of POH_3. However, the complementary hydrogen bonding of RU in O-NH with the target cytosine opposite the AP site switches on the SU fluorescence through prototropic allostery towards the phenol/isopyrrole (OH-N) conformer, while the bases of thymine, guanine, and adenine totally silence this allostery, suggesting a superb selectivity in single nucleotide polymorphism (SNP) analysis. The role of the prototropic allostery in achieving such SNP selectivity is also convinced using porphyrins with other hydroxyl substituent patterns. Because of the SU separation from RU, SU isn't directly involved in the interaction with the AP site and thus the turn-on selectivity is also realized for DNA with flanking guanine, the most easily oxidized base in DNA. This tolerance to the flanking base identity has seldom been achieved in previous reports. Additionally, other DNA structures can't bring this allostery, indicating that the combination recipe of the AP site design and the prototropically allosteric probe will find wide applications in DNA-based sensors.

Nucleotide recognition in nucleic acid has attracted much attention due to significance of point mutation in disease incidence.[1] Identifying one target from a strand with numerous extra nucleotides using a simple method is still in great demand, although currently developed genome sequencing methods can be manipulated by a skilled biochemist.[2-4] Hybridization-based strategies cooperative with mismatch-sensitive probes have been usually employed as the straightforward resolutions to report the single nucleotide polymorphism (SNP) using fluorescent, colorimetric, and electrochemical techniques as readouts.[4-7] Continuous efforts have been made to boost the SNP discrimination by enzyme manipulation and target-initiated multistage DNA assembly/deassembly.[4,8-11] Alternatively, in probe strand, nucleotide covalently labeled with fluorescently or electrochemically active tags and usage of surrogate in place of nucleotide opposite target have been proved as facile rationales in SNP detection.[12-15] When pairing with the target strands, these incorporators report the target nucleotides via their binding-sensitive properties of, for instance, hydrogen bonding and stacking interactions. This strategy is intriguingly straightforward, but its implementation into practice usually requires case-dependent pre-synthesis of the incorporators and subsequent challenging modification of DNA. Another universal and attractive strategy is introduction of local structures into DNA, such as abasic site (AP site),[16-21] bulge site,[16,22,23] gap site,[24,25] etc, which provide cavities ready for the exogenous probe binding. In these structures, only the target nucleotides are left unpaired to be recognized and the probe manipulation is more flexible since the covalent attachment of incorporators into DNA is unnecessarily required. Upon entering into these sites, the signal probes report the typing of SNP with changes in, for example, dielectric environment-dependent fluorescence, self-folding-sensitive emission, and some photophysical behaviors including excited-state intramolecular proton transfer (ESIPT), photoinduced electron transfer (PET), etc.[16-25]

Although these rationales are susceptible to mutation occurrence, serious challenge continuously arises in ideally exploring the SNP identity because of the following requirements: (1) the probe free in solution is better to be totally nonfluorescent and thus the turn-on fluorescence response upon SNP analysis arises from a background-free level; (2) high selectivity in fluorescence response to the target nucleotide must be met for analysis of nucleotide transversion (C/T↔G/A) and transition (C↔T and A↔G); (3) the turn-on selectivity in SNP analysis isn't seriously influenced by neighboring nucleotides other than the target; (4) the DNA probe strand is free of fluorophore and quencher labeling easy for universal manipulation in variant

applications. Thus, a novel rationale in probe design is needed to meet these requirements. In this work, we attempt to develop a versatile fluorescent probe with a prototropic allostery potency in analyzing the SNP identity based on the AP site design, since following the Teramae and Nishizawa's pioneer work[19] used in DNA analysis, the AP site has also received much attention in novel sensor developments.[26 - 29] Porphyrins possessing hydroxyphenyl substituents (POH_n, Scheme 1) were herein selected as the promising probes since they own large absorption coefficient due to the conjugated macrocycle π system and their structures can be easily tautomerized *via* protonation and deprotonation.[30 - 34] The used porphyrins feature a distal hydroxyphenyl recognition unit and a porphyrin fluorogenic signal unit (RU and SU, Scheme 1). We found that fluorescent analysis of the SNP identify with the ideal performance as mentioned above can be greatly facilitated only when the base opposite the AP site is cytosine that has a complementary hydrogen bonding pattern with the hydroxyphenyl RU in POH_3. Prototropic allostery to the emitting state that is caused by such selective recognition was proposed as a novel rationale for the superb SNP analysis. Our work will expand the sensor applications of porphyrins besides their wide usage in targeting G-quadruplex structures.[35-38]



Scheme 1. The used porphyrin structures at the free base state.

EXPERIMENTAL SECTION

Porphyrin Synthesis and Reagents. 5,10,15,20-tetrakis(3,5-dihydrox-yphenyl)porphyrin (POH_2) and 5,10,15,20-tetrakis(4-hydroxy-phenyl)porphyrin (POH_1) were obtained from J&K Scientific Ltd. (Shanghai, China) and used as received. 5,10,15,20-tetrakis(3,4,5-trihydroxyphenyl)porphyrin (POH_3) was synthesized from 5,10,15,20-tetrakis(3,4,5-trimethoxyphenyl) porphyrin (POM_3). POM_3 was first obtained as previously described.[39] In a three-neck round-bottom flask, 24.3 ml of pyridine and 28.5 ml of hydrochloric acid (37% in water) were heated under nitrogen atmosphere until the temperature reached 170 °C. The POM_3 (500.0 mg, 0.513 mmol) was added and the mixture was kept at 170 °C overnight. The reaction mixture was cooled down and 20 ml of methanol was added. The mixture was poured slowly into 600 ml of stirring dichloromethane. The precipitate was filtered and washed with dichloromethane and heptane. The porphyrin was dried at 120°C under vacuum overnight to yield a purple solid (150.0 mg, 36%). ^{1}H NMR spectrum for POH_3 was recorded on a Bruker DRX-300 AVANCE spectrometer at the 'Plateforme d'Analyse Chimique et de Synthèse Moléculaire de l'Université de Bourgogne (PACSMUB)'. Chemical shifts for ^{1}H NMR spectrum was measured in DMSO-*d6*. Mass spectrum was obtained on a Bruker Daltonics Ultraflex II spectrometer at the PACSMUB in the MALDI/TOF reflectron mode using dithranol as a matrix. ^{1}H NMR (300 MHz, DMSO-*d6*) (ppm): –2.94 (s, 2 H, N*H*); 7.12 (s, 8 H, $H_{o\text{-Ph}}$); 8.56 (s, 4 H, O$H_{p\text{-Ph}}$); 8.93 (s, 8 H, $H_{\beta\text{-pyrr}}$); 9.28 (s, 8 H, O$H_{m\text{-Ph}}$). MS (MALDI/TOF) *m/z* = 807.13 for $[M+H]^+$ (807.75 calculated for $C_{44}H_{31}N_4O_{12}^+$). UV–Vis (ethanol): λ_{max} (nm) ($\varepsilon \times 10^{-3}$ M^{-1} cm^{-1}) = 424 (100.12); 515 (9.82); 555 (4.58); 594 (3.88); 654 (4.97). DNA species (Table S1) were synthesized by TaKaRa Biotechnology Co., Ltd (Dalian, China) and purified by HPLC. The naturally-occurred deoxyribose structure is unstable and the tetrahydrofuran residual was used as the chemically stable AP site model. The ss-DNA was first dissolved in pure water for concentration measurement using the 260-nm extinction coefficients calculated by nearest neighbor analysis. To prepare DNA duplex solution, the equivalent probe and target strands were mixed and annealed in a thermocycler (first at 92 °C, then cooled down to room temperature slowly) in 0.1 M phosphate buffer (PBS) and stored in 4 °C overnight. All other chemicals were analytical-reagent grade (Sigma Chemical Co., St. Louis, USA) and used without further purification. Milli-Q water (18.2 mΩ; Millipore Co., Billerica, USA) was used throughout experiments.

Fluorescence Measurements. A FLSP920 spectrofluorometer (Edinburgh Instruments Ltd., Livingston, UK) was used for fluorescence measurements. The quartz cuvette temperature was kept at 20±1 °C using an equipped temperature-controlled circulator (Julabo Labortechnik GmbH, Seelbach, Germany). The DNA-containing porphyrin solution was allowed to incubate for 20 min before fluorescence measurements. The fluorescence experiments were measured in triplicate and the deviation in the fluorescence was less than 7%.

To measure the binding constants, the fluorescence intensity at 660 nm was plotted as a function of the DNA concentration under 0.5 μM porphyrin and the data were fitted by KaleidaGraph (Synergy Software, PA) according to a 1:1 binding mode. Saturation in fluorescence was obtained using excess of DNA.

The binding stoichiometry between porphyrin and DNA was determined by Job's plot analysis. The total concentration of porphyrin and DNA was maintained at 2 μM and the porphyrin-to-DNA concentration ratio was sequentially varied.

UV/Vis Absorption Spectra and DNA Melting Temperature (T_m) Measurements. UV/Vis absorption spectra were acquired at a UV2550 spectrophotometer (Shimadzu Corp., Kyoto, Japan) using a quartz cuvette of a path length of 1 cm. The melting temperatures (T_m) of DNA in the absence and presence of porphyrin were determined using the same spectrophotometer but equipped with a TMSPC-8 T_m analysis accessory. A micro multi-cell with eight chambers was used to eliminate the temperature discrepancy between samples. The DNA absorbance at 260 nm as a function of solution temperature between 5 °C and 100 °C was collected in 0.5 °C increments, with a 30-second equilibration time applied after each temperature increment. We found that the possible changes in the porphyrin aggregation state upon increasing the solution temperature had no effect on the DNA T_m measurements.

Dynamic Light Scattering Measurements. DLS measurements for porphyrin aggregation and dispersion were performed using Malvern Zetasizer Nano ZS90 analyzer (Malvern, Worcs. UK) equipped with a 4 mW He-Ne laser (633 nm) and operating at 90° angle and at temperature of 20 °C.

Quantum Chemistry Calculation. All geometry optimization were carried out using the B3LYP functional with the 6-31G(d) basis set implemented in the Gaussian 09 program. To evaluate the solvent effect of polar water, the stationary points with lowest energies were re-optimized using B3LYP/6-311G(d,p) basis set and the polarizable continuum model.[40] The nature of the optimized structures was characterized by vibrational analysis, with which all stationary points were confirmed to be minima.

RESULTS AND DISCUSSION

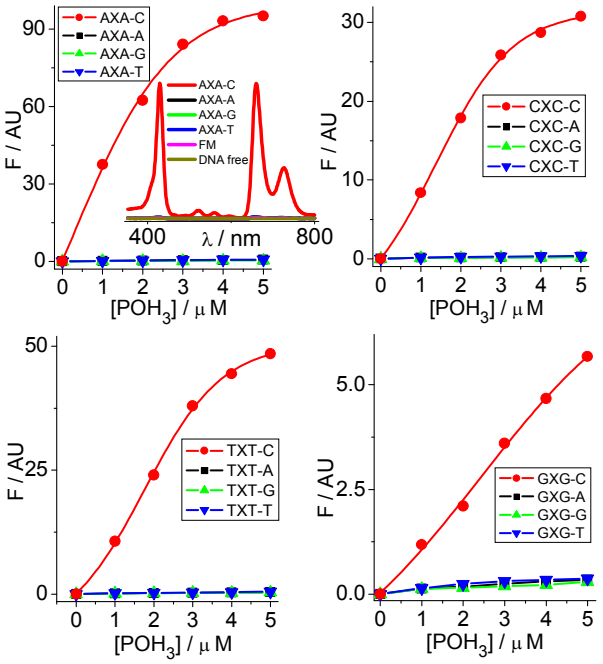


Figure 1. Dependence of POH_3 fluorescence at 660 nm in 0.1 M PBS buffer (pH 8.2) on the base identity of N and Y in duplex NXN-Y (1 μM) under various POH_3 concentrations. Excitation: 428 nm. Inset: typical excitation and emission spectra of 1 μM DNA and 5 μM POH_3.

Strong Dependence of Switching-on POH_3 Fluorescence on SNP Identity. The AP site-containing and other control DNA sequences used in this work were listed in Table S1. To evaluate the influence of neighboring bases on the SNP analysis, the target (Y) and flanking (N) bases just close to the AP site (X) were systematically varied in the DNA duplex (NXN-Y, Table S1). The SNP analysis was first investigated by measuring the fluorescence response of POH_3 in binding with NXN-Y in 0.1 M PBS buffer at pH 8.2 (Figure 1). POH_3 bears three hydroxyl groups at each phenyl ring, one at the para-position and two at the meta-positions, respectively (Scheme 1). At this condition, POH_3 is totally nonfluorescent alone in solution or in the presence of fully matched DNA (FM) without the AP site. Interestingly, only the NXN-C duplexes with the target cytosine opposite the AP site are capable of significantly lighting up the POH_3 fluorescence, although the turn-on absolute fluorescence intensity follows a flanking base N dependence in the order of A>T>C>G. Besides the strong Soret band in excitation spectra, the Q bands characteristic of porphyrin at the free base state (Scheme 1)[30-34] are also appreciable (Inset of Figure 1). In contrast, the fluorescence responses for NXN-A, -G, and -T are negligible even with the POH_3 concentration being several times higher than DNAs. Since the geometric size of the whole POH_3 molecule is larger than that of the AP site, insertion of a portion of POH_3 (most likely the hydroxyphenyl moiety) into the AP site via specific interaction with the target C should occur towards this high selectivity in SNP analysis. However, the binding mode should be also somewhat tuned by the stacking interaction with the base N flanking the AP site. Note that such superb selectivity accompanying the target turn-on response for all the flanking sequence environments hasn't been achieved in previous efforts.[16-21]

In order to evaluate the role of the hydroxyl substituent pattern of RU in achieving the SNP selectivity, POH_1 and POH_2 with respective one para-hydroxyl and two meta-hydroxyls at each phenyl ring were utilized as controls (Scheme 1). We found that although they emit somewhat higher or lower than POH_3 upon DNA binding, these hydroxyl patterns in POH_1 and POH_2 are detrimental to the fluorescence selectivity of SNP analysis (Figure S1), suggesting that the hydroxyl substituents in RU are very crucial to recognize the SNP target using the fluorogenic porphyrin as the readout SU. According to their structures, we assume that the perfectly complementary hydrogen bonding pattern of POH_3 and the target C (Scheme 2) is responsible for this superb selectivity, which is unfeasible for POH_1 and POH_2.

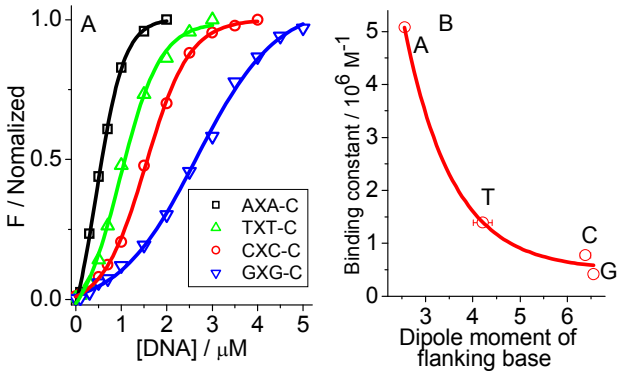


Figure 2. (A) Fluorescence titration of 0.5 μM POH_3 by NXN-C in 0.1 M PBS buffer (pH 8.2). (B) Dependence of the POH_3 binding constant of NXN-C on dipole moment of the flanking base of the AP site.

Fluorescence titration was used to estimate the binding affinity. Interestingly, as shown in Figure 2A, the POH_3 fluorescence follows a sigmoidal increase upon NXN-C titration. Assuming a 1:1 binding mode, the binding constants resolved accordingly with the Herderson-Hasselbalch equation are 5.1×10^6, 1.4×10^6, 7.7×10^5, and 4.1×10^5 M^{-1} for AXA-C, TXT-C, CXC-C, and GXG-C, respectively. This 1:1 binding mode was further confirmed by the Job's plot analysis (Figure S2). These affinities just follow the flanking N dependent fluorescence intensity responses as observed in Figure 1, suggesting that the binding affinity of POH_3 at the AP site, other than some photophysical processes such as photo-induced electron transfer, mainly determines the resultant fluorescence. We tried to resolve the critical factor that controls the affinity. As shown in Figure 2B, the obtained binding constants seem to be just exponentially and inversely dependent on the dipole moment of the flanking base N,[41,42] suggesting that POH_3 preferentially binds to the AP site with less polar flanking base. Conversely, it also means that the RU of POH_3 in binding with the AP site exists preferably at its less polar state among all possible isomers. Additionally, for NXN-Cs, the flanking base-tuned POH_3 fluorescence amplitude observed in Figure 1 doesn't strictly follow the order of the bases' electrochemical oxidation potentials (G<A<T<C).[43] Thus, the excited state electron transfer shouldn't be the critical factor to tune the fluorescence behavior of the bound POH_3 since the fluorogenic porphyrin SU in POH_3 is reasonably not the DNA binding moiety. This performance in selectivity for SNP analysis hasn't been achieved with the previously developed probes.[16-25] In those cases, except for a few inorganic probes,[21] the flanking guanine usually quenches the organic probes' fluorescence via electron transfer. The fluorescence quantum yield of POH_3 in binding with AXA-C was estimated to be about 0.11, comparable to other tetraphenylporphyrins.[44]

Mechanism of Superb Selectivity in SNP Identity Fluorescence Analysis. We can conclude that the SNP identity differentiation of C from A, G, and T can be achieved using POH_3

as the probe on the basis of the AP site design at pH≥8.2. However, the SNP selectivity is slightly tuned prior to this pH range (Figure S3) because of the easy AP site binding of POH_3 at the free base state (OH-N, Scheme 2). Simultaneously, at all the investigated pH range, POH_3 alone in solution or in the presence of FM is totally nonfluorescent. DLS experiments (Figure 3A) confirm the aggregation behavior of POH_3 in aqueous solution, although the aggregation in size is pH dependent (Figure S4). This could be caused by formation of J-aggregates.[33] At pH 8.2, AXA-C can disperse the aggregate (about 72 nm) to a much smaller size (about 1.3 nm), while AXA-A and AXA-G have a negligible effect on the POH_3 aggregate. More interestingly, AXA-T is also able to disperse the aggregate although to a size (2.7 nm) slightly higher than that caused by AXA-C. This opposite pyrimidine dependent binding was also confirmed by T_m experiments (Figure S5). However, the fluorescence results in Figure 1 suggest that the AXA-T-bound state of POH_3 at pH 8.2 exists still at the fluorescently dark state, very different from the AXA-C-determined fluorescent state. In order to have an insight into the effect of aggregation on the POH_3 fluorescence, a surfactant, TX-100, was used to disperse POH_3 in aqueous solution (Figure S6).[33] We found that TX-100 can disperse POH_3 to monomer in a pH-independent manner (Figure S6A,B) and the turn-on fluorescence is, however, strongly pH dependent (Figure S6C). At pH 8.2, although POH_3 is dispersed to the same extent as occurring at other pHs, it is still nonfluorescent, suggesting that the POH_3 monomer can isomerize to a fluorescently dark state at this pH. However, the POH_3 aggregate in the absence of TX-100 is nonfluorescent in a pH-independent manner (Figure S3). Upon binding to the AP site at pH≥8.2, only the DNA having the target base C opposite the AP site can turn POH_3 from the dark state to the emission state, as observed in Figure 1 and S3, while the binding to other DNAs with A, G, and T opposite the AP site still keeps POH_3 remaining at the dark state.

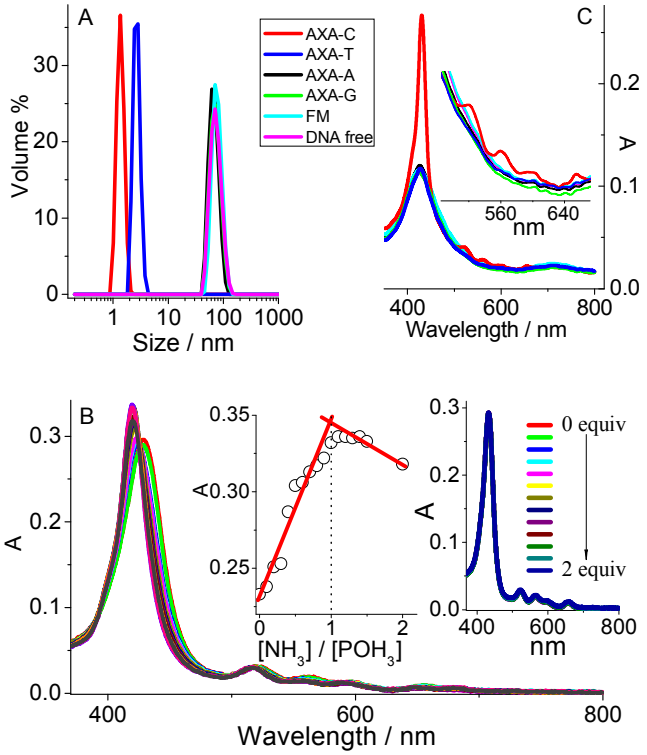


Figure 3. (A) Interaction of POH_3 (0.5 μM) with DNA (3 μM) investigated by DLS in 0.1 M PBS buffer (pH 8.2). (B) Absorption spectra of POH_3 (3 μM) in DMSO upon gradual addition of $NH_3{\cdot}H_2O$ till up to 6 μM. Inset: the absorbance at 420 nm with increasing the $NH_3{\cdot}H_2O$ concentration. Also shown in inset are the absorption spectra of POH_3 (3 μM) in DMSO upon gradual addition of NaOH till up to 6 μM. (C) Absorption spectra of 3 μM POH_3 in absence and presence of 9 μM DNA in 0.1 M PBS buffer (pH 8.2). For clarity, the changes in Q bands were enlarged as shown in Inset.

On the basis of the pH-dependent fluorescence of POH_3 in the TX-100 solution (Figure S6C), two deprotonation processes were expected when the solution pH was changed from 2 to 8. The numbers of involved protons in these deprotonation processes were estimated to be about 1.90 and 1.29, according to the calculation procedure (see Supporting Information for details),[30] suggesting that a two-proton transfer process is followed by a one-proton transfer process during deprotonation. It is well known that upon protonation of the cyclic tetrapyrrole nitrogens (two-proton process), besides the changes in the Soret band, the characteristic Q-bands are degenerated especially with disappearance of the absorption bands within the range of 500-650 nm.[31] Such protonation causes distortion of the tetrapyrrole macrocycle[31,32] and this structure distortion opens the pathway for radiationless deactivation,[32] as observed in Figure S6C with the resultant fluorescence decrease upon acidifying the solution from pH 4 to 2. We then compared the absorption spectra of POH_n (n=1, 2, 3) dispersed with TX-100. As shown in Figure S7, POH_2 that bears two hydroxyl groups only at the meta-positions of phenyl ring (Scheme 1) isn't so strongly affected in spectra profile within the pH range of 2~11 as POH_1 and POH_3. The peaks corresponding to protonation of the two nitrogens in the pyrrole macrocycle at acidic solution can't be observed within the 450-500 nm and near-infrared range,[33,34] while the peaks for the Q bands characteristic of the free base form of the macrocycle always exist in the investigated pH range. These facts suggest that the overall structure of POH_2 is negligibly affected. However, for POH_3 and POH_1, the peaked bands for the nitrogen diprotonation at acidic solution appear at 466/706 and 453/487/780 nm, respectively. Different from POH_2, the two porphyrins have a hydroxyl group at the para-position of phenyl ring in spite of other two meta-hydroxyl groups occurring to POH_3. The Q bands of the free base forms of POH_3 and POH_1 experience disappearance, appearance, and re-disappearance upon changing the solution pH from 2 to 11. POH_3 seems to be more pH-sensitive than POH_1 in disappearance of the Q bands at neutral and basic solution. These results mean that the hydroxyl group at the para-position of phenyl ring has a more electronic communication with the nitrogen atoms of the pyrrole macrocycle. We propose that the presence of this para-hydroxyl group will favor the protonation in the nitrogens by the prototropic allostery process (Scheme 2). Accordingly, in POH_3, the para-hydroxyl group should be preferably deprotonated at neutral and weakly basic pH range with respect to the hydroxyl groups at the meta-positions, as observed in the pH-dependent absorption spectra in Figure S7. The changes in absorbance between 420-430 nm within pH 7-8.5 also predict a deprotonation process involving one proton (Inset of Figure S7D). The preferential deprotonation of the para-hydroxyl groups in POH_3 was also confirmed using NMR by observation of the prior fading of the 8.57 ppm peak with respect to the 9.29 ppm peak for meta-hydroxyl groups in DMSO-*d6* after addition of trace amount of $NH_3{\cdot}H_2O$ (Figure S8). During this deprotonation process, the NMR peak at 8.93 ppm for the protons attached to the pyrrole *β*-positions becomes broader and the peak at -2.94 ppm[48] for the NH protons attached to the pyrroles slightly increases, suggesting that restructuring of POH_3 occurs to some extent upon deprotonation,[45] although this conformational change does not apparently shift the NMR peaks due to the usage of DMSO and $NH_3{\cdot}H_2O$. The absorption spectra of POH_3 in DMSO upon addition of $NH_3{\cdot}H_2O$ exhibits a slight blue-shift in the Soret band from 428 to 420 nm with an isobestic point occurring at about 425 nm (Figure 3B), suggesting conversion of POH_3 between two structures (most likely OH-N↔O^--N↔O-N^- with small amount of

O-NH, Scheme 2). However, addition of $NH_3{\cdot}H_2O$ results in relative minor changes in the Q bands, suggesting the formation of small amount of O-NH. Interestingly, The absorbance changes at 420 nm (Inset of Figure 3B) indicate that 1 equiv of $NH_3{\cdot}H_2O$ is enough to implement the first deprotonation process, suggesting preferential loss of one proton in the initial deprotonation (Scheme 2). However, interestingly, addition of up to 2 equiv of NaOH gives no alteration with the POH_3 spectra profile (Inset of Figure 3B), suggesting that the electronic structure of tetrapyrrole macrocycle is unaffected upon deprotonation using the properly quantitative NaOH and POH_3 should at the O^--N state, as opposed to the O-N^- state (although in a small amount of O-NH) stabilized by $NH_3{\cdot}H_2O$ possibly via hydrogen bonding interactions.[46] Additionally, NaOH isn't a good hydrogen bonding donor and is more alkalic than $NH_3{\cdot}H_2O$ and thus unable to support the formation of O-N^- (and also O-NH). The formation of small amount of O-NH also indicates that the basicity of the O-N^- state is weaker than $NH_3{\cdot}H_2O$ (pK_b=4.75), but in aqua buffer solution at pH 8.2, it can reach the nonfluorescent O-NH state (Figure S6C and Scheme 2).

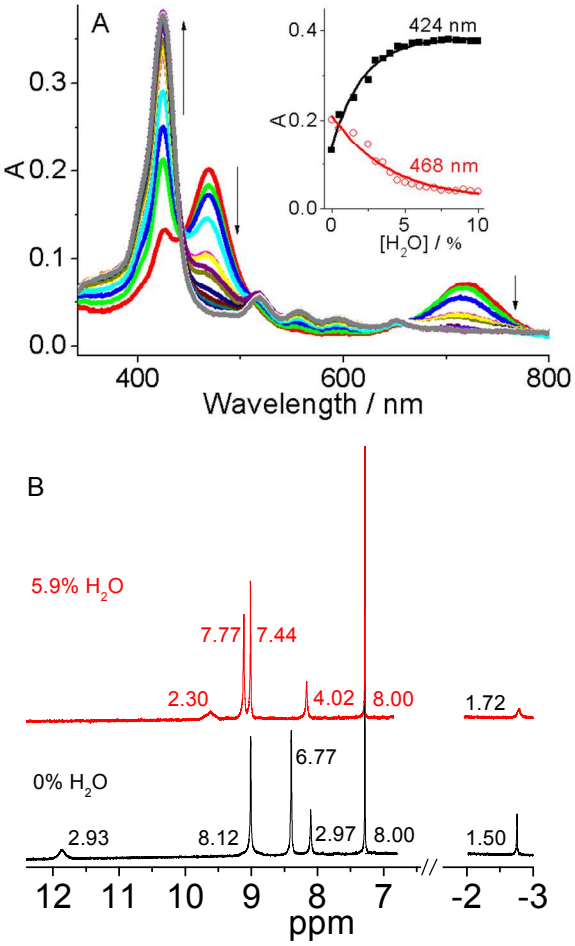
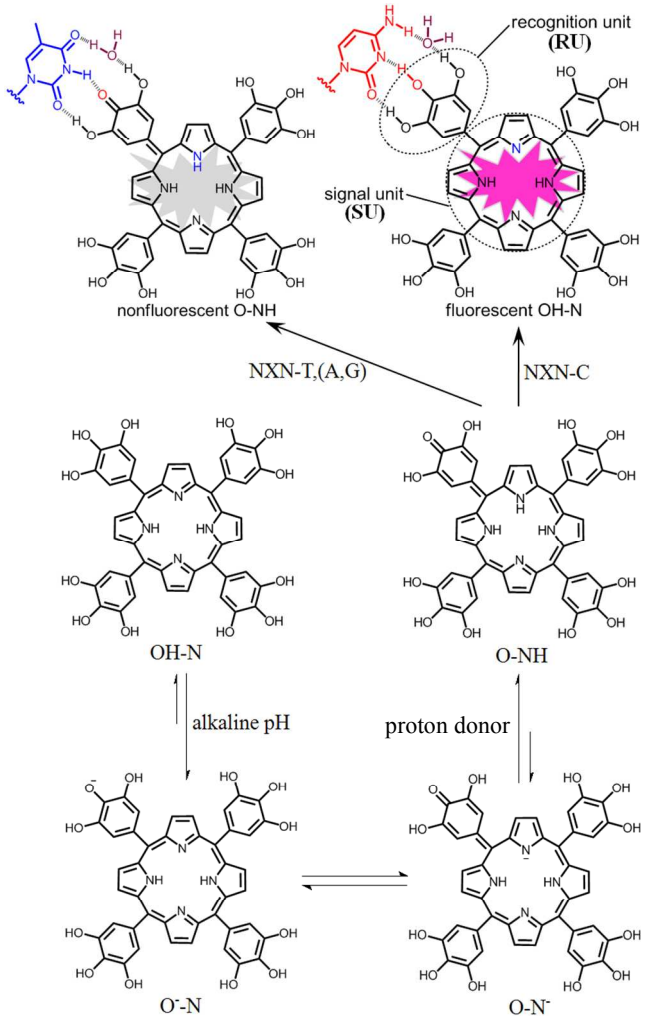


Figure 4. (A) Absorption spectra of POH_3 (3 μM) in acetone upon gradual addition of H_2O till up to 10 %. (B) ^{1}H NMR spectra of POH_3 in acetone-*d*6 upon addition of 5.9 % H_2O (v/v). All the indicated numbers are integral area reported relative to the peak at 7.29 ppm for 8 protons of $H_{o\text{-}Ph}$. Note that in either case OH-N should coexists with O-NH and the indicated proton numbers only illustrate their relative changes.

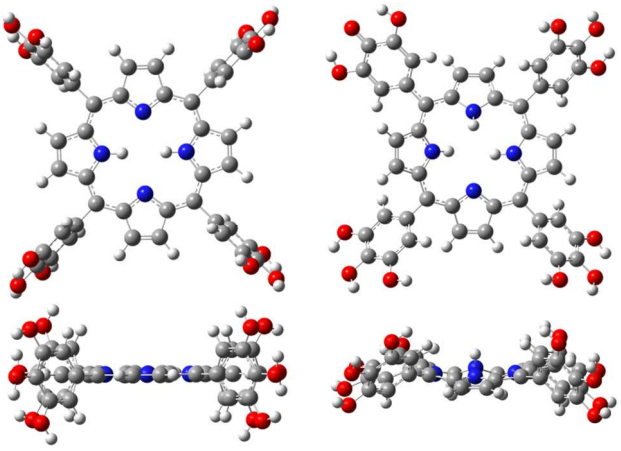
Since the NH deprotonation of the free base tetrapyrrole macrocycle occurs at pH beyond 16,[30] the disappearance of the characteristic Q bands of the tetrapyrrole macrocycle at pH 8.2 (Figure S7) should be caused by tautomerization to the quinone-like product (O-NH, Scheme 2). The fluorescent darkness of POH_3 at pH≥8.2 in either aggregate state (Figure S3) or TX-100-dispersed state (Figure S6) indicates that tautomerization of the tetrapyrrole macrocycle should also take place upon deprotonation of the para-hydroxyl substituent.[32] It has been reported that the phenyl porphyrin with para-hydroxyl substituent can tautomerize to quinone-like structure in solvent with an appropriate polarity.[45-48] Additionally, monoprotonation of the tetrapyrrole macrocycle has been observed.[32,46, 49 - 51] In place of the porphyrin aromaticity,[52 , 53] this tautomerization possibly produces an extended phenoquinoid conjugated structure with an interrupted π-system[47] via delocalizing the left electron onto the tetrapyrrole macrocycle (OH-N→O^--N→O-N^-→O-NH, Scheme 2).[54] The presence of alkaline can catalyze this deprotonation-driven tautomerization.[48] The quinone-like structure is easily further stabilized by the J-aggregate formation via intermolecular hydrogen bonding.[33]



Scheme 2. Prototropic allostery of POH_3 and the rationale for the DNA analysis with superb selectivity. During conversion from O-N^- to O-NH using a proton donor (such as HPO_4^{2-} or H_2O), the whole proton donor is most likely to totally bind at the N^- position, not only H^+. This is omitted for clarity.

Our experimental results should be explained by tautomerization of POH_3 to the quinone-like structure (O-NH, Scheme 2) at an appropriate pH range. However, the AP site with the right target base should be more acidic in microenvironment for the POH_3 binding than its existing solution of weakly basic pH, as observed in Figure 2A with a characteristic sigmoidal curve analogous to an acid-base titration interaction. Upon binding to

the AP site, the quinone-like structure (O-NH, Scheme 2) recovers again to the enol structure (OH-N, Scheme 2), as observed in Figure 3C with re-appearing of the characteristic Q bands at 520/560/600/654 nm upon binding with AXA-C. As demonstrated in Figure 3C, this recovery to OH-N occurs only to the target base C, reflecting critical role of the hydrogen bonding pattern of the target base C in converting the DNA-bound POH_3 to the fluorescently emitting state of OH-N (Scheme 2). Note that only the base C has a complementary hydrogen bonding pattern with the free base form (OH-N) of POH_3, while the base T is just complementary in hydrogen bonding with its dark state (O-NH, Scheme 2). This is in accordance with the occurrence of POH_3 binding with AXA-T as observed in Figure 3A, but giving a nonfluorescent behavior as observed in Figure 1. For DNAs with the base A and G opposite the AP site, a less space is left for the POH_3 binding at pH≥8.2 because of the deformed structure in O-NH, as observed in Figure 3A with lack of the POH_3 dispersion in the presence of AXA-A and AXA-G.



Scheme 3. DFT optimized structures of POH_3 in the OH-N (left) and O-NH (right) states. The top-view (top) and side-view (bottom) conformations are presented.

Because of the possible strong binding capacity of DMSO with the tetrapyrrole macrocycle to inhibit the POH_3 structure conversion,[46] as confirmed in Figure 3B and S8, acetone was further used as the alternative solvent (note that POH_3 is insoluble in weak polar solvents such as dichloromethane and chloroform). As shown in Figure 4A, in acetone, POH_3 displays absorption bands at 424, 468, 716 nm, respectively, somewhat different from the TX-100 dispersed POH_3 in acidic solution with the absorption bands at 428, 466, 708 nm, suggesting the varied existing state. Addition of trace amount of H_2O increases the 424 nm absorption band and also the usual Q bands at the expense of 468 and 716 nm absorption bands thus with an isobestic point occurring at 442 nm. Interestingly, ^{1}H NMR of POH_3 in acetone-*d6* exhibits a NH resonance of the quinone-like structure (O-NH) at 11.83 ppm (Figure 4B).[48] The peak at 8.10 ppm predicts three O$H_{p\text{-Ph}}$ protons relative to the eight $H_{o\text{-Ph}}$ protons at 7.30 ppm. However, addition of 5.9% H_2O increases the O$H_{p\text{-Ph}}$ resonance to four protons at the expense of the NH resonance of the quinone-like structure, strongly suggesting the conversion from the quinone-like structure (O-NH) to the enol structure (OH-N). Following this conversion, addition of H_2O shifts the other ^{1}H NMR peaks except for the $H_{o\text{-Ph}}$ and $H_{\beta\text{-pyrr}}$ resonances, since this structure conversion mainly affects the distribution of electronic densities in the involved oxygen and nitrogen atoms. These results also suggest that the existing state of POH_3 is very sensitive to the solvent environment. Additionally, although the DNA binding experiments were carried out in aqueous solution, we expected that the quinone-like structure most possibly also existed in aqueous solution since the POH_3 aggregate should provide a right polar microenvironment to favor such structure via intermolecular hydrogen bonding.[33]

Our DFT calculation (Scheme 3) shows that following structure conversion of POH_3 from OH-N to O-NH, the tetrapyrrole macrocycle distorts to a nonplanar conformer, in agreement with the darkness in fluorescence at pH 8.2. Additionally, the remained phenyl substituents can thus adopt an orientation of more coplanarity with respect to the tetrapyrrole macrocycle, as does the quinone-like phenyl substituent. Thus, in O-NH, the remained phenyl substituents should have a stronger electronic communication with the tetrapyrrole macrocycle [55] and this should increase the electron density of the three remained para-hydroxyl groups, as demonstrated in the small chemical shift of ^{1}H NMR of POH_3 in acetone-*d6* in comparison to addition of H_2O (Figure 4B). This would make these hydroxyl groups deprotonating at higher pH than 8.2. Thus, the first step in deprotonation only involves dissociation of one para-hydroxyl group in favor of monoprotonation of the tetrapyrrole macrocycle,[32,46,49-51] and the quinone-like phenyl substituent prefers to bind to the AP site. Therefore, we only observed a 1:1 binding of POH_3 (in the O-NH state) with NXN-C at pH 8.2.

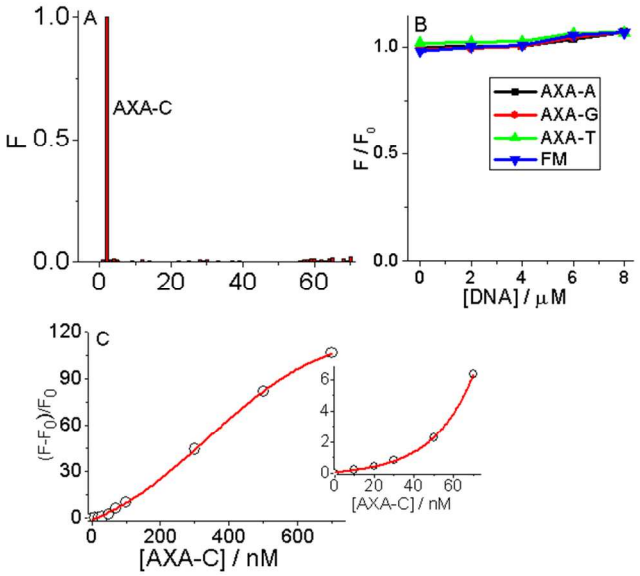


Figure 5. (A) Fluorescence response of POH_3 (5 μM) at 660 nm in 0.1 M PBS K^+ buffer (pH 8.2) to the variant DNA sequences (1 μM) including ss-DNAs and others with potentials to form hairpin, G-quadruplex, triplex, and i-motif structures. (B) Effect of coexisted DNA on the POH_3 (5 μM) fluorescence in 0.1 M PBS buffer (pH 8.2) containing AXA-C (1 μM). (C) Fluorescence response of POH_3 (5 μM) to AXA-C at low concentration. Inset: AXA-C at tens of nM level.

Performance of Superb Selectivity in SNP Analysis Using POH_3 as Probe. As discussed above, both the base stacking and pairing interactions are crucial to light up the POH_3 fluorescence. This should provide a chance to achieve a superb selectivity in SNP analysis using POH_3 as the AP site-specific probe. We then tested 70 sequences including ss-DNAs and others with potentials to form secondary structures of hairpin, G-quadruplex, triplex, and i-motif, although the herein used experimental conditions aren't appropriate to bring some of these structures (such as triplex and i-motif). The detailed sequences were listed in Table S1. As expected, only the AP site-containing ds-DNA (using AXA-C as an example) is efficient to fluorescently light up POH_3 (Figure 5A). Note that the homo-poly(dC_{15}), poly(dT_{15}), poly(dG_{15}), and poly(dA_{15}) behave as inefficiently as the two ss-DNAs composing of the AXA-C duplex in lighting up POH_3. The hairpin DNAs with the exposed C loops of variant lengths (from 6C-loop to 9C-loop, Table S1) also keep POH_3 at the dark state.

Additionally, the G-quadruplex structures involving the AP site at the different locations along the strands (from Quad-AP1 to Quad-AP4, Table S1)[56] are as well unable to bring emissive POH_3. Additionally, the superb selectivity for the C base in the SNP identity analysis using the AP site design is also confirmed by the coexistence of the target AXA-C with the excess of AXA-T, AXA-G, AXA-A, and FM, respectively (Figure 5B). The limit of detection for the AXA-C analysis is about 30 nM (Figure 5C). Therefore, our method has promising applications in sample analysis, for example, the PCR product detection.[18]

CONCLUSIONS

We develop a promising rationale using POH_3 as the novel fluorescent probe to implement a superb selectivity of nucleotide recognition in DNA. RU in POH_3 is physically separated from SU but a mutual electronic communication remains between them. The fluorescent behavior of SU is strongly dependent of its existing state with respect to RU. The quinone-like state of O-NH is nonfluorescent. The complementary hydrogen bonding of RU in O-NH with the target cytosine opposite the AP site switches on the SU fluorescence through prototropic allostery to the fluorescent OH-N state. However, this allostery can't come true when the bases opposite the AP site are thymine, guanine, and adenine. Because of the limited space provided by the AP site and the SU separation from RU, only RU mainly enters into the AP site *via* hydrogen bonding and stacking interactions and SU isn't directly involved in the interaction with the AP site. Thus, the turn-on fluorescence remains when the flanking base is guanine, the most easily oxidizable base in DNA. This tolerance to the flanking base identity has seldom been achieved with the previously developed probes. Additionally, other DNA structures are unable to bring this allostery, indicating that the combination recipe of the AP site design and the dual-function porphyrin will find wide applications in selective DNA-based sensors.

ASSOCIATED CONTENT

Supporting Information

Calculation procedures of proton transfer number and fluorescence quantum yield, oligonucleotides used in this work, emisison spectra of POH_n, Job′s plot analysis, pH dependent SNP selectivity, pH dependent DLS spectra, T_m results, DLS, fluorescence, and absorption spectra in TX-100, and ^1H NMR spectra of POH_3 in DMSO-NH_3·H_2O. This material is available free of charge via the Internet at http://pubs.acs.org.

ORCID

Yong Shao: 0000-0003-0834-6244
Zhiai Xu: 0000-0002-4391-2507

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

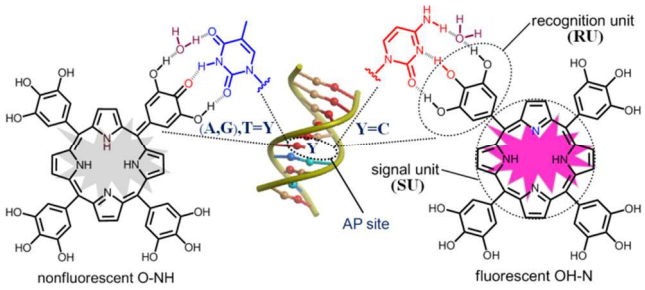
This work was supported by the National Natural Science Foundation of China (Grant No. 21675142, 21545009, and 21675055).

REFERENCES

(1) Trejo-de la, A.; Hernandez-Sancen, P.; Maldonado-Bernal, C.; *Genes*. *Immun*. **2014**, *15*, 199–209.
(2) Metzker, M. L. *Nat*. *Rev*. *Genet*. **2010**, *11*, 31–46.
(3) Venkatesan, B. M.; Bashir, R. *Nat*. *Nanotechnol*. **2011**, *6*, 615–624.
(4) Shen, W.; Tian, Y.; Ran, T.; Gao, Z. Q. *TrAC-Trend. Anal*. *Chem*. **2015**, *69*, 1–13.
(5) Jia, X. F.; Li, J.; Han, L.; Ren, J. T.; Yang, X.; Wang, E. K. *ACS Nano* **2012**, *6*, 3311–3317.
(6) Guo, W. W.; Yuan, J. P.; Dong, Q. Z.; Wang, E. K. *J*. *Am*. *Chem*. *Soc*. **2010**, *132*, 932−934.
(7) Xiao, Y.; Lou, X. H.; Uzawa, T.; Plakos, K. J. I.; Plaxco, K. W.; Soh, H. T. *J*. *Am*. *Chem*. *Soc*. **2009**, *131*, 15311–15316.
(8) Wang, H. Q.; Liu, W. Y.; Wu, Z.; Tang, L. J.; Xu, X. M.; Yu, R. Q. *Anal*. *Chem*. **2011**, *83*, 1883–1889.
(9) Li, J. S.; Deng, T.; Chu, X.; Yang, R. H.; Jiang, J. H.; Shen, G. L.; Yu, R. Q. *Anal*. *Chem*. **2010**, *82*, 2811–2816.
(10) Wu, S.; Liang, P. P.; Yu, H. X.; Xu, X. W.; Liu, Y.; Lou, X. H.; Xiao, Y. *Anal*. *Chem*. **2014**, *86*, 3461–3467.
(11) Wu, T. B.; Xiao, X. J.; Gu, F. D.; Zhao, M. P. *Chem*. *Commun*. **2015**, *51*, 17402–17405.
(12) Hocek, M.; Fojta, M. *Chem*. *Soc*. *Rev*. **2011**, *40*, 5802–5814.
(13) Greco, N. J.; Tor, Y. Z. *Nat*. *Protoc*. **2007**, *2*, 305–316.
(14) Sinkeldam, R.; Greco, W. N.; Tor, Y. *Chem*. *Rev*. **2010**, *110*, 2579–2619.
(15) Venkatesan, N.; Seo, Y. J.; Kim, B. H. *Chem*. *Soc*. *Rev*. **2008**, *37*, 648–663.
(16) Zeglis, B. M.; Boland, J. A.; Barton, J. K. *J*. *Am*. *Chem*. *Soc*. **2008**, *130*, 7530–7531.
(17) Kausar, A.; Osman, E. A.; Gadzikwa, T.; Gibbs-Davis, J. M. *Analyst* **2016**, *141*, 4272–4277.
(18) Sato, Y.; Nishizawa, S.; Yoshimoto, K.; Seino, T.; Ichihashi, T.; Morita, K.; Teramae, N. *Nucleic Acids Res*. **2009**, *37*, 1411–1422.
(19) Yoshimoto, K.; Nishizawa, S.; Minagawa, M.; Teramae, N. *J*. *Am*. *Chem*. *Soc*. **2003**, *125*, 8982–8983.
(20) Xu, S.; Shao, Yong.; Ma, K.; Cui, Q.; Liu, G.; Wu, F.; Li, M. *Analyst* **2011**, *136*, 4480–4485.
(21) Ma, K.; Cui, Q.; Liu, G.; Wu, F.; Xu, S.; Shao, Y. *Nanotechnology* **2011**, *22*, 305–502.
(22) Takei, F.; Igarashi, M.; Agihara, M.; Oka, Y.; Soya, Y.; Nakatani, K. *Angew*. *Chem*. *Int*. *Ed*. **2009**, *48*, 7822–7824.
(23) Verma, R. K.; Takei, F.; Nakatani, K. *Org*. *Lett*. **2016**, *18*, 3170–3173.
(24) Ihara, T.; Uemura, A.; Futamura, A.; Shimizu, M.; Baba, N.; Nishizawa, S.; Teramae, N.; Jyo, A. *J*. *Am*. *Chem*. *Soc*. **2009**, *131*, 1386–1387.
(25) Futamura, A.; Uemura, A.; Imoto, T.; Kitamura, Y.; Matsuura, H.; Wang, C. X.; Ichihashi, T.; Sato, Y.; Teramae, N.; Nishizawa, S.; Ihara, T. *Chem*.-*Eur*. *J*. **2013**, *19*, 10526–10535.
(26) Xiang, Y.; Tong, A. J.; Lu, Y. *J*. *Am*. *Chem*. *Soc*. **2009**, *131*, 15352–15357.
(27) Wu, T.; Xiao, X.; Gu, F.; Zhao, M. *Chem. Commun.* **2015**, *51*, 17402–17405.
(28) Xu, Z.; Morita, K.; Sato, Y.; Dai, Q.; Nishizawa, S.; Teramae, N. *Chem*. *Commun*. **2009**, *45*, 6445–6447.
(29) Li, M. J.; Sato, Y.; Nishizawa, S.; Seino, T.; Nakamura, K.; Teramae, N. *J*. *Am*. *Chem*. *Soc*. **2009**, *131*, 2448–2449.
(30) Li, Y.; Sen, D. *Chem*. *Biol*. **1998**, *5*, 1–12.
(31) Chauvin, B.; Kasselouri, A.; Chaminade, P.; Quiameso, R.; Nicolis, I.; Maillard, P.; Prognon, P. *Anal*. *Chim*. *Acta* **2011**, *705*, 306–314.
(32) Chirvony, V. S.; van Hoek, A.; Galievsky, V. A.; Sazanovich, I.V.; Schaafsma, T. J.; Holten, D. *J*. *Phys*. *Chem*. *B* **2000**, *104*, 9909–9917.
(33) Li, X.; Zheng, Z.; Han, M.; Chen, Z.; Zou, G. *J*. *Phys*. *Chem*. *B* **2007**, *111*, 4342–4348.
(34) De Luca, G.; Romeo, A.; Scolaro, L. M. *J*. *Phys*. *Chem*. *B* **2006**, *110*, 14135–14141.
(35) Li, T.; Wang, E.; Dong, S. *Anal*. *Chem*. **2010**, *82*, 7576–7580.
(36) Nicoludis, J. M.; Miller, S. T.; Jeffrey, P. D.; Barrett, S. P.; Rablen, P. R.; Lawton, T. J.; Yatsunyk, L. A. *J*. *Am*. *Chem*. *Soc*. **2012**, *134*, 20446–20456.

(37) Huang, X. X.; Zhu, L. N.; Wu, B.; Huo, Y. F.; Duan, N. N.; Kong, D. M. *Nucleic Acids Res*. **2014**, *42*, 8719–8731.
(38) Wang,Y.; Hu,Y.-H.; Wu, T.; Liu, H.; Zhang, L.-H.; Zhou, X.-S.; Shao Y. *Analyst* **2015**, *140*, 5169–5175.
(39) Banfi, S.; Caruso, E.; Caprioli, S.; Mazzagatti, L.; Canti, G.; Ravizza, R.; Gariboldi, M.; Monti, E. *Bioorg*. *Med*. *Chem*. **2004**, *12*, 4853–4860.
(40) Barone, V.; Cossi, M.; Tomasi, J. *J*. *Comput*. *Chem*. **1998**, *19*, 404–417.
(41) Hobza, P.; Šponer, J. *Chem*. *Rev*. **1999**, *99*, 3247–3276.
(42) Šponer, J.; Hobza, P. *Biopolymers* **2002**, *61*, 3–31.
(43) Seidel, C. A. M.; Schulz, A.; Sauer, M. H. M. *J*. *Phys*. *Chem*. **1996**, *100*, 5541–5553.
(44) Ormond, A. B.; Freeman, H. S. *Dyes Pigments* **2013**, *96*, 440–448.
(45) Xie, Y.; Hill, J. P.; Schumacher, A. L.; Karr, P. A.; D'Souza, F.; Anson, C. E.; Powell, A. K.; Ariga, K. *Chem*.-*Eur*. *J*. **2007**, *13*, 9824–9833.
(46) Honda, T.; Kojima, T.; Fukuzumi, S. *Chem*. *Commun*. **2009**, 4994–4996.
(47) Li, M.; Chen, X. Y.; Zou, J. Z.; Xu, Z.; You, X. Z. *Chin*. *J*. *Chem*. **1996**, *14*, 20–24.
(48) Ding, Y.; Li, X.; Hill, J. P.; Ariga, K.; Ågren, H.; Andréasson, J.; Zhu, W.; Tian, H.; Xie, Y. *Chem*.-*Eur*. *J*. **2014**, *20*, 12910–12916.
(49) De Luca, G.; Romeo, A.; Scolaro, L. M.; Ricciardi, G.; Rosa, A. *Inorg*. *Chem*. **2007**, *46*, 5979–5988.
(50) Rudine, A. B.; DelFatti, B. D.; Wamser, C. C. *J*. *Org*. *Chem*. **2013**, *78*, 6040−6049.
(51) Corwin, A. H.; Chivvis, A. B.; Poor, R. W.; Whitten, D. G.; Baker, E. W. *J*. *Am*. *Chem*. *Soc*. **1968**, *90*, 6577–6583.
(52) Milgrom, L. R. *Tetrahedron* **1983**, *39*, 3895–3898.
(53) Hill, J. P.; Schumacher, A. L.; D'Souza, F.; Labuta, J.; Redshaw, C.; Elsegood, M. R. J.; Aoyagi, M.; Nakanishi, T.; Ariga, K. *Inorg*. *Chem*. **2006**, *45*, 8288–8296.
(54) Milgrom, L. R.; Mofidi, N.; Harriman, A. *J*. *Chem*. *Soc*. *Perkin Trans*. *II*, **1989**, 805–809.
(55) Goldberg, P. K.; Pundsack, T. J.; Splan, K. E. *J*. *Phys*. *Chem*. *A* **2011**, *115*, 10452–10460.
(56) Školáková, P.; Bednářová, K.; Vorlíčková, M.; Sagi, J. *Biochem*. *Biophys*. *Res*. *Commun*. **2010**, *399*, 203–208.

For TOC only


recognition unit
(RU)
(A,G)-T=Y
Y=C
signal unit
(SU)
AP site
nonfluorescent O-NH
fluorescent OH-N