ACES Asian Chemical Editorial Society

DOI: 10.1002/asia.201600527

CHEMISTRY AN ASIAN JOURNAL Full Paper

Binding to Trinucleotide Repeats

# Naphthyridine-Benzoazaquinolone: Evaluation of a Tricyclic System for the Binding to $(CAG)_n$ Repeat DNA and RNA

Jinxing Li [+],[a] Akihiro Sakata[+],[a] Hanping He,[b] Li-Ping Bai,[c] Asako Murata,[a] Chikara Dohno,[a] and Kazuhiko Nakatani*[a]

**Abstract:** The expansion of CAG repeats in the human genome causes the neurological disorder Huntington's disease. The small-molecule naphthyridine-azaquinolone **NA** we reported earlier bound to the CAG/CAG motif in the hairpin structure of the CAG repeat DNA. In order to investigate and improve **NA**-binding to the CAG repeat DNA and RNA, we conducted systematic structure-binding studies of **NA** to CAG repeats. Among the five new **NA** derivatives we synthesized, surface plasmon resonance (SPR) assay showed that all of the derivatives modified from amide linkages in **NA** to a carbamate linkage failed to bind to CAG repeat DNA and RNA. One derivative, **NBzA**, modified by incorporating an additional ring to the azaquinolone was found to bind to both $d(CAG)_9$ and $r(CAG)_9$. **NBzA** binding to $d(CAG)_9$ was similar to **NA** binding in terms of large changes in the SPR assay and circular dichroism (CD) as well as pairwise binding, as assessed by electron spray ionization time-of-flight (ESI-TOF) mass spectrometry. For the binding to $r(CAG)_9$, both **NA** and **NBzA** showed stepwise binding in ESI-TOF MS, and **NBzA**-binding to $r(CAG)_9$ induced more extensive conformational change than **NA**-binding. The tricyclic system in **NBzA** did not show significant effects on the binding, selectivity, and translation, but provides a large chemical space for further modification to gain higher affinity and selectivity. These studies revealed that the linker structure in **NA** and **NBzA** was suitable for the binding to CAG DNA and RNA, and that the tricyclic benzoazaquinolone did not interfere with the binding.

## Introduction

The expansion of trinucleotide repeat (TNR) sequences in the human genome is known as causative of more than 40 hereditary neurological disorders.[1–4] The CAG sequence in the coding region of the protein huntingtin is 4–100 repeats for patients with Huntington's disease, whereas that is about 6–36 repeats for healthy individuals.[5,6] Aberrant and enormous expansion of the CTG sequence of 5–38 repeats in the 3′-UTR of the DMPK gene up to 40–1000 repeats was observed for patients with myotonic dystrophy type 1.[7–9] The CGG repeat expansion causes Fragile X syndrome upon expansion of up to more than 200 repeats.[10,11] These TNR sequences share the common sequence motif of CXG, where X is A, T, C, and G, while there are other types of TNR such as GAA, which causes Friedreich's ataxia upon expansion in the FXN gene.[12,13]

The mechanisms of dynamic instability of TNR sequence involving repeat expansion and contraction depend on the type of TNR sequences, the involvement of repair systems, and the type of cells and tissues.[14–17] The common feature of the mechanism in TNR expansion and contraction, however, involves non-canonical secondary structures produced on the TNR sequences.[18–23] Due to C-G base-pairing, the expanded CXG repeats can fold into a hairpin secondary structures containing a number of the palindromic CXG/CXG motif, which

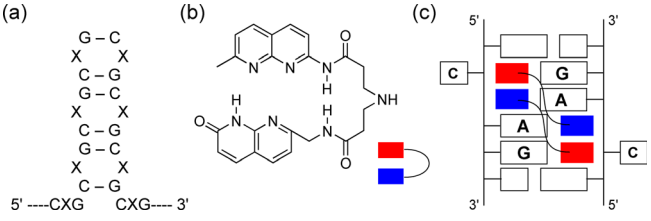



**Figure 1.** (a) A possible hairpin secondary structure of $(CXG)_n$, (b) the structure of naphthyridine-azaquinolone (**NA**), and (c) 2:1 binding mode of **NA** to a 5′-d(CAG)-3′/5′-d(CAG)-3′ motif. The red and blue blocks represent naphthyridine and azaquinolone units, respectively.

[a] Dr. J. Li,[+] A. Sakata,[+] Prof. Dr. A. Murata, Prof. Dr. C. Dohno, Prof. Dr. K. Nakatani
Department of Regulatory Bioorganic Chemistry
The Institute of Scientific and Industrial Research
Osaka University
8-1 Mihogaoka, Ibaraki 567-0047 (Japan)
E-mail: nakatani@sanken.osaka-u.ac.jp

[b] Prof. Dr. H. He
Chemistry and Chemical Engineering College
Hubei University
Road Youyi 368, Wuchang, Wuhan, Hubei 430062 (China)

[c] Prof. Dr. L.-P. Bai
State Key Laboratory of Quality Research in Chinese Medicine, and Macau Institute for Applied Research in Medicine and Health
Macau University of Science and Technology
Avenida Wai Long, Taipa, Macau (China)

[+] These authors contributed equally to these work.

Supporting information for this article can be found under http://dx.doi.org/10.1002/asia.201600527.




These are not the final page numbers! ↗↗

hold a X-X mismatch flanked by two C-G base pairs (Figure 1 a). One of the mechanisms of TNR expansion is believed to involve the polymerase stall during the replication of TNR sequence and subsequent slippage of the TNR sequence to form hairpin structures on the synthesizing strand.[22,23] Since the stability of the hairpin secondary structures involving CXG/CXG motifs increases as the repeat length increases, the expanded CXG sequence is more susceptible to the formation of the hairpin structure and more likely to undergo further expansion.[24–26]

We have discovered small molecules that selectively bind to the CXG/CXG motifs and promote the formation of hairpin secondary structures on the CXG repeat sequences.[27–30] Naphthyridine-azaquinolone **NA** is the first discovered molecule that binds to the CAG/CAG motif produced in the hairpin secondary structure of the CAG repeat DNA sequence (Figure 1 b).[27] The **NA**-bound structure on the CAG/CAG motif was determined by NMR spectroscopy to have the following structural characteristics (Figure 1 c): 1) the naphthyridine unit in **NA** bound to the guanine by three hydrogen bonds, 2) cytosine that base-paired to the guanine was forced to flip out from the π-stack and had little effect on the stability of the complex, 3) the azaquinolone unit in **NA** bound to the adenine by two hydrogen bonds, 4) two molecules of **NA** simultaneously bound to the CAG/CAG motif, and 5) naphthyridine and azaquinolone units in a different molecule were well stacked with each other and neighboring base pairs. In addition to the $d(CAG)_n$ repeat, $r(CAG)_n$ repeat, a transcript of the expanded $d(CAG)_n$ sequence has drawn significant recent scientific interests because the expanded $r(CAG)_n$ repeat was found to be translated without the presence of the AUG start codon.[31–33] This phenomenon is called repeat associated non-ATG translation (RAN translation) and is supposed to involve the non-canonical hairpin structure of the expanded $r(CAG)_n$ repeat. Small molecules binding to the hairpin secondary structure of the CAG repeat RNA with high affinity hold potentials to modulate the repeat expansion and contraction, and RAN translation.

To investigate the binding of **NA** to the $r(CAG)_n$ repeat and further improvements of the **NA** structure regarding the affinity and selectivity not only to $d(CAG)_n$ but also to $r(CAG)_n$ repeats, we conducted systematic structure-binding studies of **NA** on the binding to these CAG repeat sequences. Here we describe the effect of the structural modification of the linker connecting naphthyridine and azaquinolone units and the expanded aromatic system from bicyclic azaquinolone to the tricyclic benzoazaquinolone on the binding to $d(CAG)_9$ and $r(CAG)_9$ investigated by SPR assay, CD spectra, and ESI-TOF MS. These studies revealed that, among five derivatives of **NA**, the ring-expanded derivative **NBzA** with the amide linkage bound to $d(CAG)_n$ repeat in a pairwise manner similar to **NA**. In contrast, **NA** and **NBzA** showed the binding to $r(CAG)_n$ repeat with decreased affinity, and the binding was not pairwise but in a stepwise manner. The change in CD spectra of $r(CAG)_9$ was more extensive for **NBzA** than **NA**, suggesting the effect of tricyclic system of **NBzA**. These observations for **NBzA** binding are important clues for improving the affinity and selectivity of ligands targeting the genetically important $r(CAG)_n$ repeat.

## Results

### Design and synthesis of NA derivatives

**NA** consisted of two units of aromatic heterocycles, which were connected by the linker through an amide linkage. This linker was adopted from the first-generation molecule of our mismatch-binding ligands naphthyridine dimer, which selectively bound to the G-G mismatch in double-stranded DNA.[34–36] The amide linkage to the 2-amino group of 1,8-naphthyridine was rather weak due to the electron-withdrawing property of 1,8-naphthyridine and, therefore, underwent unfavorable hydrolysis at a high temperature even in a neutral pH. To circumvent this potential instability, the amide linkage was replaced by a carbamate linkage for connecting the amino group of 1,8-naphthyridine. A new G-G mismatch binding molecule having a carbamate linkage showed an improved chemical stability and affinity,[29,37] which is likely due to the increased π-stacking at the carbamate moiety and the longer linker length by two atoms than the amide linkage. According to the successful transformation in G-G mismatch-binding molecules, the amide linkage connecting heterocycle units in **NA** was modified to the carbamate linkage.

For the systematic structure-binding studies, we separated **NA** into two parts in design, namely, amide-linked naphthyridine **N** and amide-linked azaquinolone **A** (Figure 2). The amide-linked 1,8-naphthyridine **N** was modified to the carbamate-linked 1,8-naphthyridine **NC**, which might have a better

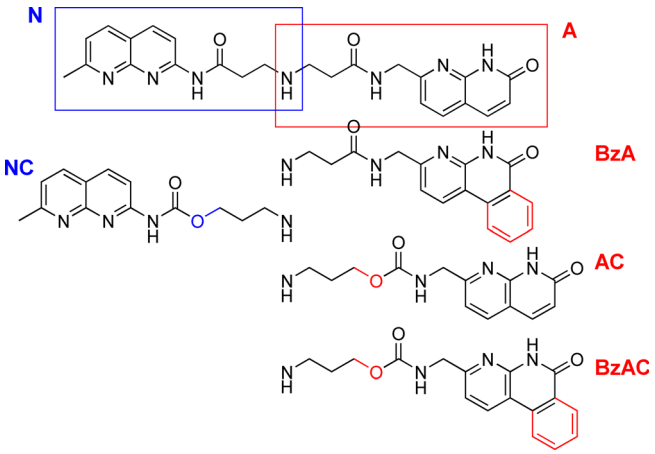



**Figure 2.** Structures of **NA** derivatives discussed in these studies. The combination of G-binding units (**N** and **NC**) shown in the left with A-binding units (**A**, **BzA**, **AC**, and **BzAC**) shown in the right provided six derivatives: **NA**, **NBzA**, **NAC**, **NBzAC**, **NCA**, and **NCAC**.

stacking ability with the neighboring bases and flexibility in the linker. With regard to the 8-azaquinolone part, the amide-linked 8-azaquinolone **A** was also changed to the carbamate-linked 8-azaquinolone **AC**. For these structural changes in the linker moiety, however, we could not keep the linker length same as that of parent **NA**, because of the instability of hemiaminal toward the acid hydrolysis and the facile cyclization into cyclic carbamate[37] (Figure 3).

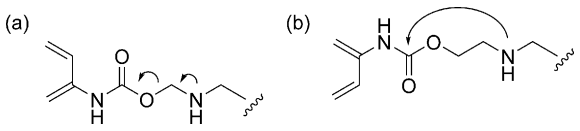



**Figure 3.** Limitation in the linker structures due to (a) instability of the hemiaminal structure and (b) facile intramolecular nucleophilic attack of nitrogen to a carbamate carbonyl group.

In addition, the 8-azaquinolone part was further modified by introducing a third ring system in benzoazaquinolone **BzA**. In the **NA**-CAG/CAG complex, NMR studies clearly revealed that 8-azaquinolone bound to the adenine through two hydrogen bonds mimicking the T-A base pairing.[27] However, the azaquinolone-A binding was likely an additional binding concomitant with the strong 1,8-naphthyridine-G binding based on the previous structure–activity studies on the **NA** binding to the G-A mismatch.[38] In fact, a naphthyridine dimer, where the 8-azaquinolone part of **NA** was substituted with 1,8-naphthyridine, still showed the affinity to the G-A mismatch, although the affinity was not as high as that of **NA**. To enhance the binding of the 8-azaquinolone part and, eventually, increase the affinity of the molecule to the CAG/CAG motif, we investigated the effect of the third ring system that may provide a better stacking possibility with the neighboring bases. The **BzA** moiety was also transformed into the carbamate-linked unit **BzAC**. With these basic designs in mind, combination of the 1,8-naphthyridine parts (**N** and **NC**) with the 8-azaquinolone parts (**A**, **BzA**, **AC**, and **BzAC**) yielded five new molecules, **NBzA**, **NAC**, **NBzAC**, **NCA**, and **NCAC**. The synthesis of these molecules is described in the Experimental Section. In brief, introduction of a carbamate linkage was accomplished by using *N*-Boc-((3-hydroxypropyl)amino)propanoic acid (for **NCA**, **NAC**, and **NBzAC**) and *N*-Boc-bis(3-hydroxypropyl)amine (for **NCAC** and **NCAC**). The benzoazaquinolone unit was synthesized from 2-amino-6-methylpyridine by way of 3-methylbenzo[*c*][1,8]naphthyridin-6(5*H*)-one and subsequent transformation of the methyl group.

### Binding analysis of NA derivatives to $(CAG)_n$ repeat DNA by SPR assay

At first, we investigated the binding of **NA** and five **NA** derivatives to $d(CAG)_9$ by the SPR assay using $d(CAG)_9$-immobilized surface prepared on a streptavidin-coated sensor (SA chip, GE Healthcare). In general, the analysis of SPR sensorgrams provides with kinetic parameters for the binding. This requires fitting the sensorgrams to the theoretical binding isotherm. When the host has a single binding site for the ligand, we could use the binding stoichiometry of 1:1. In the CXG repeat DNA and RNA, we speculated the potential binding site as the X-X mismatch flanked by C-G base pairs (CXG/CXG) when the CXG repeat folds into a hairpin secondary structure. The number of potential binding sites, however, could not be explicitly proven due to the possibility of forming diverse hairpin structures. Furthermore, the sensor surface with hairpin secondary structures of CXG repeat DNA and RNA may not reproducible upon denaturation and renaturation processes. With these ambiguities in the binding stoichiometry and reproducibility of the surface, the SPR sensorgrams obtained for ligand binding to CXG repeat sensor were discussed only in qualitative aspects.

SPR sensorgrams of ligand binding to $d(CAG)_9$-immobilized surface showed strong responses for **NA** and **NBzA** with the response unit reaching about 140 and 225 response units (RU), respectively, when both ligands were applied at 2 μM concentration. (Figure 4) In contrast, the other four ligands **NCA**, **NAC**, **NCAC**, and **NBzAC** showed a much decreased response below

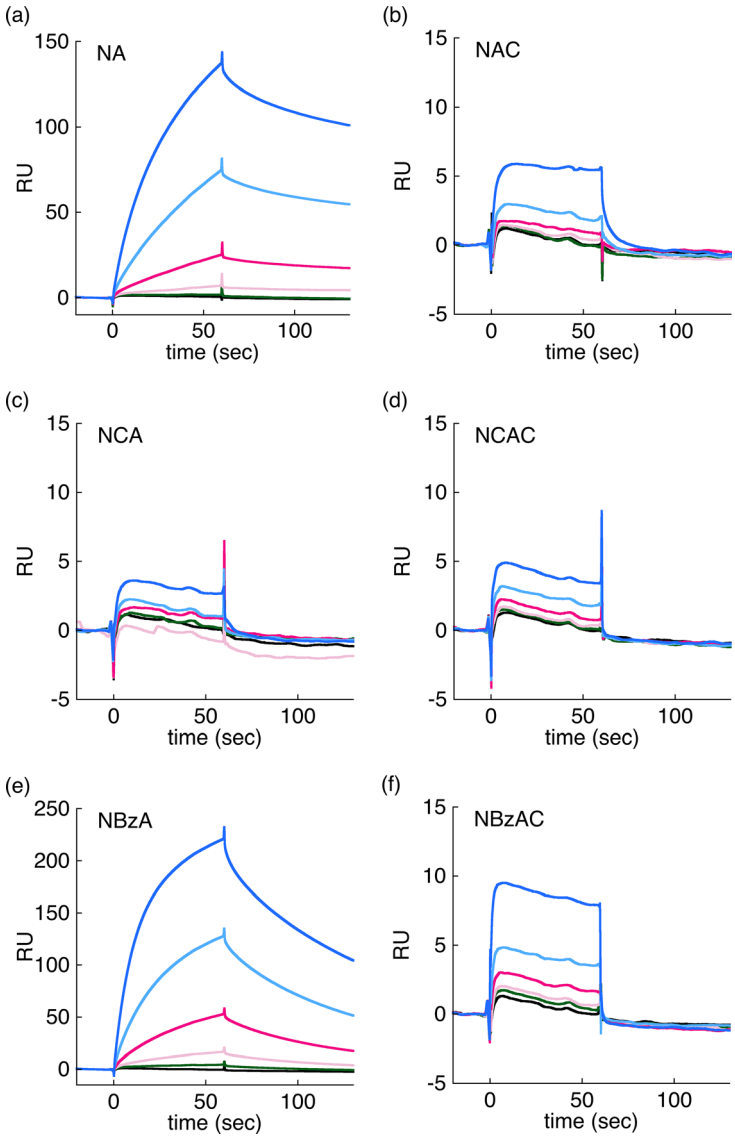



**Figure 4.** SPR analysis of ligand binding to the $d(CAG)_9$-immobilized surface. Each ligand was added stepwise at 0 (black), 0.13 (green), 0.25 (pink), 0.5 (red), 1.0 (pale blue), and 2.0 μM (blue). The sensors were exposed to the ligand for 0–60 s. $d(CAG)_9$ was immobilized on the SA sensor chip for 497 RU. (a) **NA**, (b) **NAC**, (c) **NCA**, (d) **NCAC**, (e) **NBzA**, and (f) **NBzAC**.

10 RU even when 2 μM ligand was applied. Furthermore, the shape of response curves of these weak binders showing fast association and dissociation was totally different from those observed for **NA** and **NBzA**. The large responses observed for **NA** and **NBzA** with slow association and dissociation showed the strong binding of these ligands to the $d(CAG)_9$ repeat. These SPR data for **NA** were fully consistent with those we reported before.[34] The similarity in the SPR response curves obtained for **NBzA** and **NA** suggested that **NBzA** binding to $d(CAG)_9$ shares the characteristics of mode of the **NA** binding to $d(CAG)_9$.

### Binding analysis of NA derivatives to $(CAG)_n$ repeat RNA by SPR assay

Then, we investigated the binding of **NA** and five **NA** derivatives to the $r(CAG)_9$ repeat by the SPR assay using $r(CAG)_9$-immobilized surface prepared on the SA chip (GE healthcare). SPR sensorgrams for the binding of six ligands to $r(CAG)_9$-immobilized surface showed a marked ligand dependence (Figure 5). Among six ligands, the strongest response of ca. 13 RU was observed for **NBzA** when 2 μM of the ligand was used. **NA** also showed a response to the $r(CAG)_9$ repeat with lower efficiency. **NA** and **NBzA** showed a concentration-dependent SPR response, showing that **NA** and **NBzA** bound to the $r(CAG)_9$. All four ligands **NCA**, **NAC**, **NCAC**, and **NBzAC** did not show significant binding in the SPR analyses to $r(CAG)_9$ at the concentration below 2 μM, while these ligands showed SPR responses when the concentration increased up to 50 μM of ligand (data not shown). These data indicated that four ligands should have a much lower affinity to $r(CAG)_9$ than that of **NA** and **NBzA**.

We also conducted single cycle SPR analysis for **NA** and **NBzA** binding to $d(CAG)_n$ and $r(CAG)_n$, confirming their binding (Figure 6 and 7). According to the shape of response curves, both association to and dissociation from $d(CAG)_9$ seems slower for **NA** than **NBzA**. Finally, we investigated if **NA** and **NBzA** might bind to other trinucleotide repeat sequences. Single cycle kinetic analysis firmly confirmed that **NA**- and **NBzA**-binding to $r(CCG)_9$, $r(CGG)_9$, and $r(CUG)_9$ were not significant up to 3 μM ligand concentration (Figures S1 and S2).

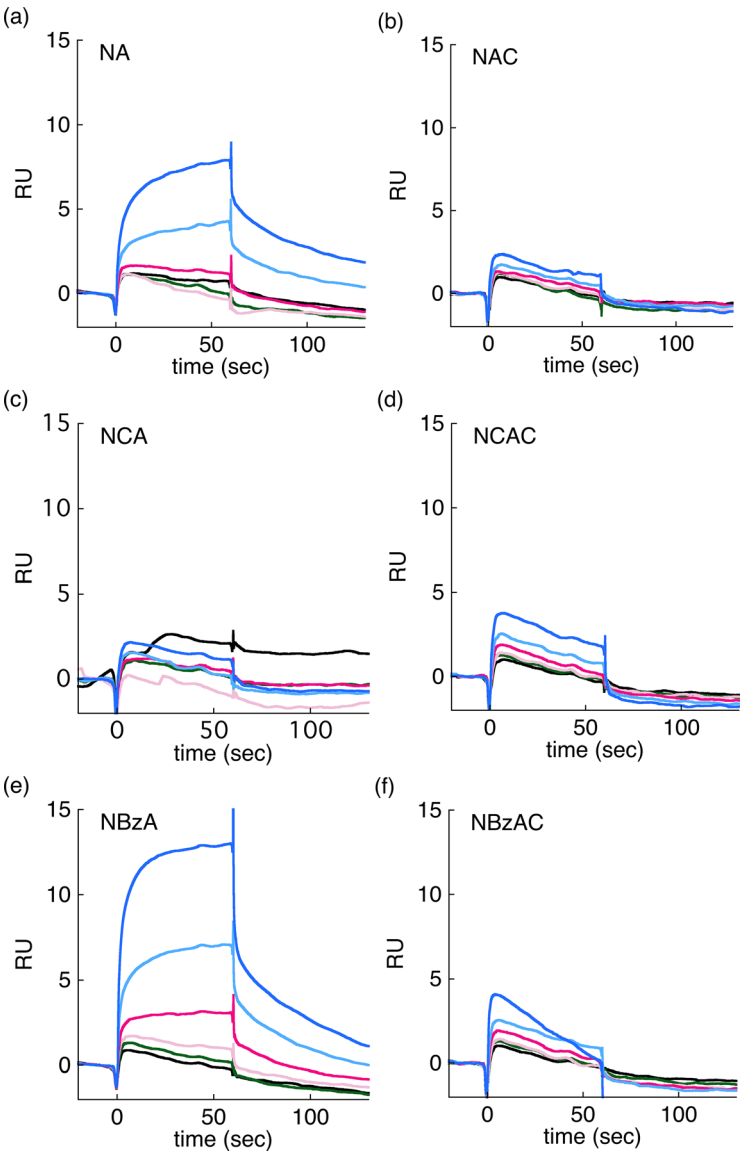



**Figure 5.** SPR analysis of ligand binding to the $r(CAG)_9$-immobilized surface. Each ligand was added stepwise at 0 (black), 0.13 (green), 0.25 (pink), 0.5 (red), 1.0 (pale blue), and 2.0 μM (blue). The sensors were exposed to the ligand for 0-60 s. $r(CAG)_9$ was immobilized on the SA sensor chip for 524 RU. (a) **NA**, (b) **NAC**, (c) **NCA**, (d) **NCAC**, (e) **NBzA**, and (f) **NBzAC**.

### CD spectra change on ligand binding

To gain information into the secondary structure, we measured the CD spectra of $d(CAG)_9$ and $r(CAG)_9$ in the presence and absence of **NA** and **NBzA**. As we reported, the CD spectra of $d(CAG)_9$ changed markedly upon binding with **NA**[27] (Figure 8a). Thus, $d(CAG)_9$ showed CD spectra similar to that of B-form dsDNA with a negative peak at 250 nm and a positive peak at 270 nm. Upon binding with **NA**, the negative peak at 250 nm became a positive peak and the peak at 270 nm became negative in addition to the appearance of a large negative induced CD at 320 nm. **NBzA**-biding to $d(CAG)_9$ also produced a negative CD signal at 270 nm and an induced signal at 320 nm. (Figure 8c) The positive peak observed at 250 nm for **NA**-binding to $d(CAG)_9$ shifted to 235 nm in the **NBzA**-d(CAG) complex probably because of a large absorption band of **NBzA** at 230 nm.

The CD spectra of $r(CAG)_9$ showed an A-form like structure in the absence of ligands (Figure 8b). A negative peak was observed at 240 nm and a positive peak was at 265 nm. Upon addition of **NA**, both the positive and negative peaks shifted toward a short wavelength by 5 to 10 nm with concomitant appearance of broad negative induced CD peak at 320 nm. All the CD peaks were rather broad. In contrast, the addition of

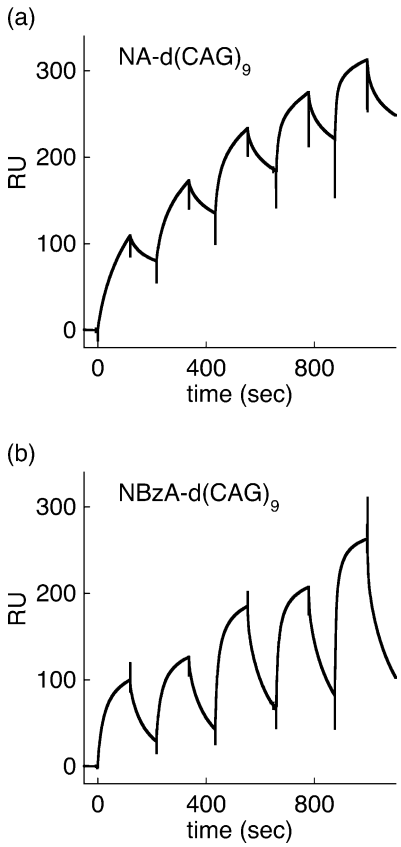



**Figure 6.** Single-cycle kinetic analysis of the ligands to $d(CAG)_9$ (497 RU) by SPR assay. The ligand (a) **NA** or (b) **NBzA** was sequentially added at 0.75, 1.0, 1.5, 2.0, and 3.0 μM.

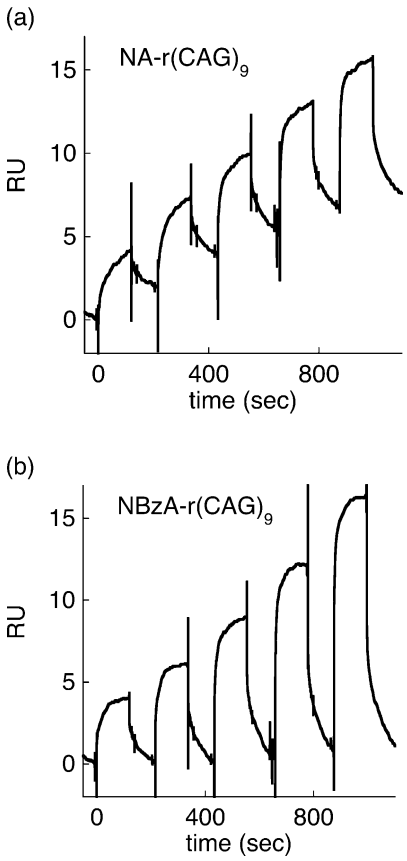



**Figure 7.** Single-cycle kinetic analysis of the ligands to $r(CAG)_9$ (524 RU) by SPR assay. The ligand (a) **NA** or (b) **NBzA** was sequentially added at 0.75, 1.0, 1.5, 2.0, and 3.0 μM.

**NBzA** increased the molar ellipticity of the positive peak at 280 nm with a shift toward a longer wavelength by 15 nm (Figure 8d). Since the absorption at 280 nm was very low for **NA** and **NBzA** (Figure S3), the strong positive CD peak at 280 nm in **NBzA**-bound $r(CAG)_9$ suggested the structural change on RNA. The induced CD signals at 320 nm also became much more apparent than that of **NA**-bound complex. In addition to the major induced CD peak, two weak but significant peaks at 345 nm (negative) and 355 nm (positive) were clearly seen for **NBzA**. By comparing the CD spectra of **NA**- and **NBzA**-bound $r(CAG)_9$ repeat, it is likely that the binding of **NBzA** induced more extensive conformational change on RNA.

### Stoichiometry of NA- and NBzA-binding to $d(CAG)_n$ and $r(CAG)_n$ repeats

On the basis of SPR and CD measurements, the binding of **NA** and **NBzA** to $r(CAG)_n$ repeat is suggested to be significantly different from that to $d(CAG)_n$ in terms of the mode of binding. To gain further insight into the mode of the binding of the two ligands to $d(CAG)_n$ and $r(CAG)_n$ repeat, ESI-TOF MS was carried out for these ligands and repeats. In our previous studies on **NA**, we have revealed that **NA** binds to $d(CAG)_n$ repeat in a pairwise manner, showing intense molecular ions containing an even number of **NA** molecules.[27] This pairwise binding of **NA** to $d(CAG)_n$ repeat was supported by the NMR structure obtained for the complex with a A-A mismatch flanking two C-G base pairs, a basic unit of CAG/CAG triad produced in the hairpin secondary structure of $d(CAG)_n$ repeat (cf. Figure 1).

ESI-TOF MS measurements of $d(CAG)_n$ repeat with **NA** successfully reproduced the molecular ions containing an even number of **NA** molecules as shown in Figure 9b. Molecular ions containing an odd number of **NA** molecules were almost negligible in terms of intensity. Similar mass spectra showing molecular ions containing an even number of **NBzA** molecules were obtained for **NBzA** (Figure 9c), suggesting that the mode of **NBzA** binding to $d(CAG)_n$ repeat is most likely involved in a pairwise binding as we observed for **NA**.

In marked contrast, the MS results for $r(CAG)_n$ repeat are totally different from the ESI-TOF MS of $d(CAG)_n$ repeat in terms of the number of ligand molecules included in the observed molecular ions. Thus, $r(CAG)_9$ showed 5- and 6-ions in the MS (Figure 10a). These ions decreased the intensity in the presence of **NA** with concomitant appearance of new ions corresponding to 5- and 6-ions of $r(CAG)_9 + 1$**NA** and $r(CAG)_9 + 2$**NA** complexes (Figure 10b). In the presence of **NBzA**, ions corresponding to complexes containing 1 to 3 **NBzA** were observed as well (Figure 10c). However, the intensity of ions corresponding to the complex containing 2 **NBzA** molecules was much stronger in the intensity than those containing 1 and 3 **NBzA** molecules. With an increase of **NBzA** concentration, ions of the

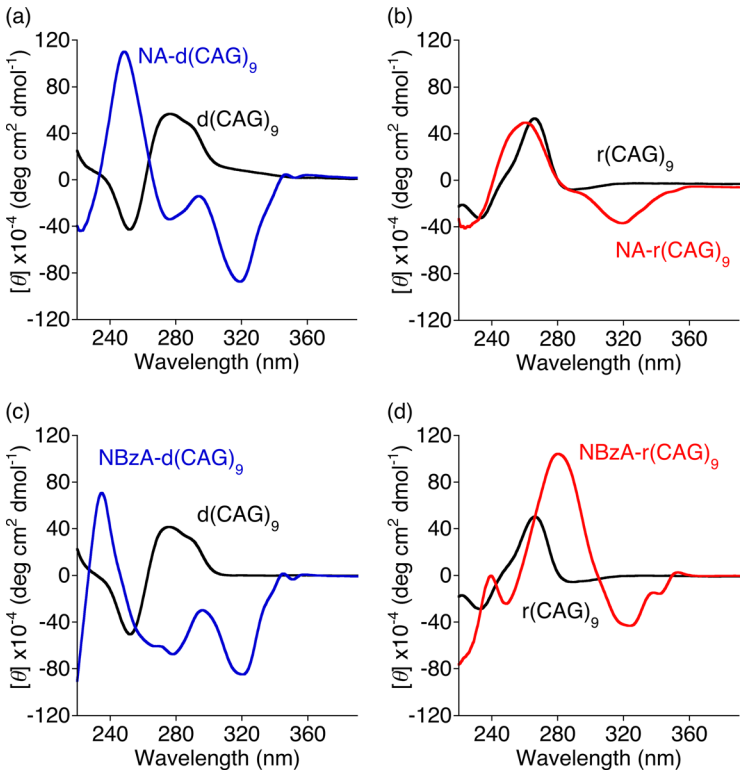



**Figure 8.** CD spectra of $d(CAG)_9$ and $r(CAG)_9$ (3 μM) in the presence of **NA** and **NBzA** (50 μM) in 10 mM sodium cacodylate buffer (pH 7.0) and 100 mM NaCl at room temperature. (a) **NA**-$d(CAG)_9$, (b) **NA**-$r(CAG)_9$, (c) **NBzA**-$d(CAG)_9$, and (d) **NBzA**-$r(CAG)_9$

complexes containing 4 and 5 **NBzA** were also observed (data not shown), suggesting that the preference of an even number of ligands in the binding to $d(CAG)_n$ repeat was not likely applicable to **NA**- and **NBzA**-binding to $r(CAG)_9$ repeat. On the basis of these observations, we concluded that the mode of **NA** and **NBzA** binding to $r(CAG)_9$ does not involve a pairwise binding.

### The effect of an additional ring system in NBzA in the binding to $r(CAG)_n$ repeat

For the binding to $r(CAG)_9$, **NBzA** induced more prominent structural changes upon binding as shown in CD spectra and the **NBzA**-bound complex showed better stability in ESI-TOF MS than the **NA**-bound complex. Single-cycle kinetic analysis by SPR suggested faster association and dissociation kinetics for **NBzA** than **NA**. These different binding characteristics of **NBzA** from those of **NA** are obviously due to the tricyclic benzoazaquinolone system implemented into the structure. The benzoazaquinolone in **NBzA** is anticipated to have a better stacking interaction with neighboring bases than the parent azaquinolone.

Ding and Ellestad studied the effect of the hydrophobicity of the molecule in the binding to DNA by changing the counter anion of sodium cation from a weakly hydrated $ClO_4^-$ anion to a strongly hydrated

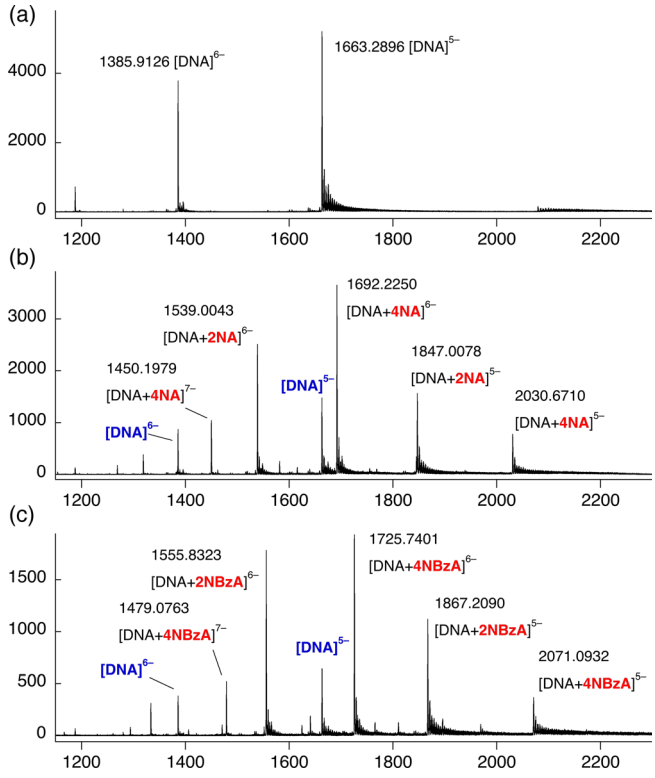



**Figure 9.** ESI-TOF-MS analysis of (a) $d(CAG)_9$ (2 μM), (b) $d(CAG)_9$ with **NA** (10 μM), and (c) $d(CAG)_9$ with **NBzA** (10 μM).

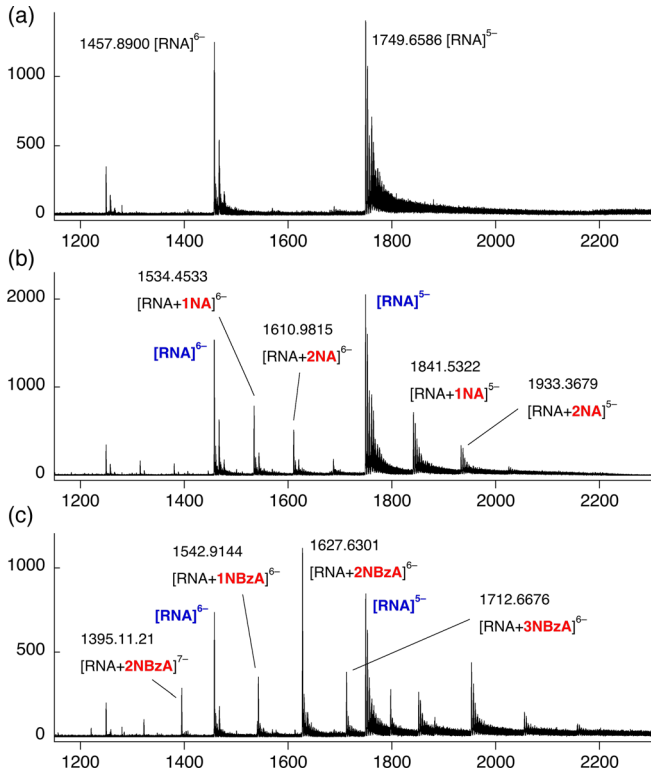



**Figure 10.** ESI-TOF-MS analysis of (a) $r(CAG)_9$ (2 μM), (b) $r(CAG)_9$ with **NA** (10 μM), and (c) $r(CAG)_9$ with **NBzA** (10 μM).

$SO_4^{2-}$ anion.[39] The $SO_4^{2-}$ anion showed "salting out" effect by increasing the hydrophobic interaction, whereas the $ClO_4^-$ anion showed the opposite effect.[40] The $Cl^-$ anion had intermediate effect between these two anions. The CD spectra of $r(CAG)_9$ (3 μM) in the presence of **NA** and **NBzA** (50 μM) were measured in the presence of 100 mM of $NaClO_4$, NaCl, and $Na_2SO_4$. While the CD spectra of $r(CAG)_9$ was not affected by the anion at all (Figure S4), significant differences were observed in the presence of **NA** (Figure 11 a). The CD bands increased the intensity in the presence of $SO_4^{2-}$ anion compared with those in $Cl^-$ and $ClO_4^-$ anions. The intensity changes were especially remarkable in the induced CD bands suggesting that the **NA**-binding to $r(CAG)_9$ was most favorable in the medium of increased hydrophobic interaction with 100 mM $Na_2SO_4$ among three conditions. In contrast, the CD spectra of $r(CAG)_9$ in the presence of **NBzA** was much less sensitive to the change from $Cl^-$ to $SO_4^{2-}$, and virtually no change from $Cl^-$ to $ClO_4^-$ (Figure 11 b). We also looked at the binding of **NA** and **NBzA** by SPR with the buffer containing $SO_4^{2-}$ anion to see if there were any effects of counter anion on the binding, but the SPR signals obtained for $r(CAG)_9$-immobilized sensor with **NA** and **NBzA** up to 3 μM (cf. SPR conditions in Figure 7) were so weak that we could not discuss the effect of counter anions.

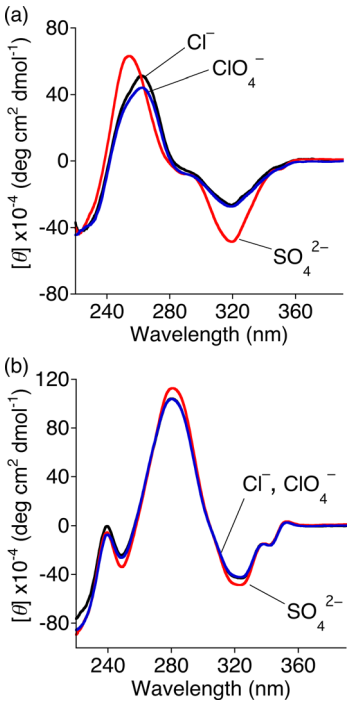



**Figure 11.** CD spectra of $r(CAG)_9$ (3 μM) with (a) **NA** and (b) **NBzA** (50 μM) in the presence of 100 mM NaCl (black line), $Na_2SO_4$ (red line) or $NaClO_4$ (blue line) in 10 mM sodium cacodylate buffer (pH 7.0) at room temperature.

### In vitro translation assay

The effect of **NA** and **NBzA** on in vitro translation of genes upstream or downstream from $(CAG)_n$ repeat RNA was investigated by a reporter system involving *Renilla* luciferase (*R*luc) and firefly luciferase (Fluc)[41,42] (Figure 12). We have constructed a vector containing $d(CAG)_{89}$ repeat between *R*luc and Fluc genes and a control vector that did not contain the repeat (Figure 12 A and 12 B). After in vitro transcription, the RNA was in vitro translated with Rabbit Reticulocyte Lysate System in the absence (w/o) or in the presence of **NA** or **NBzA** in different concentrations. The activities of *R*luc and Fluc were measured by the luciferase assay kit and compared with those obtained by in vitro translation of control vector. For the translation of luciferases on the control RNA (Figure 12 B), both **NA** and **NBzA** showed a tendency of suppressing the translation of both luciferases as the concentration of ligands increased. The suppression rate of the control RNA was 0.53 and 0.41 for *R*luc and Fluc, respectively, (Figure 12 B-a and B-c) at 100 μM of **NA**, giving the ratio of Fluc and *R*luc activities of 0.78 (Figure 12 B-e). **NBzA** also showed the similar effects on translation of luciferases on control RNA, providing 0.41 and 0.30 for *R*luc and Fluc, respectively, (Figure 12 B-b and B-d) at 100 μM. The ratio of two luciferase activities was 0.71. Comparing **NA** and **NBzA**, the translation suppression was more significant for **NBzA** (cf. B-a and B-b, and B-c and B-d).

For the translation of luciferases containing $(CAG)_{89}$ repeat between *R*luc and Fluc (Figure 12 A), translation suppression of Fluc by **NA** and **NBzA** was somewhat less efficient as compared to those observed for the translation of Fluc on the control RNA. Thus, the activity of Fluc translated from repeat containing RNA was 0.50 and 0.49 with **NA** and **NBzA**, respectively, (Figure 12 A-c and A-d) at 100 μM concentration and that obtained for control RNA was 0.41 and 0.30 (Figure 12, B-c and B-d) This tendency was more marked for the translation suppression of *R*luc located upstream of $(CAG)_{89}$. The activity of *R*luc was 0.75 and 0.71 (Figure 12 A-a and A-b, respectively) at the same ligand concentration, which were significantly higher than those observed for the translation of *R*luc on control RNA (cf. 0.53 with **NA** (B-a) and 0.41 with **NBzA** (B-b)). While the molecular mechanism was not clear, the presence of $(CAG)_{89}$ repeat somehow repressed the translation suppression by both **NA** and **NBzA**. The differences between two ligands were not obvious in the translation of repeat containing RNA.

## Discussion

In our previous studies on the molecules binding to the G-G mismatched DNA, the modification of the amide linkage to the carbamate linkage significantly increased the binding affinity.[29,37] While we anticipated the increased affinity by additional stacking interaction and flexibility by a longer linker length, it was not the case for the four molecules **NCA**, **NAC**, **NCAC**, and **NBzAC** having a carbamate linkage between two heterocycles. SPR analysis of **NA** and five newly synthesized ligands clearly indicated that only **NA** and **NBzA** did show the affinity to $d(CAG)_9$ and $r(CAG)_9$. Failure of the binding of four ligands with the carbamate linkage is possibly due to the difference in conformations of two heterocycles from those of **NA** and **NBzA**, which altered the required energy for the rearrangement of two chromophores into the appropriate position for the bind-

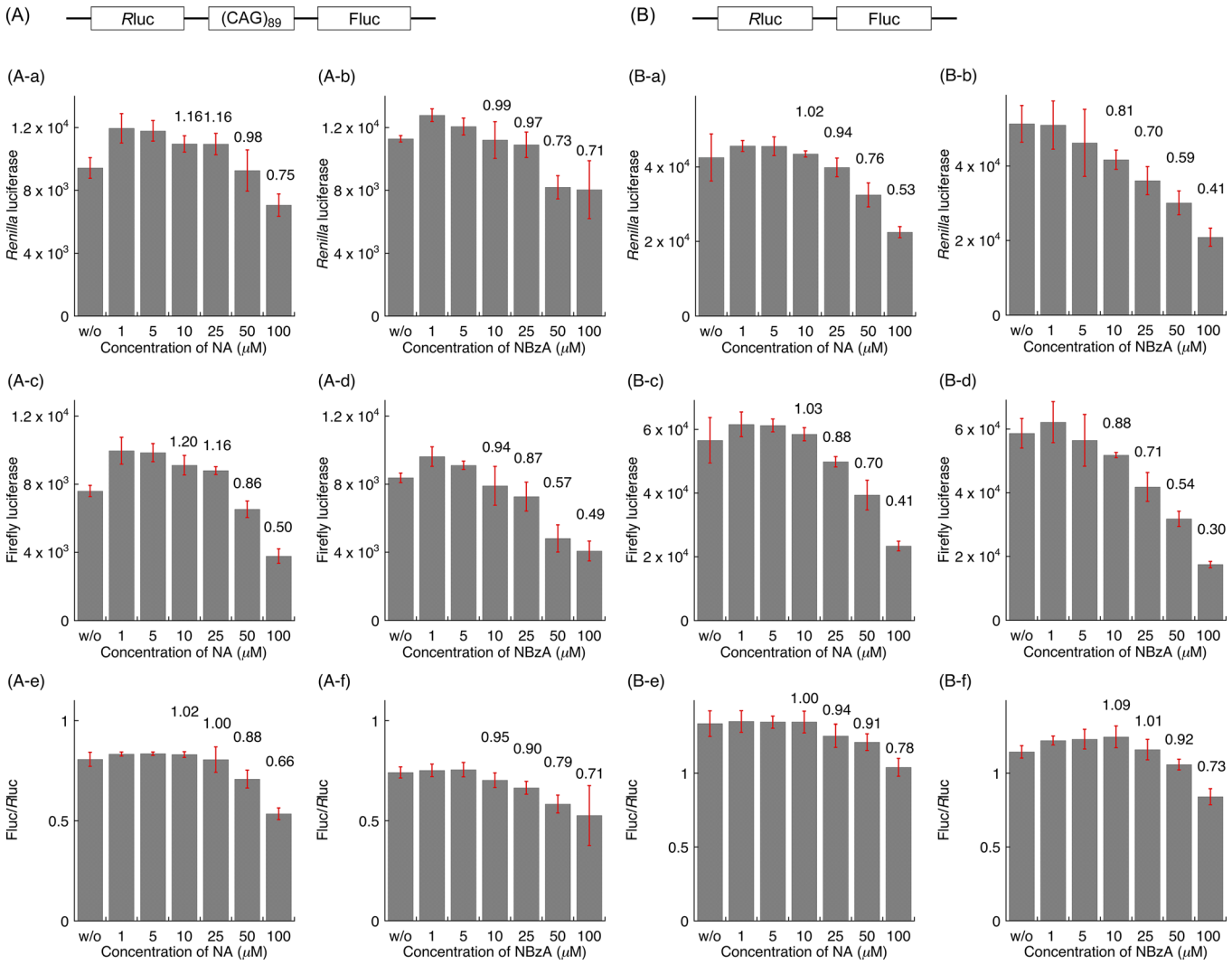



**Figure 12.** The effect of **NA** and **NBzA** on the in vitro translation of RNA template containing *R*luc and Fluc genes (A) with or (B) without the intervening sequence of $(CAG)_{89}$ repeat. The activities of reporter proteins were examined in the presence of **NA** and **NBzA**, and plotted against ligand concentration. Key: *R*luc activities with (a) **NA** and (b) **NBzA**, Fluc activities with (c) **NA** and (d) **NBzA**, the ratio of two luciferases (Fluc/*R*luc) with (e) **NA** and (f) **NBzA**. The number shown above the bar indicated the ratio of luciferase activity (a–d) and Fluc/*R*luc (e and f) relative to those without ligand.

ing. Since **NA** and **NBzA** have the same linker structure with the shortest length among molecules we examined, one of factors determining the binding to CAG repeat DNA and RNA would be likely the spatial distance and arrangement of two chromophores provided by the linker.

The SPR sensorgrams obtained for $d(CAG)_9$ and $r(CAG)_9$ with **NA** and **NBzA** were quite different in terms of the magnitude of the response. The SPR response for $d(CAG)_9$ reaches about 225 RU for the binding of **NBzA**, whereas that for $r(CAG)_9$ is only about 13 RU. The amount of immobilized $d(CAG)_9$ and $r(CAG)_9$ was 497 and 524 RU, respectively, suggesting that the large difference is not likely due to the amount of immobilized polynucleotides. Other possibilities are the difference in the folded structure and the amount of bound ligands.

ESI-TOF-MS analysis clearly showed that the binding stoichiometry of **NA** and **NBzA** to $d(CAG)_9$ and $r(CAG)_9$ was quite different. A pairwise binding of **NA** and **NBzA** was confirmed for the binding to $d(CAG)_9$, whereas a stepwise binding of ligands to $r(CAG)_9$ without any apparent cooperativity was suggested. In our previous studies, the large SPR response was rationalized by the formation of hairpin secondary structure on the sensor surface inducing a large change in dielectric constant near the surface.[27] The large change in SPR response for $d(CAG)_9$ might be due to the induced formation of hairpin secondary structure on the surface upon ligand binding. Since $r(CAG)_9$ is reported to favorably take a hairpin secondary structure,[19] such ligand-induced hairpin formation may not be conceivable for $r(CAG)_9$.

In addition to the stoichiometry of the binding, the MS analysis may suggest that **NBzA**-bound complexes to $r(CAG)_9$ are thermodynamically more stable than **NA**-bound complexes if we could assume a comparable ionization efficiency for both **NA**- and **NBzA**-bound complexes, because the intensity of ions corresponding to complex containing 2 and 3 **NBzA** molecules is higher than those containing **NA**.

The results obtained from CD measurements in the different counter anions showed that the increased hydrophobic interaction by a salting out $SO_4^{2-}$ anion promotes the **NA**-binding to r(CAG)$_9$ but not the **NBzA**-binding significantly. This observation suggests that **NBzA** intrinsically gained such hydrophobic effect but **NA**-binding was further promoted by the salting out effect.[39,40] It is reasonable to assume that the expanded tricyclic system has more chances to gain hydrophobic interaction than the parent two-ring system, the characteristic **NBzA**-binding to r(CAG)$_9$ is most likely due to the stacking interaction of the benzoazaquinolone system with the neighboring bases.

To investigate the effect of **NA** and **NBzA** binding to (CAG)$_n$ repeat RNA, we have established the in vitro translation assay, where two reporter genes, for example, *Renilla* luciferase (*R*luc) and firefly luciferase (Fluc), were arranged upstream and downstream from (CAG)$_n$ repeat. Simultaneous suppression of both *R*luc and Fluc expression by ligand suggests that ligand interfered the entire translation system, whereas pronounced suppression of expression of one of two genes could be due to the presence of (CAG)$_n$ repeat.

For the control RNA that lacked the intervening CAG repeat sequence between two reporter genes, both **NA** and **NBzA** suppressed the translation of both *R*luc and Fluc in concentration-dependent manner, suggesting that both ligands have inhibitory effects on the translation. The ratio of two reporter genes was 0.78 for **NA** and 0.73 for **NBzA** at 100 μM concentration. These ratios are the intrinsic number for the translation of tandemly arranged *R*luc and Fluc under the conditions of our in vitro translation assay. For the translation of RNA containing CAG intervening repeat sequence between two luciferases, the effect of ligands on the translation suppression was less efficient. Especially, the translation of *R*luc located upstream of CAG repeat was kept in high level of translation than that in the RNA without CAG repeat (e.g., 0.75 vs. 0.53 for **NA**, and 0.71 vs. 0.41 for **NBzA** treatment at 100 μM). The repressive effect on translation suppression of *R*luc gene is due to the presence of CAG repeat, though the molecular mechanism is not fully understood. Differences in two ligands **NA** and **NBzA** for the translation were not clearly observed. It was apparent that the tricyclic system in **NBzA** did not have any significant positive or negative effects on translation, but provides large chemical space for further modification to gain higher affinity and selectivity.

## Conclusions

Among the five new **NA** derivatives we synthesized, those modified by the carbamate linkage failed to bind to r(CAG)$_9$. The **NBzA** having a tricyclic system of benzoazaquinolone showed comparable binding affinity and repeat selectivity with **NA**, and induced a more marked and intense CD signal change than **NA**. ESI-TOF MS analysis revealed that the mode of binding of **NA** and **NBzA** to an r(CAG)$_n$ repeat is not the pairwise binding we observed for the binding to d(CAG)$_n$ repeat. The intensity of ion peaks was higher for the **NBzA**-bound r(CAG)$_9$ than that of **NA**-bound r(CAG)$_9$, suggesting that **NBzA**-bound r(CAG)$_9$ would be more stable than **NA**-bound complexes. These characteristics of **NBzA**-binding are due to the presence of the tricyclic system, by which the stacking interaction with the neighboring base pairs would likely be promoted. All the data described here suggested that the linker structure of **NA** and **NBzA** was suitable for the ligand binding to CAG repeat DNA and RNA based on our molecular design involving two heterocycles, and the tricyclic system of benzoazaquinolone does not interfere with binding to the CAG repeat. While the binding of **NBzA** was not markedly increased in terms of the affinity, the tricyclic ring system should provide us with a larger chemical space than the bicyclic system to design compounds, which eventually may exhibit exclusive binding to CAG repeat RNA with high affinity.

## Experimental Section

**Surface plasmon resonance assay for the r(CAG)$_9$-immobilized surface**. A streptavidin-coated sensor chip (SA chip, GE Healthcare) was washed with HBS-EP buffer (10 mM HEPES pH 7.4, 0.15 M NaCl, 3 mM EDTA and 0.005% v/v Surfactant P20) for 6 min and then activated with three consecutive 1 min injections of 30 μL activation buffer (50 mM NaOH and 1 M NaCl). 5′-Biotinylated r(CAG)$_9$ (purchased from Gene Design Inc., Osaka Japan) was diluted to 1 μM with HBS-EP buffer and flowed onto the SA chip at 5 μL min$^{-1}$ for 60 s. The amount of r(CAG)$_9$ immobilized on the chip surface was 524 response units (RU). Binding of the ligands to the surface was analyzed by using a BIAcore T200 SPR system (GE Healthcare).

Multi-cycle kinetic analysis was carried out at 25 °C under the continuous flow of HBS-EP buffer at a flow rate of 30 μL min$^{-1}$. After the surface had been conditioned by exposure to the buffer flow for 60 s, a solution of each concentration of the ligand in HBS-EP buffer was injected for 60 s to analyze association to the sensor chip. The buffer was subsequently injected for another 180 s in order to determine the dissociation of the bound ligand from the surface.

Single-cycle kinetic analysis was carried out at 25 °C under the continuous flow of HBS-EP buffer. To avoid mass transport limitations, the flow rate of HBS-EP buffer was 90 μL min$^{-1}$.[43] The surface was conditioned by exposure for 120 s to analyze the association of ligands to the sensor chip. Each ligand was dissolved in HBS-EP buffer at the concentration of 0.75, 1.0, 2.0, 2.5, and 3.0 μM, and the resulting solutions were sequentially injected over flow cells on the sensor surface for 120 s at a flow rate of 90 μL min$^{-1}$ in single-cycle mode. The obtained data were analyzed using the BIAcore T200 evaluation software, version 2.0, and kinetic parameters were determined by fitting the data to a 1:1 binding model.

**Circular dichromic spectra of r(CAG)$_9$ in the presence of NA and NBzA.** Circular dichroism (CD) measurements were carried out on a J-725 CD spectropolarimeter (JASCO, Japan) using a 1.0 cm path length cell. CD spectra of r(CAG)$_9$ (3 μM) in sodium cacodylate buffer (10 mM, pH 7.0) and NaCl (100 mM) were measured in the absence and presence of **NA** or **NBzA** (50 μM) at 25 °C.

**ESI-TOF MS measurements.**[44] ESI-TOF MS experiments were performed on a Bruker maXis impact mass spectrometer in the negative mode. The source parameters were set as follows: capillary voltage of +3500 V, nebulizer of 1.0 bar, dry gas flow of 4.0 L min$^{-1}$ at 160 °C, and end plate offset voltage of 500 V. The measurements were operated at a background pressure of $5\times10^{-7}$ mbar. The rate of sample injection into the mass spectrometer was 3 μL min$^{-1}$. The mixture of a final concentration of 2 μM r(CAG)$_9$ repeat with 10 μM **NA** or 10 μM **NBzA** in 50 mM ammonium acetate solution

(pH 7.0) was directly injected into mass spectrometer. Each mass spectrum was an average of at least 60 scans. Data were analyzed with the software of Compass DataAnalysis.

**Preparation of RNA template.** The pCMVTnT_RFluc vector carrying *Renilla* and firefly luciferase reporter gene in tandem downstream of the T7 promoter sequence was sequentially digested with *Sal*I and *Eco*RI (Takara) and purified by using an NucleoSpin Gel and PCR Clean-up kit (Macherey–Nagel). A dsDNA fragment containing an *Eco*NI site was prepared by annealing the two oligonucleotides 5′-AAT TCT CCT GGG GGA GGT G-3′ and 5′-TCG ACA CCT CCC CCA GGA G-3′ (purchased from Gene Design), producing sticky ends for *Sal*I and *Eco*RI. The dsDNA fragment was ligated into the *Sal*I/*Eco*RI-digested vector with T4 DNA ligase (New England BioLabs) to create an *Eco*NI site between *Renilla* and firefly luciferase reporter gene in the pCMVTnT_RFluc vector. The obtained vector was digested with *Eco*NI (New England BioLabs) and dephosphorylated with Antarctic Phosphatase (New England BioLabs), followed by purification with NucleoSpin Gel and PCR Clean-up kit (Macherey–Nagel). DNA fragments containing the $(CAG)_n$ sequence were prepared by ligation of double-stranded DNA containing 5′-G $(CAG)_9$ CA-3′ and 5′-$(CTG)_{10}$-3′ (purchased from Gene Design) with T4 DNA ligase (New England BioLabs). The ligation products were separated by native PAGE followed by staining with SYBR®Gold, and the bands ranging from 150 to 300 bp were excised and eluted from the gel. The obtained DNA fragments were ligated into the *Eco*NI-digested vector and the product was then used to transform NEB Stable Competent *E. coli* (New England Biolabs). The individual clones were partially sequenced and the number of CAG repeats was determined. Among the clones isolated, the clone carrying $(CAG)_{89}$ sequence (pCMVTnT_RFluc_$CAG_{89}$) was used for preparing the RNA template for in vitro translation. pCMVTnT_RFluc_$CAG_{89}$ was linearized by digestion with *Bam*HI (Takara) prior to in vitro transcription. The linearized pCMVTnT_RFluc_$CAG_{89}$ was transcribed with T7 RNA polymerase (HiScribe T7 Quick High Yield RNA Synthesis Kit, New England BioLabs) to produce RNA containing the $CAG_{89}$ repeat. After in vitro transcription, the reaction mixture was treated with RNase-free DNase I (QIAGEN) to digest the DNA template, and precipitated with ammonium acetate/2-propanol. The precipitates were dissolved in RNase-free water and the mixture was applied to a NAP-5 gel filtration column (GE Healthcare) to remove unincorporated NTPs. The purified RNA was dissolved in water, analyzed by 1.5% (w/v) agarose gel electrophoresis, and stored at −80 °C.

**In vitro translation assay.** In vitro translation was performed using Rabbit Reticulocyte Lysate System (Promega). Translation reactions (10 μL) contained 70% lysate, 20 μM amino acid mixture, and 85 fmol template RNA in the presence of the ligand at the final concentration of 0, 1, 5, 10, 25, 50, and 100 μM. The reaction mixtures were incubated at 30 °C for 60 min and were quenched by adding RNase A (10 μL) to a final concentration of 0.2 mg $mL^{-1}$. Firefly and *Renilla* luciferase activities were measured with a Mithras LB940 multimode microplate reader using the Dual-Glo Luciferase Assay System (Promega).

## Acknowledgements

This work was supported by JSPS KAKENHI Grant-in-Aid for Specially Promoted Research (26000007) and partially by the Advanced Research for Medical Products Mining Program of the National Institute of Biomedical Innovation of the National Institute of Biomedical Innovation (NIBIO, 10–22) for KN. Authors also thank Platform for Drug Discovery, Informatics, and Structural Life Science from the Ministry of Education, Culture, Sports, Science and Technology, Japan for the technical support on using Molecular modelling software (SCHRÖDINGER Suit).

**Keywords:** CAG repeats · recognition · RNA · small molecules · surface plasmon resonance · trinucleotide repeats

[1] G. R. Wilmot, S. T. Warren in *Genetic Instabilities and hereditary neurological diseases*, (Eds, : R. D. Wells, S. T. Warren), Academic Press, San Diego, **1998**, pp. 3–12.
[2] C. J. Cummings, H. Y. Zoghbi, *Annu. Rev. Genomics Hum. Genet.* **2000**, *1*, 281–328.
[3] S. M. Mirkin, *Nature* **2007**, *447*, 932–940.
[4] H. T. Orr, H. Y. Zoghbi, *Annu. Rev. Neurosci.* **2007**, *30*, 575–621.
[5] S. W. Davies, M. Turmaine, B. A. Cozens, M. DiFiglia, A. H. Sharp, C. A. Ross, E. Scherzinger, E. F. Wanker, L. Mangiarini, G. P. Bates, *Cell* **1997**, *90*, 537–548.
[6] F. O. Walker, *Lancet* **2007**, *369*, 218–228.
[7] J. D. Brook, M. E. McCurrach, H. G. Harley, A. J. Buckler, D. Church, H. Aburatani, K. Hunter, V. P. Stanton, J. P. Thirion, T. Hudson, R. Sohn, B, Zemelman, R. G. Snell, S. A. Rundle, S. Crow, J. Davies, P. Shelbourne, J. Buxton, C. Jones, V. Juvonen, K. Johnson, P. S. Harper, D. J. Shaw, D. E. Housman, *Cell* **1992**, *68*, 799–808.
[8] H. Jiang, A. Mankodi, M. S. Swanson, R. T. Moxley, C. A. Thornton, *Hum. Mol. Genet.* **2004**, *13*, 3079–3088.
[9] M. B. de León, B. Cisneros, *J. Neurosci. Res.* **2008**, *86*, 18–26.
[10] Y. H. Fu, D. P. Kuhl, A. Pizzuti, M. Pieretti, J. S. Sutcliffe, S. Richards, A. J. Verkerk, J. J. Holden, R. G. Fenwick, S. T. Warren, B. A. Oostra, D. L. Nelson, C. T. Caskey, *Cell* **1991**, *67*, 1047–1058.
[11] R. Willemsen, J. Levenga, B. A. Oostra, *Clin. Genet.* **2011**, *80*, 214–225.
[12] S. I. Bidichandani, T. Ashizawa, P. I. Patel, *Am. J. Human Genet.* **1998**, *62*, 111–121.
[13] M. B. Delatycki, R. Williamson, S. M. Forrest, *J. Med. Genet.* **2000**, *37*, 1–8.
[14] A. R. La Spada, J. P. Taylor, *Nat. Rev. Genet.* **2010**, *11*, 247–258.
[15] C. T. McMurray, *Nat. Rev. Genet.* **2010**, *11*, 786–799.
[16] V. Dion, J. H. Wilson, *Trends Genet.* **2009**, *25*, 288–297.
[17] J. D. Cleary, K. Nichol, Y. H. Wang, C. E. Pearson, *Nat. Genet.* **2002**, *31*, 37–46.
[18] K. Ohshima, R. D. Wells, *J. Biol. Chem.* **1997**, *272*, 16798–16806.
[19] C. T. McMurray, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 1823–1825.
[20] K. Sobczak, M. de Mezer, G. Michlewski, J. Krol, W. J. Krzyzosiak, *Nucleic Acids Res.* **2003**, *31*, 5469–5482.
[21] C. E. Pearson, R. R. Sinden, *Biochemistry* **1996**, *35*, 5041–5053.
[22] S. Kang, A. Jaworski, K. Ohshima, R. D. Wells, *Nat. Genet.* **1995**, *10*, 213–218.
[23] M. J. Hartenstine, M. F. Goodman, J. Petruska, *J. Biol. Chem.* **2000**, *275*, 18382–18390.
[24] A. Marquis Gacy, G. Goellner, N. Juranić, S. Macura, C. T. McMurray, *Cell* **1995**, *81*, 533–540.
[25] C. E. Pearson, R. R. Sinden, *Curr. Opin. Struct. Biol.* **1998**, *8*, 321–330.
[26] C. E. Pearson, M. Tam, Y. H. Wang, S. E. Montgomery, A. C. Dar, J. D. Cleary, K. Nichol, *Nucleic Acids Res.* **2002**, *30*, 4534–4547.
[27] K. Nakatani, S. Hagihara, Y. Goto, A. Kobori, M. Hagihara, G. Hayashi, M. Kyo, M. Nomura, M. Mishima, C. Kojima, *Nat. Chem. Biol.* **2005**, *1*, 39–43.
[28] M. Hagihara, H. He, K. Nakatani, *ChemBioChem* **2011**, *12*, 1686–1689.
[29] T. Peng, K. Nakatani, *Angew. Chem. Int. Ed.* **2005**, *44*, 7280–7283; *Angew. Chem.* **2005**, *117*, 7446–7449.
[30] M. Hagihara, H. He, M. Kimura, K. Nakatani, *Bioorg. Med. Chem. Lett.* **2012**, *22*, 2000–2003.
[31] T. Zu, B. Gibbens, N. S. Doty, M. Gomes-Pereira, A. Huguet, M. D. Stone, J. Margolis, M. Peterson, T. W. Markowski, M. A. Ingram, Z. Nan, C. Forster, W. C. Low, B. Schoser, N. V. Somia, H. B. Clark, S. Schmechel, P. B. Bitterman, G. Gourdon, M. S. Swanson, M. Moseley, L. W. Ranum, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 260–265.

[32] C. E. Pearson, *Plos Genet.* **2011**, *7*, e1002018.
[33] M. Wojciechowska, M. Olejniczak, P. Galka-Marciniak, M. Jazurek, W. J. Krzyzosiak, *Nucleic Acids Res.* **2014**, *42*, 11849 – 11864.
[34] K. Nakatani, S. Sando, I. Saito, *Nat. Biotechnol.* **2001**, *19*, 51 – 55.
[35] K. Nakatani, S. Sando, H. Kumasawa, J. Kikuchi, I. Saito, *J. Am. Chem. Soc.* **2001**, *123*, 12650 – 12657.
[36] K. Nakatani, *Bull. Chem. Soc. Jpn.* **2009**, *82*, 1055 – 1069.
[37] T. Peng, T. Murase, Y. Goto, A. Kobori, K. Nakatani, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 259 – 262.
[38] S. Hagihara, H. Kumasawa, Y. Goto, G. Hayashi, A. Kobori, I. Saito, K. Nakatani, *Nucleic Acids Res.* **2004**, *32*, 278 – 286.
[39] W. D. Ding, G. A. Ellestad, *J. Am. Chem. Soc.* **1991**, *113*, 6617 – 6620.
[40] R. Breslow, C. J. Rizzo, *J. Am. Chem. Soc.* **1991**, *113*, 4340 – 4341.
[41] B. A. Sherf, S. L. Navarro, R. R. Hannah, K. V. Wood, *Promega Notes* **1996**, *57*, 2 – 9.
[42] A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11593 – 11598.
[43] R. Karlsson, A. Fält, *J. Immunol. Methods* **1997**, *200*, 121 – 133.
[44] L.-P. Bai, M. Hagihara, K. Nakatani, Z.-H. Jiang, *Sci. Rep.* **2014**, *4*, 6767.

Manuscript received: April 16, 2016
Accepted Article published: May 16, 2016
Final Article published: ■■ ▮▮, 0000

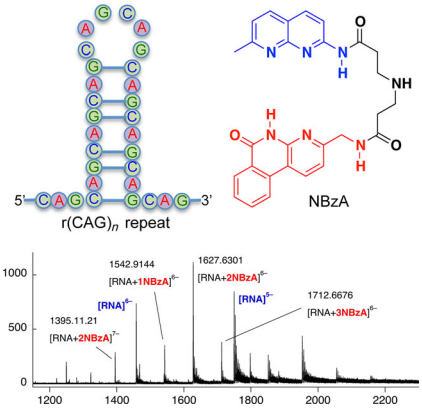
# FULL PAPER

**Binding to Trinucleotide Repeats**

*Jinxing Li , Akihiro Sakata, Hanping He, Li-Ping Bai, Asako Murata, Chikara Dohno, Kazuhiko Nakatani**

■■ – ■■

**Naphthyridine-Benzoazaquinolone: Evaluation of a Tricyclic System for the Binding to $(CAG)_n$ Repeat DNA and RNA**




**Repeat RNA Binding**: Structure-binding studies of a series of naphthyridine-azaquinolone derivatives are described. **NBzA** with the tricyclic system of benzoazaquinolone induced a more extensive conformational change on the CAG repeat. The linker structure in **NA** and **NBzA** was suitable for the binding to CAG DNA and RNA, and the tricyclic benzoazaquinolone did not interfere with the binding.