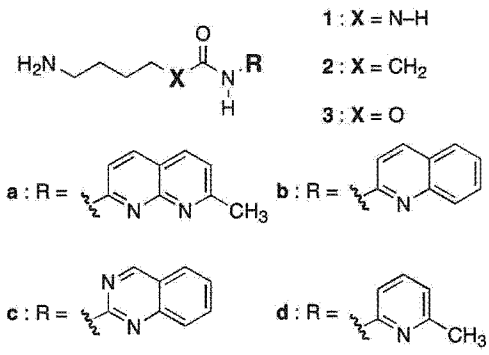


Figure 2. Molecules examined in these studies.

of a single nucleotide bulge in duplex DNA was compared with corresponding amide compounds **2a** and **2b** as well as quinolinylcarbamate **3b**. Amide compounds were synthesized by the coupling of amino heterocycles with pentafluorophenyl *N*-Boc-6-aminohexanoate, whereas **3b** was obtained by Curtius rearrangement of quinaldic acid followed by a coupling with *N*-Boc-aminobutanol. Deprotection of the Boc group furnished the synthesis of all molecules. The number of atoms between the carbonyl carbon to the terminal amino group in the linker was kept constant for all molecules. The effects of these molecules on a stabilization of a single nucleotide bulge were investigated by measuring a melting temperature (T_m) of duplexes 5′-d(TCC AG GCA AC)-3′/3′-d(AGG TCX CGT TG)-5′ containing a single nucleotide bulge at the position of **X**. The difference of the melting temperature (ΔT_m) in the absence and presence of the molecule (100 μM) in sodium cacodylate buffer (10 mM, pH 7.0) was summarized in Table 1.

In the presence of **1a** consisting of a naphthyridine ring and a urea linker, T_m was increased for all duplexes-containing single nucleotide bulges, but not for the fully matched 11-mer duplex. ΔT_m for cytosine, thymine, guanine, and adenine bulges were 4.8, 3.7, 2.4, and 3.1 °C, respectively. Ureidoquinoline **1b** that was lacked by one ring nitrogen compared to **1a** showed a comparable ΔT_m for cytosine and thymine bulges, but a decreased ΔT_m for guanine and adenine bulges. Ureidoquinazoline **1c**, where only one of two ring nitrogens can simultaneously participate in the hydrogen bonding to the bulged nucleotide, showed much lower ΔT_m for all bulges than **1b**. Truncation of naphthyridine and quinoline rings to a pyridine ring as in **1d** resulted in a complete loss of stabilizing effects of bulged duplexes. The significance of a urea structure of **1a** and **1b** for the stabilization of single nucleotide bulges was clearly demonstrated by comparing the ΔT_m values with those obtained by the corresponding amide compounds **2a** and **2b**. The ΔT_m obtained for the cytosine, thymine, and adenine bulges in the presence of **2a** was decreased by 3.8, 3.7, and 3.2 °C, respectively, from those obtained in the presence of **1a**, but the ΔT_m for the guanine bulge was slightly increased instead. A similar propensity of ΔT_m values was observed for **2b**. Quinolinylcarbamate **3b** where one hydrogen donor in a urea structure was replaced by hydrogen acceptor exhibited a median effect on the stabilization of single nucleotide bulges.

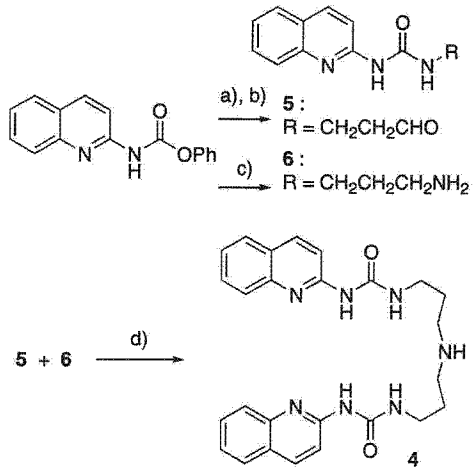
Comparing **1a** with **1b** and also **1b** with **2b**, the nitrogen at a position 8 of 1,8-naphthyridine was not necessary for the stabilization of cytosine, thymine, and adenine bulges, but hydrogen-bonding donor of a urea group was essential. The low ability of **1c** for the stabilization of single nucleotide bulges is most likely due to a conformational variety regarding the urea group. Conformational calculations of **1c** indicated that hydrogen bonding between N3 of quinazoline and N–H in urea would be stable by forming six-membered ring.[16] In that conformation, hydrogen-bonding groups aligned in the order of acceptor–donor–acceptor. Preorganization of hydrogen groups was reported an important factor for producing a stable hydrogen-bonded complex.[17–20] Marked modulations of the ΔT_m observed by changing N–H in **1b** into methylene and oxygen in **2b** and **3b**, respectively, clearly demonstrated the effect of hydrogen-bonding interactions on the stabilization of single nucleotide bulges.

Having found that ureidoquinoline **1b** effectively stabilized the cytosine and thymine bulges, the potential for the stabilization of a base mismatch by a dimeric form of **1b** was investigated. Ureidoquinoline dimer **4** was synthesized by a reductive amination of aldehyde **5** with primary amine **6**, which were obtained by a common precursor of phenylcarbamate (Scheme 1).[21] ΔT_m values for the base mismatches were obtained from the melting temperatures of 11-mer duplex containing a single base mismatch (**X**–**Y**) in the middle of the sequence (Table 2).

Table 1. Increased T_m (°C) of bulge-containing duplexes in the presence of testing molecules[a]

Drug	5′-d(TCCAG_GCAAC)-3′ 3′-d(AGGTCXCGTTG)-5′				Match (11-mer)
	X: C	T	G	A	
1a	4.8	3.7	2.4	3.1	−0.3
1b	4.9	4.2	1.8	2.6	−0.4
1c	1.1	1.4	−1.9	−0.1	−0.6
1d	0.8	0.0	−0.1	0.1	−0.7
2a	1.0	0.0	2.9	−0.1	0.2
2b	1.0	0.2	0.6	0.1	−0.1
3b	2.2	1.9	1.0	0.9	0.0

[a] [DNA base] = 100 μM, [ligand] = 100 μM, [NaCl] = 100 mM, [sodium cacodylate] = 10 mM (pH 7.0).



Scheme 1. Reagents and conditions: (a) 3,3-diethoxypropylamine, (b) AcOH, H_2O, 99% (two steps), (c) propylenediamine, 79%, (d) **5**+**6**, $NaBH_3CN$, AcOH, MeOH, 54%.

Table 2. Increased T_m (°C) of mismatch-containing duplexes in the presence of **4**[a]

5′-d(CTAACXGAATG)-3′ 3′-d(GATTGYCTTAC)-5′			
X–Y	ΔT_m	X–Y	ΔT_m
C–C	6.8	G–T	1.0
C–T	6.1	T–T	0.5
C–A	2.4	A–A	−0.2
G–G	2.4	G–C	0.0
G–A	2.0	A–T	−0.8

[a] [DNA base] = 100 μM, [ligand] = 100 μM, [NaCl] = 100 mM, [sodium cacodylate] = 10 mM (pH 7.0).

The ΔT_m of 6.8 °C was observed for the C–C mismatch, whereas the fully matched duplexes where **X**–**Y** were G–C and A–T were not stabilized at all under the same conditions. In marked contrast to dimer **4**, little increase of the T_m was observed for the C–C mismatch in the presence of ureidoquinoline **1b**, demonstrating that a covalent connection of two ureidoquinolines is quite effective for the stabilization of the C–C mismatch.[22] The ΔT_m of 6.1 °C was also observed for the C–T mismatch. Other mismatches including C–A, G–G, G–A, G–T, A–A, and T–T showed a small increase of their T_m values. Stabilization of C–C and C–T mismatches by **4** is consistent with that a monomeric form **1b** stabilize cytosine and thymine bulges. A complete failure of stabilizing the thymine–thymine mismatch by **4** implies that the stabilization of thymine–thymine mismatch needs a different strategy from that used for the G–G and C–C mismatches, although we have so far succeeded in stabilizing base mismatches by utilizing a dimeric form of bulge-stabilizing molecules.

In conclusion, ureidoquinolines were found a good molecular element for stabilizing single cytosine and thymine bulges. Integration of this binding element into its dimeric form provides a molecule stabilizing cytosine–cytosine and cytosine–thymine mismatches.

Acknowledgements

This work was supported by Grant-in-Aid for Scientific Research on Priority Areas (C) 'Medical Genome Science' from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References and notes

1. Syvänen, A.-C. *Nature Rev. Genet.* **2001**, *2*, 930–942.
2. Kwok, P. Y. *Annu. Rev. Genom. Hum. G.* **2001**, *2*, 235–258.
3. Schafer, A. J.; Hawkins, J. R. *Nat. Biotechnol.* **1998**, *16*, 33–39.
4. Jackson, B. A.; Barton, J. K. *J. Am. Chem. Soc.* **1997**, *119*, 12986–12987.
5. Jackson, B. A.; Alekseyev, V. Y.; Barton, J. K. *Biochemistry* **1999**, *38*, 4655–4662.
6. Lacy, E. R.; Cox, K. K.; Wilson, W. D.; Lee, M. *Nucleic Acids Res.* **2002**, *30*, 1834–1841.
7. Nakatani, K.; Sando, S.; Saito, I. *J. Am. Chem. Soc.* **2000**, *122*, 2172–2177.
8. Nakatani, K.; Sando, S.; Saito, I. *Nat. Biotechnol.* **2001**, *19*, 51–55.
9. Nakatani, K.; Sando, S.; Kumasawa, H.; Kikuchi, J.; Saito, I. *J. Am. Chem. Soc.* **2001**, *123*, 12650–12657.
10. Nakatani, K.; Sando, S.; Saito, I. *Bioorg. Med. Chem.* **2001**, *9*, 2381–2385.
11. Smith, E. A.; Kyo, M.; Kumasawa, H.; Nakatani, K.; Saito, I.; Corn, R. M. *J. Am. Chem. Soc.* **2002**, *124*, 6810–6811.
12. Nakatani, K.; Hagihara, S.; Sando, S.; Sakamoto, S.; Yamaguchi, K.; Maesawa, C.; Saito, I. *J. Am. Chem. Soc.* **2003**, *125*, 662–666.
13. Kobori, A.; Horie, S.; Suda, H.; Saito, I.; Nakatani, K. *J. Am. Chem. Soc.* **2004**, *126*, 557–562.
14. Hagihara, S.; Kumasawa, H.; Goto, Y.; Hayashi, G.; Kobori, A.; Saito, I.; Nakatani, K. *Nucleic Acids Res.* **2004**, *32*, 278–286.
15. Nakatani, K.; Kobori, A.; Kumasawa, H.; Saito, I. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1105–1108.
16. Although a similar cyclic conformation involving internal hydrogen bonding was also conceivable for **1b**, energy differences between cyclic and linear conformations were smaller for **1b** than **1c**.
17. Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008–2010.
18. Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810–2819.
19. Sartorius, J.; Schneider, H. J. *Chem. Eur. J.* **1996**, *2*, 1446–1452.
20. Murray, T. J.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1992**, *114*, 4010–4011.
21. Sigmund, H.; Pfleiderer, W. *Helv. Chim. Acta* **1994**, *77*, 1267–1280.
22. ΔT_m of 0.1 °C was observed for C–C mismatch in the presence of ureidoquinoline **1b** at the concentration of 100 μM in 100 mM sodium cacodylate buffer (pH 7.0) containing 100 mM NaCl.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters
ELSEVIER
Bioorganic & Medicinal Chemistry Letters 15 (2005) 259–262

A new ligand binding to G–G mismatch having improved thermal and alkaline stability

Tao Peng,[b] Takashi Murase,[a] Yuki Goto,[a] Akio Kobori[b] and Kazuhiko Nakatani[a,b,*]

[a]*Department of Synthetic Chemistry and Biological Chemistry, Faculty of Engineering, Kyoto University, Kyoto 606-8501, Japan*
[b]*PRESTO, Japan Science and Technology Agency (JST), Kyoto 615-8510, Japan*

Received 4 October 2004; revised 28 October 2004; accepted 30 October 2004
Available online 21 November 2004

Abstract—Naphthyridine dimer (**ND**) specially binds to guanine–guanine (G–G) mismatch in duplex DNA. In order to improve the thermal and alkaline stability and binding ability of the ligand, we have examined structural modification of the linker. A new ligand (**NNC**) possessing 2-amino-1,8-naphthyridines and a carbamate linker is much more thermally stable than **ND**. The half-life of **NNC** is 2.5 times longer than that of **ND** at 80 °C. **NNC** is also much more stable than **ND** under alkaline conditions. In addition, **NNC** binds to G–G mismatch more strongly than **ND**. The improved stability and the binding of **NNC** to the G–G mismatch would be suitable for the practical use of **NNC**-immobilized sensor.

Since a draft sequence of the human genome was determined,[1,2] some 1.6 million human single nucleotide polymorphisms (SNPs) have been found in the human genome and deposited to public databases.[3] SNPs became extremely important as a genetic marker for the identification of disease genes and detection of genetic mutations.[4,5] Thus, simple and rapid detection of a single nucleotide difference in the DNA sequences is an indispensable technique for both SNP mapping and typing. Although a number of methods have been developed for SNPs typing,[4,6,7] there is still a great need for designing new typing methods that are simple in operation, rapid and accurate in analysis, and low in cost.

We have recently reported a novel approach for the detection of SNPs by sensing guanine–guanine (G–G) mismatches in duplex DNA.[8] We have developed a sensor chip that can detect G–G mismatches in duplex DNA by means of surface plasmon resonance (SPR).[9,10] The sensor was prepared by immobilizing mismatch binding ligand naphthyridine dimer (**ND**) onto the carboxylated dextran matrix on the gold surface. We have reported that **ND** binds selectively to G–G mismatches in duplex DNA.[8,11] During the regeneration process of the **ND**-immobilized surface under alkaline conditions after each mismatch analysis, it was observed that the immobilized **ND** was slowly degraded under the conditions. We also found that high temperature necessary for denaturing the bound duplex on **ND**-immobilized sensor induced the **ND** degradation. Improving the thermal and alkaline stability of the mismatch-binding ligand eventually leads to a prolonged sensor lifetime. We report here a novel G–G mismatch binding ligand (**NNC**) that has not only greatly improved thermal and alkaline stability but also the higher affinity and selectivity to the G–G mismatch compared to **ND** (Fig. 1).

To gain insights into the degradation pathway, we first examined the thermal reaction of **ND** at 80 °C in 100 mM sodium cacodylate (pH 7.0) by HPLC (Fig. 2).

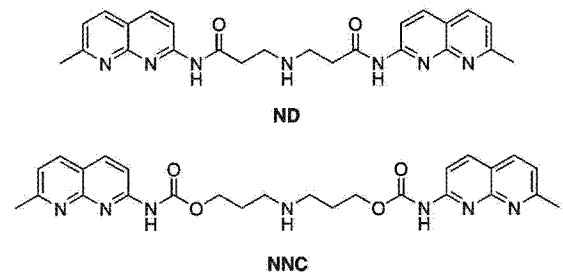


Figure 1. **ND** and **NNC**.

Keywords: DNA; Recognition; Mismatch.
* Corresponding author. Tel.: +81 753832756; fax: +81 753832759; e-mail: nakatani@sbchem.kyoto-u.ac.jp

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.11.003

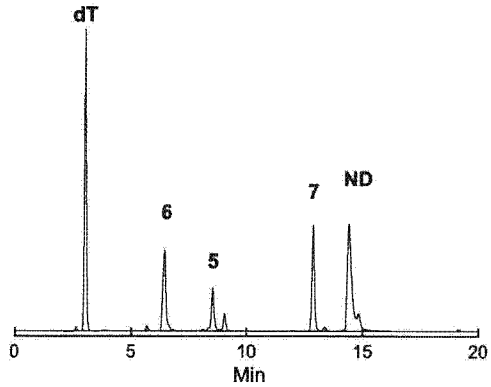


Figure 2. HPLC profile for the thermolysis of **ND** (0.71 mM) in 100 mM sodium cacodylate buffer (pH 7.0) for 45 min at 80 °C. dT was added as an internal standard.

The dT was selected as an internal standard for the thermolysis so that the reproducible and quantitative data could be obtained from the chromatographs. We detected three major products, which were identified as 2-amino-7-methyl-1,8-naphthyridine (**5**), 3-amino-*N*-(7-methyl-1,8-naphthyridin-2-yl)-propionamide (**6**), and *N*-(7-methyl-1,8-naphthyridin-2-yl)-acrylamide (**7**). The formation of **5** suggested the hydrolysis of the amide linkage, where as β-elimination was another degradation pathway producing **6** and **7** (Fig. 3).

To suppress both degradation processes and retain the binding ability to the G–G mismatch, a new molecule **NNC**, where amide linkage was substituted by a carbamate linkage, was synthesized. In addition, the alkyl chain length was further extended by one carbon for each side to slow down the nucleophilic addition of the secondary amino group in the linker to the carbonyl group leading to a release of **5**.

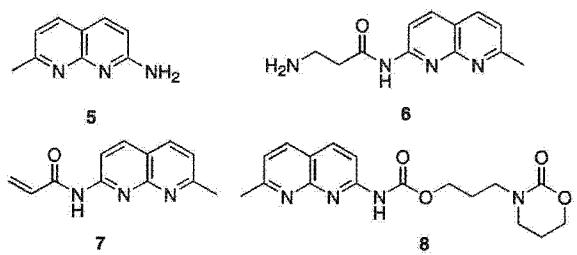


Figure 3. The products derived from thermolysis of **ND** and **NNC**.

NNC was synthesized as shown in Scheme 1. *N*-Boc-dipropanolamine was reacted with *N*,*N*′-disuccinimidyl carbonate (DSC) in dry acetonitrile to produce carbonate,[12] which was then reacted with 2-amino-7-methyl-1,8-naphthyridine to afford Boc-protected **NNC**. Deprotection by hydrogen chloride in ethyl acetate gave hydrochloride salt of **NNC**.[13]

The thermal reaction of **NNC** was examined under the same condition as that of **ND** (Fig. 4).

The major products of the **NNC** degradation after incubating at 80 °C for 120 min were identified as **5** and (7-methyl-1,8-naphthyridin-2-yl)-carbamic acid 3-(2-oxo-1,3-oxazinan-3-yl)-propyl ester (**8**) (Fig. 3). After a periodic incubation, the amount of **ND** and **NNC** were analyzed by HPLC. The rate of thermolysis could be determined from the decrease of the ligands. The half-life curves for the thermolysis of **NNC** and **ND** were shown in Figure 5. It is clearly shown that the half-life of **ND** is about 40 min at 80 °C, whereas the half-life of

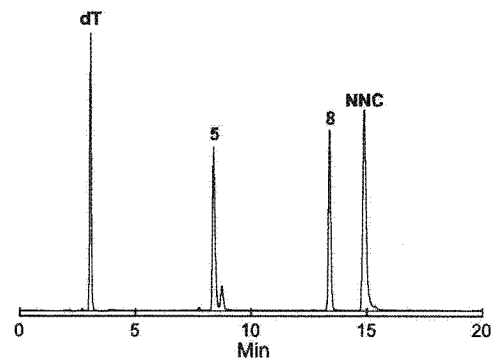
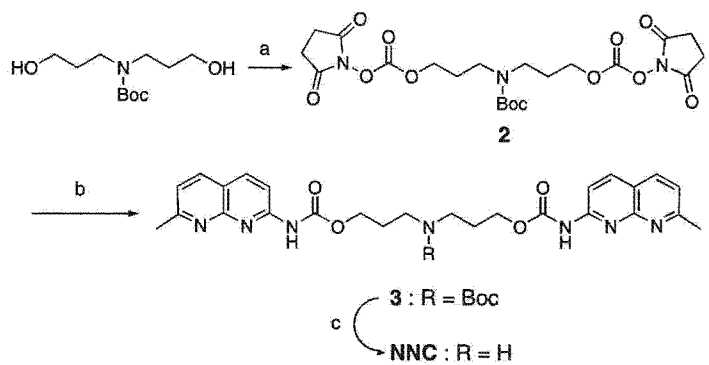


Figure 4. HPLC profile for the thermolysis of **NNC** (0.71 mM) in 100 mM sodium cacodylate buffer (pH 7.0) for 120 min at 80 °C.



Scheme 1. Reagents and conditions: (a) *N*,*N*′-disuccinimidyl carbonate, CH_3CN, Et_3N; (b) 2-amino-7-methyl-1,8-naphthyridine, CH_2Cl_2, Et_3N, 49% for two steps; (c) HCl, AcOEt, $CHCl_3$, quantitative.

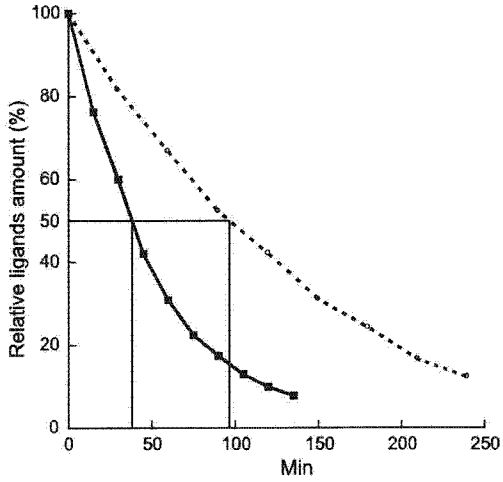


Figure 5. The half-life curves of **ND** (—) and **NNC** (···) in 100 mM sodium cacodylate buffer (pH 7.0) at 80 °C. *Y*-Axis represents the relative amount of ligands remained after incubation.

NNC is about 100 min. These results showed that the ligand **NNC** is much more thermally stable than **ND**.

The stabilities of the ligands in alkaline conditions were also examined incubating at room temperature in 50 mM sodium hydroxide (Fig. 6). Under the alkaline conditions, only 29% of **ND** remained after 2 h incubation, whereas 86% of **NNC** still remained under the same conditions. After incubation for 4 h, the amount of **ND** and **NNC** remaining was 7% and 79%, respectively. It is very clear that **NNC** is much more stable than **ND** in an alkaline solution.

Having confirmed the improved stability of **NNC** under the thermal and alkaline conditions, we then looked at

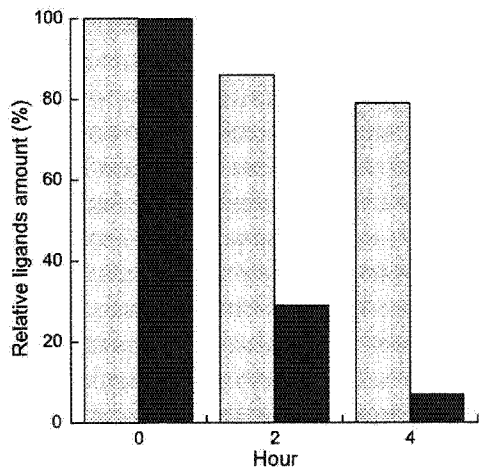


Figure 6. The amount of **ND** (solid bar) and **NNC** (shaded bar) (100 μM) remaining after incubation in 50 mM sodium hydroxide and 100 mM sodium chloride at room temperature. The amount was obtained by HPLC relative to the dT added as an internal standard. *Y*-Axis represents the relative amount of ligands remaining after incubation.

Table 1. ΔT_m of mismatch-containing duplexes in the presence ligands[a]

X–*Y*	T_m	ΔT_m[b]	
		NNC	ND
A–A	17.8 (1.4)	1.5 (0.2)	−0.8 (1.1)
A–C	16.1 (0.8)	4.1 (0.2)	2.1 (0.2)
C–C	18.2 (0.2)	6.1 (0.3)	6.7 (0.6)
G–A	25.7 (0.2)	7.0 (0.2)	8.6 (1.3)
G–G	25.6 (1.3)	29.1 (0.2)	23.7 (1.2)
G–T	28.3 (0.2)	1.2 (0.8)	10.2 (1.3)
T–C	18.6 (0.2)	3.7 (0.3)	5.5 (0.7)
T–T	25.1 (0.2)	0.7 (0.7)	0.6 (1.2)
A–T	34.3 (0.3)	0.0 (0.3)	−1.5 (0.3)
G–C	40.3 (0.0)	−2.0 (0.4)	2.0 (0.9)

[a] The UV-melting curve was measured for a duplex of d(CTA AC*X* GAA TG)/d(CAT TC*Y* GTT AG) at a total base concentration of 100 μM in a 10 mM sodium cacodylate buffer (pH 7.0) containing 0.1 M NaCl. A mismatch (*X*–*Y*) is produced in the middle of the duplex. Temperature was increased at a rate of 1 °C/min. All measurements were taken three times, and standard deviations are shown in the parentheses.

[b] ΔT_m is calculated as a difference of T_m in the presence and absence of drugs (100 μM), respectively.

the selective binding of **NNC** to the G–G mismatch. The assay was carried out by measuring the melting temperature (T_m) of 11-mer duplexes d(CTA AC*X* GAA TG)/d(CAT TC*Y* GTT AG) (where *X*, *Y* = A, G, T, or C) containing mismatches in the absence and presence of **NNC** (Table 1). T_m increase (ΔT_m) of the duplex containing a G–G mismatch was 23.7 °C in the presence of **ND** (100 μM). Under identical conditions, ΔT_m of 29.1 °C was recorded in the presence of **NNC**. The difference of ΔT_m ($\Delta\Delta T_m$) between **NNC** and **ND** for the G–G mismatch was 5.4 °C, suggesting that the modification of the linker structure of **ND** to that of **NNC** has a positive effect for the thermodynamic stabilization of the G–G mismatch by **NNC**. This is most likely due to an expanded π-surface in a carbamate linkage and a release of the linker strain involved in the bound DNA–**ND** complex. The affinity of **NNC** to the G–G mismatch was calculated by the curve fitting of the UV-melting curve obtained in the absence and presence of **NNC** to the theoretical equation.[14] The K_a obtained for the assumed 1:1 binding between **NNC** to the G–G mismatch was $>10^7\,M^{-1}$, that is larger than the K_a we reported for the **ND** binding to the G–G mismatch.

All of the data presented here showed that the ligand **NNC** is a better ligand than **ND** in terms of the affinity and selectivity to the G–G mismatch, and the thermal and alkaline stability. The use of **NNC** for the SPR sensor would enhance the sensitivity to the G–G mismatch by eliminating the unnecessary binding to DNA containing other mismatches. Furthermore, with a higher thermal and alkaline stability, **NNC** can be applied more expansively than **ND**.

References and notes

1. Lander, E. S. et al. *Nature* **2001**, *409*, 860–921.
2. Venter, C. J. et al. *Science* **2001**, *291*, 1304–1351.

3. http://www.ncbi.nlm.nih.gov/SNP/.
4. Schafer, A. J.; Hawkins, J. R. *Nat. Biotechnol.* **1998**, *16*, 33–39.
5. Collins, F. S.; Guyer, M. S.; Chakravarti, A. *Science* **1997**, *278*, 1580–1581.
6. SyvaÈnen, A. C. *Nature Rev. Genet.* **2001**, *2*, 930–942.
7. Kwok, P. Y. *Annu. Rev. Genom. Hum. Genet.* **2001**, *2*, 235–258.
8. Nakatani, K.; Sando, S.; Saito, I. *Nat. Biotechnol.* **2001**, *19*, 51–55.
9. Nice, E. C.; Catimel, B. *Bioessays* **1999**, *21*, 339–352.
10. Fivash, M.; Towler, E. M.; Fisher, R. *Curr. Opin. Biotechnol.* **1998**, *9*, 97–101.
11. Nakatani, K.; Sando, S.; Kumasawa, H.; Kikuchi, J.; Saito, I. *J. Am. Chem. Soc.* **2001**, *123*, 12650–12657.
12. Ghosh, A. K.; Duong, T. T.; McKee, S. P.; Thompson, W. J. *Tetrahedron Lett.* **1992**, *33*, 2781–2784.
13. The data of the hydrochloride salt of NNC: ^{1}H NMR (CD_3OD, 400 MHz): $\delta = 8.15$ (m, 6H), 7.35 (d, 2H, $J = 8.4$ Hz), 4.34 (t, 4H, $J = 6.0$ Hz), 3.15 (t, 4H, $J = 7.2$ Hz), 2.71 (s, 6H), 2.11 (br s, 4H). ^{13}C NMR (CD_3OD, 400 MHz): $\delta = 163.0$, 154.5, 154.3, 154.0, 139.1, 137.6, 121.4, 118.2, 113.1, 62.7, 45.4, 26.5, 23.8. HR-FABMS calcd for $C_{26}H_{30}N_7O_6$ [$(M+H)^+$], 504.2359. Found: 504.2361.
14. Nakatani, K.; Horie, S.; Murase, T.; Hagihara, S.; Saito, I. *Bioorg. Med. Chem.* **2003**, *11*, 2347–2353.

Nucleic Acids Symposium Series No. 48 123-124

Separation of mismatched DNA by using the affinity column immobilizing mismatch-binding ligands

Yuki Goto[1], Akio Kobori[2], Hitoshi Suda[1] and Kazuhiko Nakatani[1,2]
[1]Department of Synthetic Chemistry and Biological Chemistry, Faculty of Engineering, Kyoto University, Kyoto 615-8510, Japan and [2]PRESTO, Japan Science and Technology Agency (JST), Kawaguchi City, Saitama 332-0012, Japan

ABSTRACT

Mismatch-binding ligands (MBL), ND and NA, bind to DNA containing mismatch base pairs. These MBLs were immobilized to NHS-activated affinity column via amine linker. Affinity chromatographic analyses of mismatched DNAs by using the ND-immobilized column showed clear separation of G-G and G-A from other mismatched DNAs. The results of the NA-immobilized column also showed good separation of A-A, G-A, G-G, A-C from other mismatched DNAs. The order of mobility of mismatched duplexes on MBL-immobilized column was in good agreement with that of the binding affinity in a solution phase obtained by the melting temperature measurements.

INTRODUCTION

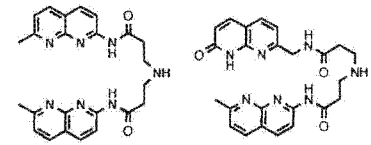
High-throughput SNPs typing methods are expected to be needed for realizing personalized medicine. One of methods for SNPs typing is heteroduplex analyses which detect heteroduplex DNA, containing a single mismatched site, generated by hybridization of two sets of duplex DNAs that differ from each other by a single nucleotide. Conventional low-throughput methods for the detection of mismatched DNAs in heteroduplex analyses were enzymatic and chemical cleavage at the mismatched site, gel electrophoresis, and the binding of mismatch-binding proteins.

We have reported mismatch-binding ligands (MBL) that bind to specific mismatch base pairs in the DNA duplex.[1-3] We have immobilized MBLs on SPR sensors providing novel mismatch-detecting sensors for high-throughput heteroduplex analyses.[1-3]

We here report novel applications of MBLs in the separation of mismatched DNA with the MBL-immobilized affinity chromatography. This technique will provide high-throughput and inexpensive method of SNPs typing.

RESULTS AND DISCUSSION

We have reported two MBLs, naphthyridine dimer (**ND**) and naphthyridine-azaquinolone (**NA**).



ND **NA**

ND strongly binds to G-G mismatch, whereas **NA** binds to A-A and G-A mismatches.

ND and **NA** were first attached to the short amine linker, and then immobilized on HiTrap NHS-activated HP Column (Amersham Biosciences). These MBL-immobilized columns were connected to ÄKTAexplorer system (Amersham Biosciences). The performance of the MBL-immobilized columns were examined by the analyses of hairpin DNAs, 5'-d(CTA ACX GAA TGT TTT CAT TCY GTT AG)-3', containing mismatched base pairs (***Figure 1***). The DNAs were eluted with a solvent mixture of running buffer [100 mM NaCl, 10 mM Na phosphate pH = 7.0], 0-10% elution buffer [100 mM NaCl, 50 mM NaOH]/3-4 min, 10-35%/4-14 min, and 35-50%/14-16 min at a flow rate 1.0 mL/min. Retention time (t_R) and half width ($W_{1/2}$) of the peak of each mismatched DNA was summarized in ***Figure 2***. G-G and G-A mismatched DNAs were strongly retained on **ND**-immobilized column and were eluted with flowing the elution buffer. Other mismatched and full matched DNAs were eluted from the column with the running buffer. t_R of C-C, G-T, and T-C mismatched DNAs (1.32-1.84 min) obviously differed from that of full matched DNAs (1.15-1.17 min), and $W_{1/2}$ of C-C, G-T, and T-C mismatched DNAs (0.44-0.87 min) were wider than that of full matched DNAs (0.35 min).

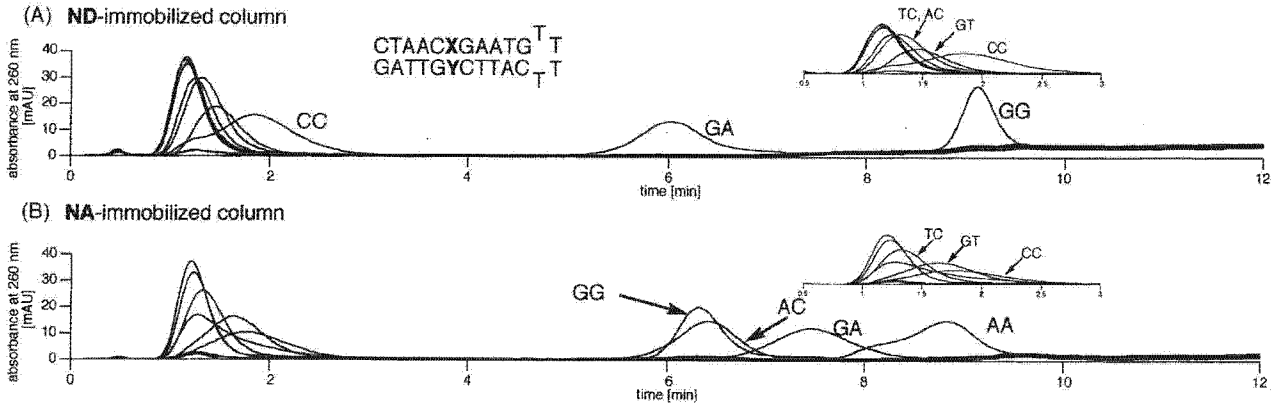


Figure 1. The affinity chromatography analyses of hairpin DNAs containing mismatched base pairs by using (A) **ND**-immobilized column and (B) **NA**-immobilized column. Inset: for clarity, the range of time from 0.5 to 3.0 was shown.

	ND			NA		
sequence	t_R (min)	$W_{1/2}$ (min)	ΔT_m (°C)	t_R (min)	$W_{1/2}$ (min)	ΔT_m (°C)
GG	9.12	0.37	24.0	6.33	0.52	6.8
GA	6.03	0.69	11.7	7.46	0.89	12.1
GT	1.45	0.51	9.9	1.78	0.89	1.4
GC	1.17	0.35	2.6	1.22	0.35	0.9
AA	1.17	0.36	2.4	8.83	0.85	31.8
AT	1.15	0.35	-1.3	1.28	0.59	0.4
AC	1.25	0.39	4.8	6.42	0.69	6.7
TT	1.17	0.36	-0.8	1.24	0.35	0.4
TC	1.32	0.44	2.7	1.33	0.44	2.3
CC	1.84	0.87	5.5	1.64	0.69	4.2

Figure 2. The retention time (t_R) and half width ($W_{1/2}$) of the peak of each mismatched DNA in the affinity chromatograph were shown. ΔT_m values of 11-mer duplex with MBLs were also shown.

These observations suggested that these mismatched DNAs were separated by the weak interactions between DNAs and **ND** surface on the solid support. In the case of **NA** column, A-A, G-A, G-G, and A-C mismatched DNAs were retained on **NA** surface, whereas other mismatches were eluted with running buffer. C-C, G-T, and T-C mismatched DNAs were not retained strongly, but their t_R (1.33-1.78 min) were larger than that of full matched DNA (1.22-1.28 min). These results of the affinity chromatography indicated that the MBL-immobilized columns are useful to separate not only mismatched DNAs strongly binding but also those weakly binding to the surface.

Basic information regarding MBL-mismatch binding was obtained by measuring the difference in melting temperature (ΔT_m) of 11-mer duplex, 5'-d(CTA ACX GAA TG)-3'/3'-d(GAT TGY CTT AC)-5', in the absence and presence of MBLs. **ND** bound not only to G-G mismatched duplex as reported, but weakly to G-A, G-T, and C-C mismatched duplexes. On the other hand, **NA** bound not only to A-A and G-A mismatched duplexes but also G-G, A-C, and C-C mismatched duplexes with decreased efficiency. The order of mobility of mismatched duplex on MBL-immobilized column was in good agreement with that of the binding affinity in a solution phase obtained by T_m measurements.

The results described here demonstrated the potential of MBLs and MBL-immobilized surfaces for the affinity separation of single mismatched duplexes. The advantage of the use of the MBL-immobilized surfaces is the redundancy of any DNA labeling, which significantly reduces the cost and simplify the assay procedure of SNPs typing.

REFERENCES

1. Nakatani, K., Sando, S. and Saito, I. (2001) *Nat. Biotechnol.*, **19**, 51-55.
2. Hagihara, S., Kumasawa, H., Goto, Y., Hayashi, G., Kobori, A., Saito, I. and Nakatani, K. (2004) *Nucleic Acids Res.*, **32**, 278-286.
3. Kobori, A., Horie, S., Suda, H., Saito, I. and Nakatani, K. (2004) *J. Am. Chem. Soc*, **126**, 557-562

Nucleic Acids Symposium Series No. 48 129-130

SPR fingerprinting of mismatched base pair

Akio Kobori[1], Tao Peng[1], Gosuke Hayashi[2] and Kazuhiko Nakatani[1,2]
[1]PRESTO, Japan Science and Technology Agency and [2]Department of Synthetic Chemistry and Biological Chemistry, Faculty of Engineering, Kyoto University, Kyoto 615-8510, Japan

ABSTRACT

Discrimination of mismatched base pairs having various flanking sequence from normal Watson-crick base pairs were examined using SPR sensor surface where mismatch binding ligand (MBL) were immobilized. CC, CT, and TT mismatched base pair are distinguishable from wild type DNA by means of diaminonaphthyridine dimer (damND) immobilized sensor surface. By comparing the component ratio of responses obtained by MLB immobilized sensors, distinct differences could be observed between mismatched base pairs having almost the same total SPR response.

INTRODUCTION

Surface Plasmon Resonance (SPR) is an effective tool for directly visualizing the ligand-substrate interactions occurring on the sensor surface. We have studied the mismatch binding ligands (MBL) that selectively and strongly bind to the mismatched base pairs, and developed the SPR sensor surface where MBL are immobilized to detect the mismatched base pairs. Naphthyridine dimer (ND) that contains two naphthyridine chromophores and linker selectively binds to G-G mismatched base pair.[1] Naphthyridine-azaquinolone (NA) and aminonaphthyridine (amND) dimer bind to G-A and C-C mismatched base pairs, respectively.[2,3] Mismatch-detecting sensors were synthesized by immobilizing these ligands on gold surface. To distinguish all eight base mismatches from Watson-Crick base pairs, we have investigated the fingerprinting of base mismatches by these MBL. In addition to three MBL, a new MBL diaminonaphthyridine dimer (damND) was used for the fingerprinting. We here report the fingerprinting of mismatched base pairs flanked by all possible sequences.

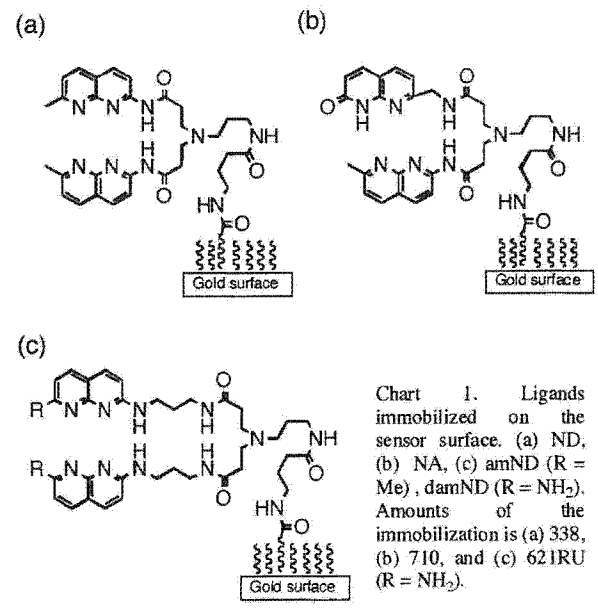


Chart 1. Ligands immobilized on the sensor surface. (a) ND, (b) NA, (c) amND (R = Me), damND (R = NH_2). Amounts of the immobilization is (a) 338, (b) 710, and (c) 621RU (R = NH_2).

RESULTS and DISCUSSION

Preparation of the ligand immobilized surface

MBL was immobilized on an activated carboxyl terminal attached to the dextran surface according to the procedure

recommended by Biacore. A 1 mM solution of MBL having a primary amino group at the linker termini in 10 mM of a borate buffer (pH = 9.2) was applied to the carboxylic acid of the CM5 sensor chip activated with *N*-hydroxysuccinimide and EDCI.

SPR fingerprinting of mismatched base pairs

The effect of the mismatched base pair with various flanking sequences on the binding to the ligand was examined using ND, NA, amND, and damND immobilized sensor surface. SPR analyses of mismatch-containing 27-mer 5'-d(GTT ACA GAA TCT N_1XN_2 AAG CCT AAT ACG)-3'/3'-d(CAA TGT TTC AGA N_3YN_4 TTC GGA TTA TGC)-5' were performed in HEPES buffer (pH 7.4) containing 500 mM NaCl. In addition to the GG and CC mismatches examined by ND and amND immobilized surface, the response at 180 sec after injection of DNA duplexes containing CT or TT mismatch were larger than 170 RU for all flanking sequences. These responses are more than 4 times larger than that for AT matched base pair (50 RU). Therefore, GG, CC, CT, and TT mismatched base pair are distinguishable from normal Watson-Crick base pair by single ND or damND immobilized sensor surface.

Fingerprinting of the remaining four mismatches, CA, GA, AA, and GT, which showed insufficient responses to each MBL-immobilized surface was carried out by taking a sum of responses obtained by damND, ND, and NA immobilized sensor surfaces at 180 sec. Typical examples of fingerprinting were shown for CCC/GAG (N_1XN_2/N_3YN_4), CGG/GAC, CCG/GAC, AGG/TTC, and CAG/GAC containing duplexes. (Figure 1) The sum of responses for the five mismatched base pairs was more than 200 RU and obviously distinguishable from the back ground noise. By comparing

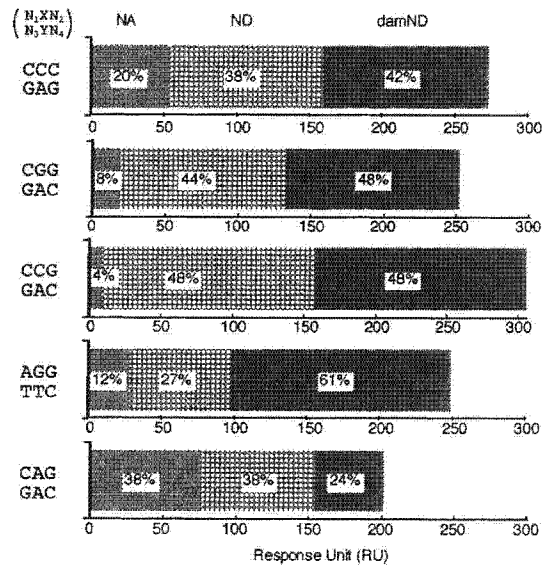


Figure 1. The sum of the responses obtained by damND, ND, and NA immobilized sensor surface at 180 sec after injection of 27mer 5'-d(GTT ACA GAA TCT N_1XN_2 AAG CCT AAT ACG)-3'/3'-d(CAA TGT CTT AGA N_3YN_4 TTC GGA TTA TGC)-5' (1 μM). Binding was measured in HBS-N (pH 7.4) containing 500 mM NaCl. Ratio of the response obtained by each sensor surfaces are shown in bar graph.

the component ratio of responses obtained by MLB immobilized surface, distinct differences were observed between individual mismatched base pairs, although the total response were almost the same for these mismatches. These results suggested that SPR fingerprinting obtained by MLB immobilized surface would be the effective way to differentiate the mismatched base pairs.

REFERENCES

1. Nakatani, K., Kobori, A., Kumasawa, H., Saito, I. (2004) *Bioorg. Med. Chem. Lett.*, **14**, 1105-1108.
2. Hagihara, S., Kumasawa, H., Hayashi, G., Kobori, A., Saito, I., Nakatani, K. (2004) *Nucleic Acids Res*, **32**, 273-286.
3. Kobori, A., Horie, S., Suda, H., Saito, I., Nakatani, K. (2004) *J. Am. Chem. Soc*, **126**, 557-562.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters
ELSEVIER
Bioorganic & Medicinal Chemistry Letters 15 (2005) 259–262

A new ligand binding to G–G mismatch having improved thermal and alkaline stability

Tao Peng,[b] Takashi Murase,[a] Yuki Goto,[a] Akio Kobori[b] and Kazuhiko Nakatani[a,b,*]

[a]*Department of Synthetic Chemistry and Biological Chemistry, Faculty of Engineering, Kyoto University, Kyoto 606-8501, Japan*
[b]*PRESTO, Japan Science and Technology Agency (JST), Kyoto 615-8510, Japan*

Received 4 October 2004; revised 28 October 2004; accepted 30 October 2004
Available online 21 November 2004

Abstract—Naphthyridine dimer (**ND**) specially binds to guanine–guanine (G–G) mismatch in duplex DNA. In order to improve the thermal and alkaline stability and binding ability of the ligand, we have examined structural modification of the linker. A new ligand (**NNC**) possessing 2-amino-1,8-naphthyridines and a carbamate linker is much more thermally stable than **ND**. The half-life of **NNC** is 2.5 times longer than that of **ND** at 80 °C. **NNC** is also much more stable than **ND** under alkaline conditions. In addition, **NNC** binds to G–G mismatch more strongly than **ND**. The improved stability and the binding of **NNC** to the G–G mismatch would be suitable for the practical use of **NNC**-immobilized sensor.

Since a draft sequence of the human genome was determined,[1,2] some 1.6 million human single nucleotide polymorphisms (SNPs) have been found in the human genome and deposited to public databases.[3] SNPs became extremely important as a genetic marker for the identification of disease genes and detection of genetic mutations.[4,5] Thus, simple and rapid detection of a single nucleotide difference in the DNA sequences is an indispensable technique for both SNP mapping and typing. Although a number of methods have been developed for SNPs typing,[4,6,7] there is still a great need for designing new typing methods that are simple in operation, rapid and accurate in analysis, and low in cost.

We have recently reported a novel approach for the detection of SNPs by sensing guanine–guanine (G–G) mismatches in duplex DNA.[8] We have developed a sensor chip that can detect G–G mismatches in duplex DNA by means of surface plasmon resonance (SPR).[9,10] The sensor was prepared by immobilizing mismatch binding ligand naphthyridine dimer (**ND**) onto the carboxylated dextran matrix on the gold surface. We have reported that **ND** binds selectively to G–G mismatches in duplex DNA.[8,11] During the regeneration process of the **ND**-immobilized surface under alkaline conditions after each mismatch analysis, it was observed that the immobilized **ND** was slowly degraded under the conditions. We also found that high temperature necessary for denaturing the bound duplex on **ND**-immobilized sensor induced the **ND** degradation. Improving the thermal and alkaline stability of the mismatch-binding ligand eventually leads to a prolonged sensor lifetime. We report here a novel G–G mismatch binding ligand (**NNC**) that has not only greatly improved thermal and alkaline stability but also the higher affinity and selectivity to the G–G mismatch compared to **ND** (Fig. 1).

To gain insights into the degradation pathway, we first examined the thermal reaction of **ND** at 80 °C in 100 mM sodium cacodylate (pH 7.0) by HPLC (Fig. 2).

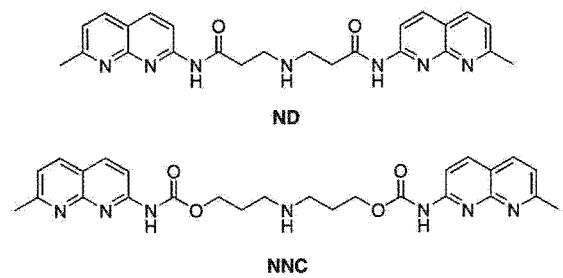


Figure 1. **ND** and **NNC**.

Keywords: DNA; Recognition; Mismatch.
* Corresponding author. Tel.: +81 753832756; fax: +81 753832759; e-mail: nakatani@sbchem.kyoto-u.ac.jp

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.11.003

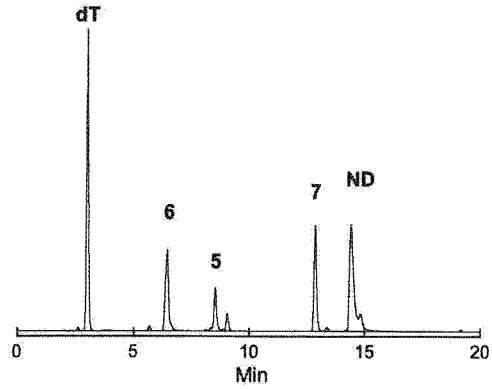


Figure 2. HPLC profile for the thermolysis of **ND** (0.71 mM) in 100 mM sodium cacodylate buffer (pH 7.0) for 45 min at 80 °C. dT was added as an internal standard.

The dT was selected as an internal standard for the thermolysis so that the reproducible and quantitative data could be obtained from the chromatographs. We detected three major products, which were identified as 2-amino-7-methyl-1,8-naphthyridine (**5**), 3-amino-*N*-(7-methyl-1,8-naphthyridin-2-yl)-propionamide (**6**), and *N*-(7-methyl-1,8-naphthyridin-2-yl)-acrylamide (**7**). The formation of **5** suggested the hydrolysis of the amide linkage, where as β-elimination was another degradation pathway producing **6** and **7** (Fig. 3).

To suppress both degradation processes and retain the binding ability to the G–G mismatch, a new molecule **NNC**, where amide linkage was substituted by a carbamate linkage, was synthesized. In addition, the alkyl chain length was further extended by one carbon for each side to slow down the nucleophilic addition of the secondary amino group in the linker to the carbonyl group leading to a release of **5**.

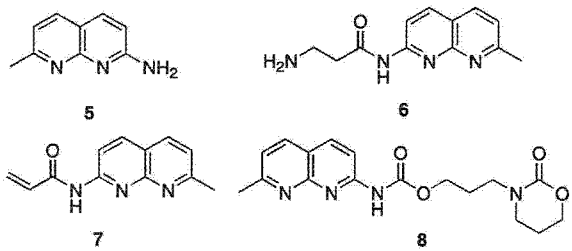


Figure 3. The products derived from thermolysis of **ND** and **NNC**.

NNC was synthesized as shown in Scheme 1. *N*-Boc-dipropanolamine was reacted with *N*,*N*′-disuccinimidyl carbonate (DSC) in dry acetonitrile to produce carbonate,[12] which was then reacted with 2-amino-7-methyl-1,8-naphthyridine to afford Boc-protected **NNC**. Deprotection by hydrogen chloride in ethyl acetate gave hydrochloride salt of **NNC**.[13]

The thermal reaction of **NNC** was examined under the same condition as that of **ND** (Fig. 4).

The major products of the **NNC** degradation after incubating at 80 °C for 120 min were identified as **5** and (7-methyl-1,8-naphthyridin-2-yl)-carbamic acid 3-(2-oxo-1,3-oxazinan-3-yl)-propyl ester (**8**) (Fig. 3). After a periodic incubation, the amount of **ND** and **NNC** were analyzed by HPLC. The rate of thermolysis could be determined from the decrease of the ligands. The half-life curves for the thermolysis of **NNC** and **ND** were shown in Figure 5. It is clearly shown that the half-life of **ND** is about 40 min at 80 °C, whereas the half-life of

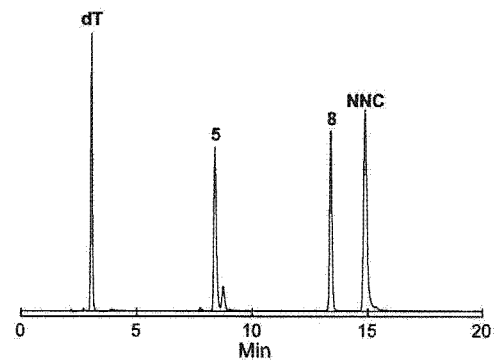
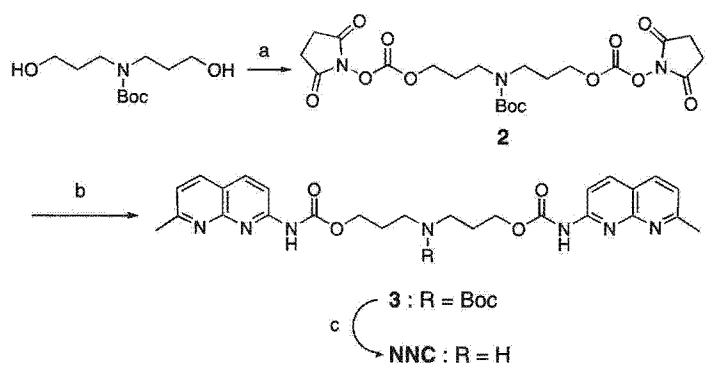


Figure 4. HPLC profile for the thermolysis of **NNC** (0.71 mM) in 100 mM sodium cacodylate buffer (pH 7.0) for 120 min at 80 °C.



Scheme 1. Reagents and conditions: (a) *N*,*N*′-disuccinimidyl carbonate, CH_3CN, Et_3N; (b) 2-amino-7-methyl-1,8-naphthyridine, CH_2Cl_2, Et_3N, 49% for two steps; (c) HCl, AcOEt, $CHCl_3$, quantitative.

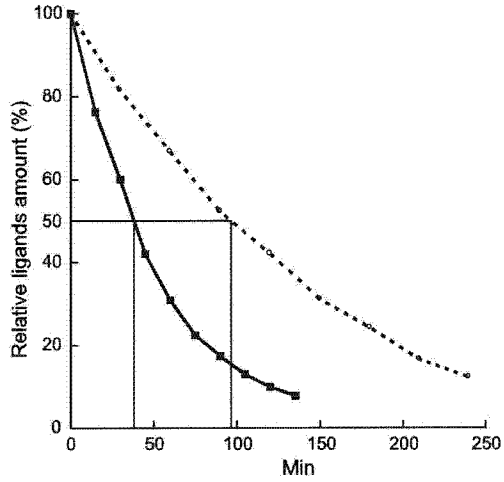


Figure 5. The half-life curves of **ND** (—) and **NNC** (···) in 100 mM sodium cacodylate buffer (pH 7.0) at 80 °C. *Y*-Axis represents the relative amount of ligands remained after incubation.

NNC is about 100 min. These results showed that the ligand **NNC** is much more thermally stable than **ND**.

The stabilities of the ligands in alkaline conditions were also examined incubating at room temperature in 50 mM sodium hydroxide (Fig. 6). Under the alkaline conditions, only 29% of **ND** remained after 2 h incubation, whereas 86% of **NNC** still remained under the same conditions. After incubation for 4 h, the amount of **ND** and **NNC** remaining was 7% and 79%, respectively. It is very clear that **NNC** is much more stable than **ND** in an alkaline solution.

Having confirmed the improved stability of **NNC** under the thermal and alkaline conditions, we then looked at

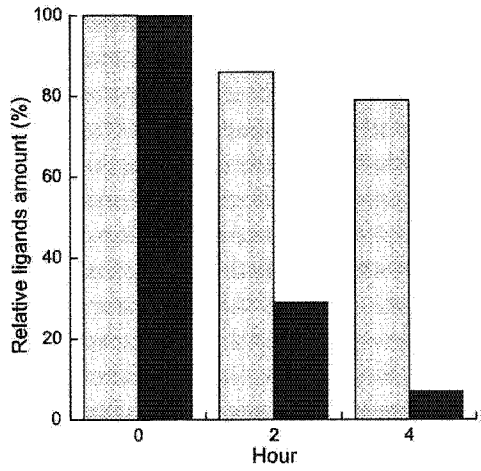


Figure 6. The amount of **ND** (solid bar) and **NNC** (shaded bar) (100 μM) remaining after incubation in 50 mM sodium hydroxide and 100 mM sodium chloride at room temperature. The amount was obtained by HPLC relative to the dT added as an internal standard. *Y*-Axis represents the relative amount of ligands remaining after incubation.

Table 1. ΔT_m of mismatch-containing duplexes in the presence ligands[a]

X–*Y*	T_m	ΔT_m[b]	
		NNC	ND
A–A	17.8 (1.4)	1.5 (0.2)	−0.8 (1.1)
A–C	16.1 (0.8)	4.1 (0.2)	2.1 (0.2)
C–C	18.2 (0.2)	6.1 (0.3)	6.7 (0.6)
G–A	25.7 (0.2)	7.0 (0.2)	8.6 (1.3)
G–G	25.6 (1.3)	29.1 (0.2)	23.7 (1.2)
G–T	28.3 (0.2)	1.2 (0.8)	10.2 (1.3)
T–C	18.6 (0.2)	3.7 (0.3)	5.5 (0.7)
T–T	25.1 (0.2)	0.7 (0.7)	0.6 (1.2)
A–T	34.3 (0.3)	0.0 (0.3)	−1.5 (0.3)
G–C	40.3 (0.0)	−2.0 (0.4)	2.0 (0.9)

[a] The UV-melting curve was measured for a duplex of d(CTA AC*X* GAA TG)/d(CAT TC*Y* GTT AG) at a total base concentration of 100 μM in a 10 mM sodium cacodylate buffer (pH 7.0) containing 0.1 M NaCl. A mismatch (*X*–*Y*) is produced in the middle of the duplex. Temperature was increased at a rate of 1 °C/min. All measurements were taken three times, and standard deviations are shown in the parentheses.

[b] ΔT_m is calculated as a difference of T_m in the presence and absence of drugs (100 μM), respectively.

the selective binding of **NNC** to the G–G mismatch. The assay was carried out by measuring the melting temperature (T_m) of 11-mer duplexes d(CTA AC*X* GAA TG)/d(CAT TC*Y* GTT AG) (where *X*, *Y* = A, G, T, or C) containing mismatches in the absence and presence of **NNC** (Table 1). T_m increase (ΔT_m) of the duplex containing a G–G mismatch was 23.7 °C in the presence of **ND** (100 μM). Under identical conditions, ΔT_m of 29.1 °C was recorded in the presence of **NNC**. The difference of ΔT_m ($\Delta\Delta T_m$) between **NNC** and **ND** for the G–G mismatch was 5.4 °C, suggesting that the modification of the linker structure of **ND** to that of **NNC** has a positive effect for the thermodynamic stabilization of the G–G mismatch by **NNC**. This is most likely due to an expanded π-surface in a carbamate linkage and a release of the linker strain involved in the bound DNA–**ND** complex. The affinity of **NNC** to the G–G mismatch was calculated by the curve fitting of the UV-melting curve obtained in the absence and presence of **NNC** to the theoretical equation.[14] The K_a obtained for the assumed 1:1 binding between **NNC** to the G–G mismatch was $>10^7\ M^{-1}$, that is larger than the K_a we reported for the **ND** binding to the G–G mismatch.

All of the data presented here showed that the ligand **NNC** is a better ligand than **ND** in terms of the affinity and selectivity to the G–G mismatch, and the thermal and alkaline stability. The use of **NNC** for the SPR sensor would enhance the sensitivity to the G–G mismatch by eliminating the unnecessary binding to DNA containing other mismatches. Furthermore, with a higher thermal and alkaline stability, **NNC** can be applied more expansively than **ND**.

References and notes

1. Lander, E. S. et al. *Nature* **2001**, *409*, 860–921.
2. Venter, C. J. et al. *Science* **2001**, *291*, 1304–1351.

3. http://www.ncbi.nlm.nih.gov/SNP/.
4. Schafer, A. J.; Hawkins, J. R. *Nat. Biotechnol.* **1998**, *16*, 33–39.
5. Collins, F. S.; Guyer, M. S.; Chakravarti, A. *Science* **1997**, *278*, 1580–1581.
6. SyvaÈnen, A. C. *Nature Rev. Genet.* **2001**, *2*, 930–942.
7. Kwok, P. Y. *Annu. Rev. Genom. Hum. Genet.* **2001**, *2*, 235–258.
8. Nakatani, K.; Sando, S.; Saito, I. *Nat. Biotechnol.* **2001**, *19*, 51–55.
9. Nice, E. C.; Catimel, B. *Bioessays* **1999**, *21*, 339–352.
10. Fivash, M.; Towler, E. M.; Fisher, R. *Curr. Opin. Biotechnol.* **1998**, *9*, 97–101.
11. Nakatani, K.; Sando, S.; Kumasawa, H.; Kikuchi, J.; Saito, I. *J. Am. Chem. Soc.* **2001**, *123*, 12650–12657.
12. Ghosh, A. K.; Duong, T. T.; McKee, S. P.; Thompson, W. J. *Tetrahedron Lett.* **1992**, *33*, 2781–2784.
13. The data of the hydrochloride salt of NNC: ^{1}H NMR (CD_3OD, 400 MHz): $\delta = 8.15$ (m, 6H), 7.35 (d, 2H, $J = 8.4$ Hz), 4.34 (t, 4H, $J = 6.0$ Hz), 3.15 (t, 4H, $J = 7.2$ Hz), 2.71 (s, 6H), 2.11 (br s, 4H). ^{13}C NMR (CD_3OD, 400 MHz): $\delta = 163.0$, 154.5, 154.3, 154.0, 139.1, 137.6, 121.4, 118.2, 113.1, 62.7, 45.4, 26.5, 23.8. HR-FABMS calcd for $C_{26}H_{30}N_7O_6$ [$(M+H)^+$], 504.2359. Found: 504.2361.
14. Nakatani, K.; Horie, S.; Murase, T.; Hagihara, S.; Saito, I. *Bioorg. Med. Chem.* **2003**, *11*, 2347–2353.

nature chemical biology

Small-molecule ligand induces nucleotide flipping in $(CAG)_n$ trinucleotide repeats

Kazuhiko Nakatani[1,2,5], Shinya Hagihara[1], Yuki Goto[1], Akio Kobori[2], Masaki Hagihara[2,5], Gosuke Hayashi[1], Motoki Kyo[3], Makoto Nomura[4], Masaki Mishima[4] & Chojiro Kojima[4]

DNA trinucleotide repeats, particularly CXG, are common within the human genome. However, expansion of trinucleotide repeats is associated with a number of disorders, including Huntington disease, spinobulbar muscular atrophy and spinocerebellar ataxia[1–4]. In these cases, the repeat length is known to correlate with decreased age of onset and disease severity[5,6]. Repeat expansion of $(CAG)_n$, $(CTG)_n$ and $(CGG)_n$ trinucleotides may be related to the increased stability of alternative DNA hairpin structures consisting of CXG-CXG triads with X-X mismatches[7–11]. Small-molecule ligands that selectively bound to CAG repeats could provide an important probe for determining repeat length and an important tool for investigating the *in vivo* repeat extension mechanism. Here we report that napthyridine-azaquinolone (NA, 1) is a ligand for CAG repeats and can be used as a diagnostic tool for determining repeat length. We show by NMR spectroscopy that binding of NA to CAG repeats induces the extrusion of a cytidine nucleotide from the DNA helix.

The expansion of $(CAG)_n$ trinucleotide repeats in genomic DNA is related to the pathogenesis of Huntington disease, the spinobulbar muscular atrophy known as Kennedy disease, and spinocerebellar ataxia. The $(CAG)_n$ repeat in normal IT15 genes ranges from 6 to 39 trinucleotide repeats, but the trinucleotide is repeated up to 121 times in patients with Huntington disease[3]. Although the mechanism of repeat expansion remains unclear, it is believed to involve strand slippage during DNA synthesis mediated by the formation of an alternative DNA hairpin structure. Because the stability of the hairpin

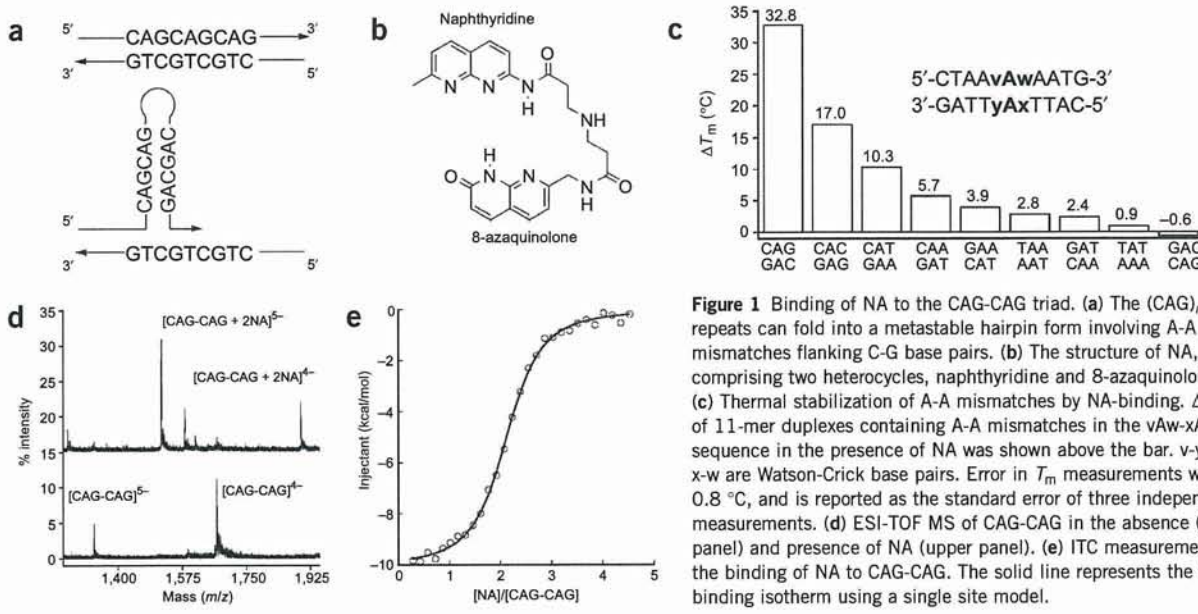


Figure 1 Binding of NA to the CAG-CAG triad. (**a**) The $(CAG)_n$ repeats can fold into a metastable hairpin form involving A-A mismatches flanking C-G base pairs. (**b**) The structure of NA, comprising two heterocycles, naphthyridine and 8-azaquinolone. (**c**) Thermal stabilization of A-A mismatches by NA-binding. ΔT_m of 11-mer duplexes containing A-A mismatches in the vAw-xAy sequence in the presence of NA was shown above the bar. v-y and x-w are Watson-Crick base pairs. Error in T_m measurements was $\pm$ 0.8 °C, and is reported as the standard error of three independent measurements. (**d**) ESI-TOF MS of CAG-CAG in the absence (lower panel) and presence of NA (upper panel). (**e**) ITC measurements for the binding of NA to CAG-CAG. The solid line represents the best-fit binding isotherm using a single site model.

[1]Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University, Kyoto 615-8510, Japan. [2]PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho Kawaguchi, Saitama 332-0012, Japan. [3]Biotechnology Frontier Project, TOYOBO Co. Ltd., Tsuruga, Fukui 914-0047, Japan. [4]Graduate School of Biological Science, Nara Institute of Science and Technology, Nara 630-0101, Japan. [5]Present address: The Institute of Scientific and Industrial Research (SANKEN), Osaka University, 8-1 Mihogaoka, Ibaraki 567-0047, Japan. Correspondence should be addressed to K.N. (nakatani@sanken.osaka-u.ac.jp) or C.K. (kojima@bs.naist.jp).

Published online 24 May 2005; doi:10.1038/nchembio708

form increases as the repeat expands, the repeat length is one of the most important determinants for diagnosis of disease severity and investigations of the expansion mechanism. The hairpin form of $(CAG)_n$ repeats involves the intramolecular pairing of CAG-CAG triads, with central A-A mismatches being flanked by two G-C base pairs (**Fig. 1a**). Thus, ligands that bind to the CAG-CAG triad also are expected to bind to the hairpin form of the $(CAG)_n$ repeat.

We previously discovered a naphthyridine-azaquinoline ligand (NA, **1**), which binds with high affinity to the CAG-CAG triad, during our investigations of small-molecular ligands that bind to base mismatches[12–15] as molecular elements for single-nucleotide polymorphism sensors[16,17] (**Fig. 1b**). The two heterocycles of NA, 2-amino-1,8-naphthyridine and 8-azaquinolone, present hydrogen bonding surfaces fully complementary to those of guanine and adenine, respectively. We assessed the binding of NA to CAG repeats through UV thermal denaturation studies. The thermal stability of a 11-mer duplex, 5′-d(CTAACAGAATG)-3′/5′-d(CATTCAGTTAG)-3′ (CAG-CAG), that contained a central CAG-CAG triad was enhanced by 34.8 °C in the presence of the NA ligand. NA-binding was highly sensitive to the identity of the base pairs flanking the central A-A mismatch (**Fig. 1c**). Qualitative analysis of the ΔT_m values suggested that NA binding involved interactions not only with the mismatched A-A pair but also with the 3′ G residue.

The stoichiometry of NA binding to the CAG-CAG triad was determined by electrospray ionization time-of-flight mass spectrometry (ESI-TOF MS) and isothermal titration calorimetry (ITC). Under ESI-MS conditions (**Fig. 1d**), CAG-CAG showed 4^- and 5^- ions of the duplex at *m/z* values of 1,668.78 and 1,334.70, respectively. Upon binding of NA to CAG-CAG, new ions appeared at *m/z* of 1,898.19 and 1,518.65, which correspond to the 4^- and 5^- ions of a complex consisting of two NA molecules and CAG-CAG. Complexes with other binding stoichiometry were not observed, even at higher NA concentrations. The association constant (K_a) of each NA molecule for the CAG-CAG triad was determined by ITC as $1.8 \times 10^6\ M^{-1}$ (**Fig. 1e**). The titration curve of the heat produced versus the DNA/ligand ratio supported a binding model involving a single set of identical sites, suggesting that two equivalent molecules of NA bind to and stabilize the CAG-CAG triad.

We carried out NMR spectroscopic structure determination to understand further how NA recognizes the CAG-CAG triad. We monitored the formation of a complex of the 11-mer duplex CAG-CAG and NA by observing shifts in one-dimensional ^1H imino proton resonances as increasing concentrations of NA were added. The imino proton chemical shifts of the NA-CAG-CAG complexes were completely different from those of the free DNA duplex. The existence of such large differences indicated that NA may intercalate into the DNA duplex[18]. Signals from free CAG-CAG and the NA-CAG-CAG complex were observed separately on a slow-exchange timescale. The stoichiometry was unambiguously determined to be 1:2 (DNA:NA), and no intermediate state was observed (**Fig. 2a** and **Supplementary Fig. 1** online). DNA signals were assigned by conventional procedures[19,20] using both unlabeled and $^{13}C^{15}N$-labeled DNA. NA signals were assigned by nuclear Overhauser effect spectroscopy (NOESY) and total correlation spectroscopy (TOCSY). Hydrogen bonding between NA and CAG-CAG was determined by analyzing the imino proton region (10–15 ppm) of NOESY (**Supplementary Fig. 2** online) and confirmed by a ^1H-^{15}N heteronuclear single-quantum coherence (HSQC) spectrum of $^{13}C^{15}N$-labeled CAG-CAG complexed with unlabeled NA (**Fig. 2b**). The solution structures of NA-CAG-CAG complexes were determined using 602 distance and dihedral constraints including 67 NA-DNA intermolecular distances (**Supplementary Fig. 3** online). Calculations of the final structure converged well, and the root-mean-square (r.m.s.) distance values were 0.79 Å for all heavy atoms of all residues, and as small as 0.54 Å for those in a well-converged region that includes A6, G7, A17, G18, NA1 and NA2

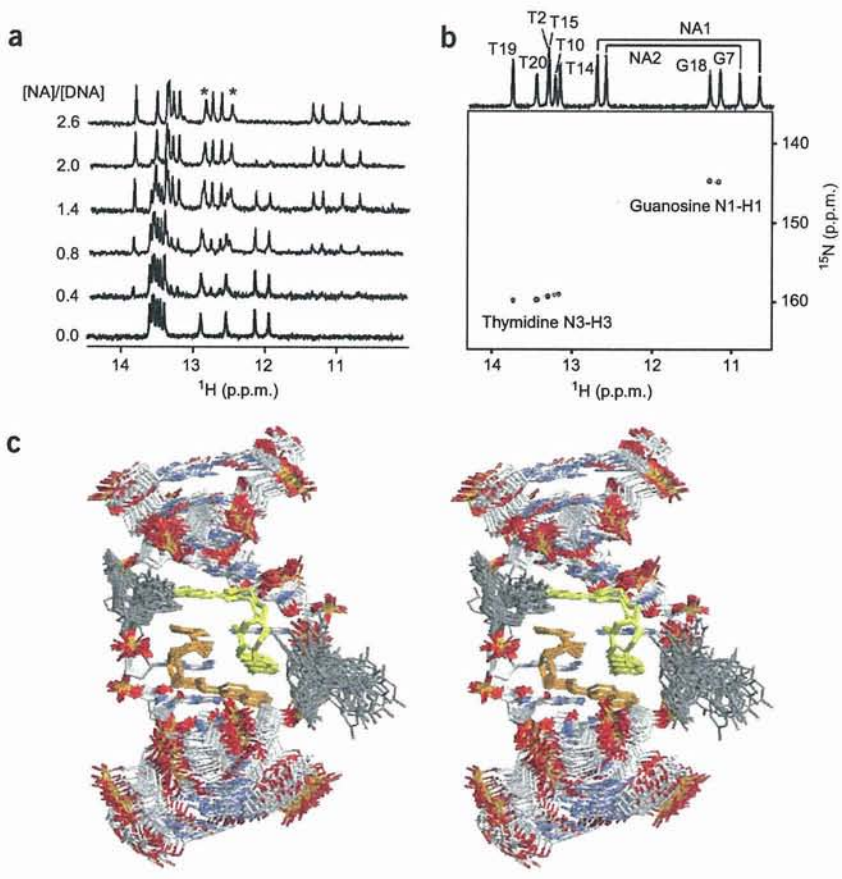


Figure 2 NMR structural analysis of NA-CAG-CAG complex. (**a**) One-dimensional ^1H spectra of 0.1 mM unlabeled 11-mer CAG-CAG at different NA concentrations at 275 K. The concentration ratios of NA to DNA are shown at left. The peaks marked with asterisks could not be observed at 293 K, the temperature at which the two-dimensional experiments were carried out. (**b**) A one-dimensional ^1H spectrum of 0.1 mM unlabeled 2:1 NA-CAG-CAG complex (top) and a ^1H-^{15}N HSQC spectrum of 0.2 mM $^{13}C^{15}N$-labeled DNA bound with the unlabeled NA (2:1 NA-CAG-CAG complex) (bottom) at 293 K. (**c**) NMR structures of NA-CAG-CAG complex. DNA is colored white, blue, red or gray except for the phosphate group, which is colored orange and red. Two NA molecules are colored yellow and orange. 30 complex structures are superimposed, focusing on A6, G7, A17, G18, NA1 and NA2 residues.

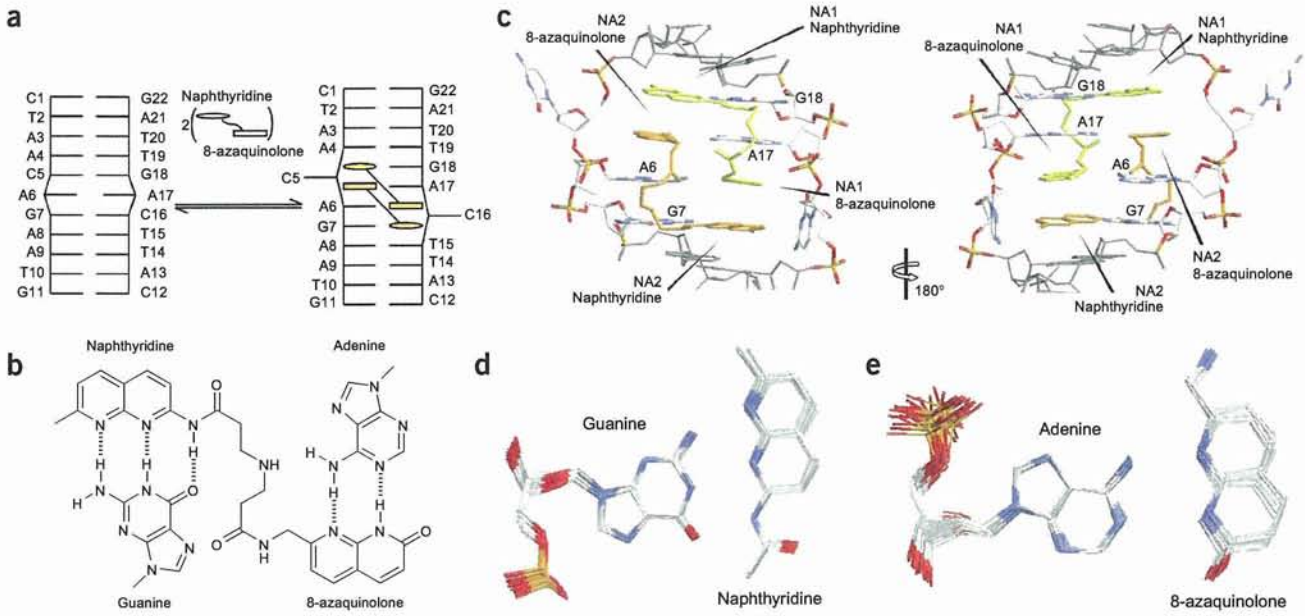


Figure 3 NA-CAG-CAG complex. (**a**) Schematic representation of the DNA conformation change induced by the complex formation, showing the free DNA (left) and NA-DNA complex (right). The face-to-face and overriding crossbars denote base pairing and base stacking, respectively. The bars sticking out to the sides, C5 and C16 of the NA-DNA complex, are the flipped-out bases. (**b**) Naphthyridine chromophore is complementary in hydrogen bonding surface to guanine, whereas 8-azaquinolone is complementary to adenine. (**c**) Major- and minor-groove views (left and right, respectively) of the NA-DNA complex structure. Two NA molecules are colored yellow (NA1) and orange (NA2), respectively. The focused and named DNA residues and NA molecules form intermolecular hydrogen bonds. (**d**) Superimposed NMR structures for the hydrogen bonding between guanine and naphthyridine. (**e**) Adenine and 8-azaquinolone.

(**Fig. 2c**). Structural and refinement statistics can be found in **Supplementary Tables 1** and **2** online.

The solution structures of the NA-CAG-CAG complex reveal how two NA molecules bind to one A-A mismatch and the 3′-G in CAG-CAG. NMR analyses showed that free CAG-CAG has a canonical B-type DNA conformation and the mismatched adenosine bases are stacked in the helix (**Fig. 3a**). In the NA-CAG-CAG complex, two mismatched adenosine bases form intermolecular hydrogen bonds with the 8-azaquinolone units of two NA molecules (**Fig. 3b**). The most unusual structural feature of the NA-CAG-CAG complex is the invasion of the G-C base pair by naphthyridine moieties. Thus, A6 and G18 bases were bound to the 8-azaquinolone and naphthyridine chromophores of NA1, respectively (**Fig. 3c**, yellow), and G7 and A16 were similarly bound to NA2 (**Fig. 3c**, orange) As a consequence, the two widowed cytidine nucleotides were extruded from the π-stack and produced two C-bulge loops. Base flipping has been observed in the complex between DNA and repair proteins[21]. The NA-CAG-CAG structure determined by NMR is notable because invasion of a small-molecule naphthyridine chromophore in NA forced the cytosine to flip out of the helix. We believe this is the first observation of induction of nucleotide flipping by a small-molecular ligand.

Naphthyridine-guanine and 8-azaquinolone–adenine pairs are well stacked in the right-handed DNA helix, showing structural mimicry of Watson-Crick base pairing (**Fig. 3d,e**). The structural studies indicated that each NA binds to an adenine base on one strand and a guanine base on the opposite strand. Neither the NMR data nor the UV melting experiment supports the possibility of an alternative intrastrand NA-CAG-CAG complex (**Supplementary Fig. 4** online). These cross-stranded interactions may explain the notably greater stability of the NA-CAG-CAG complex, as compared to the uncomplexed DNA, observed in the UV thermal denaturation

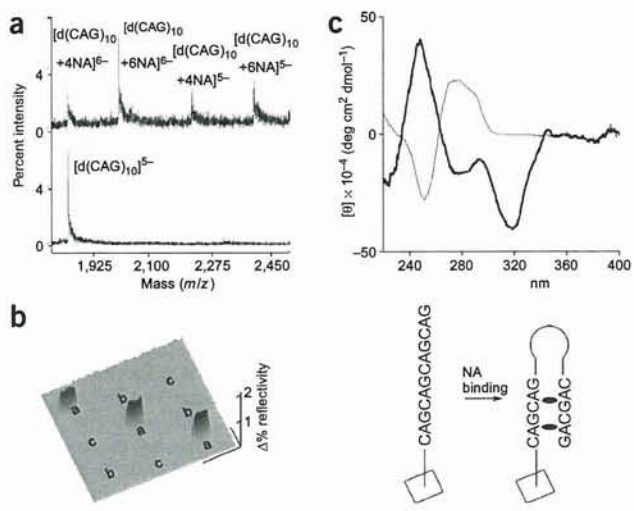


Figure 4 NA binding to the $(CAG)_n$ repeats. (**a**) ESI-TOF MS of $d(CAG)_{10}$ in the absence (below) and presence (above) of NA. (**b**) Left, SPR difference images of CAG and CTG repeats immobilized on a gold surface upon binding of NA. Key: a, HS-$d(T_{15}(CAG)_{10})$; b, HS-$d(T_{15}(CTG)_{10})$; c, blank. Right, an image of induced hairpin formation on the sensor resulting from NA binding. (**c**) CD spectral change induced on $(CAG)_{10}$ (5 μM) upon binding of NA (50 μM) with 100 mM NaCl in sodium cacodylate (pH 7.0, 10 mM). Key: $(CAG)_{10}$, thin line; $(CAG)_{10}$ + NA, bold line.

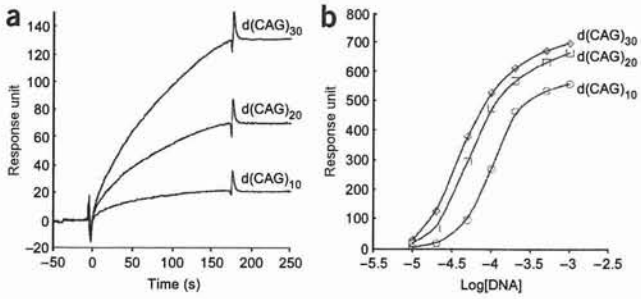


Figure 5 SPR analyses of the $(CAG)_n$ repeats by NA-immobilized sensor surface. (**a**) Binding of $d(CAG)_{10}$, $d(CAG)_{20}$ and $d(CAG)_{30}$ (each 20 nM) was measured in HEPES buffer (pH 7.4) containing 150 mM NaCl with the sensor surface where a dimeric form of NA was immobilized for 364 response units. (**b**) Dynamic range of NA-immobilized SPR sensor for the $(CAG)_n$ repeats detection. The signal intensity at 200 s of analysis was plotted. DNA concentrations were 10, 20, 50, 100, 200, 500 and 1,000 nM.

studies. These structural features also are consistent with ITC results indicating that the NA-CAG-CAG complex is stabilized by a large negative ΔH term (–16.2 kcal mol^{-1}), compensating for a negative ΔS term (–28.8 cal mol^{-1} K^{-1}).

Strong NA binding to the CAG-CAG triad induced formation of the NA-bound hairpin form in long $(CAG)_n$ repeats. ESI-TOF MS of a 30-mer single-stranded $d(CAG)_{10}$ containing ten CAG repeats in the presence of NA showed a series of ions corresponding to NA adducts of (4NA + $d(CAG)_{10}$) and (6NA + $d(CAG)_{10}$) (**Fig. 4a**). Each of these ions contained an even number of NA molecules, supporting the idea that the binding of NA to the $(CAG)_n$ repeat proceeds in a pairwise combination, as we demonstrated for the NA-CAG-CAG complex. To determine whether any conformational change is induced in the $(CAG)_n$ repeat upon binding of NA, we monitored surface plasmon resonance (SPR) difference images[22,23] of a gold surface modified with oligomers containing $d(CAG)_{10}$ and $d(CTG)_{10}$ (**Fig. 4b**). Upon exposure of this surface to an NA solution, SPR difference images were observed selectively at sites of $d(CAG)_{10}$ but not $d(CTG)_{10}$ immobilization. The reflectivity change in SPR at $d(CAG)_{10}$ spots upon NA binding was larger than that produced by hybridization with $d(CTG)_{10}$ (data not shown). Circular dichroism spectra of $d(CAG)_{10}$ also showed a large conformational change upon NA binding (**Fig. 4c**). Measurements of UV melting, ESI-TOF and fluorescence resonance energy transfer (FRET) labeling with the fluorophores 6-FAM and TAMRA at the 5′ and 3′ ends, respectively, also supported hairpin formation by $(CAG)_n$ (**Supplementary Fig. 5** online).

Given the ability of NA to bind CAG repeats, we created a sensor in which NA was immobilized on an SPR chip and assessed its utility for diagnosis of the CAG repeat length by SPR analysis. Because of the pairwise binding mechanism, we immobilized NA in dimeric form on the sensor surface (**Supplementary Methods** online). SPR analyses of the binding of $d(CAG)_{10}$, $d(CAG)_{20}$ and $d(CAG)_{30}$ to the immobilized NA dimer showed that signal intensities increased with repeat length (**Fig. 5a**). The SPR intensities of $d(CAG)_{30}$ were stronger than those of $d(CAG)_{10}$ and $d(CAG)_{20}$ at a wide range of DNA concentrations (**Fig. 5b**), suggesting that it may be possible to use the dimeric NA-immobilized SPR sensor for the rapid diagnosis of CAG repeat length.

Currently, there is no effective therapeutic agent for treating diseases caused by triplet repeat expansion. Recently, DNA alkylating agents have been shown to greatly reduce the $(CTG)_n$ repeat length in lymphoblast cells[24]. The discovery of the small-molecular ligand NA, which binds with high affinity to repeat sites, may be a substantial step toward developing effective therapeutic agents for these hereditary diseases.

METHODS

Materials. The ligand NA was synthesized as we have described[14]. All commercially available buffers and other chemicals were of the highest quality available. The oligonucleotides were purchased from Fasmac.

Melting temperature measurements. Melting temperatures of duplexes containing A-A mismatches (5 μM each strand) were recorded on a Shimadzu UV-2550 spectrophotometer with the TMSPC-8 analysis system in the absence and presence of NA (200 μM) in 10 mM sodium cacodylate buffer (pH 7.0) and 100 mM NaCl. ΔT_m is calculated as the difference of T_m in the presence and the absence of NA.

ESI-TOF MS measurements. Samples were prepared by mixing DNA (20 μM) and NA (120 μM) in 45–50% methanol in water containing 100 mM ammonium acetate. Mass spectra were obtained with an Applied Biosystems Mariner mass spectrometer and JEOL AccuTOF JMS-T100N mass spectrometer.

ITC measurements. A solution of CAG-CAG (10 μM) was titrated with NA solution (200 μM) at 5 °C in 10 mM sodium cacodylate buffer (pH 7.0) and 100 mM NaCl on a MicroCal VP-ITC calorimeter. Thermodynamic parameters were calculated from the binding curve using analytical software supplied with the instrument with a binding model involving a single set of identical sites.

NMR experiments. All data were collected on Bruker AVANCE500 and DRX800 NMR spectrometers. Titration experiments were carried out at 275 K using 0.1 and 0.8 mM unlabeled DNA duplex. Water-flipback NOESY spectra with mixing times of 30, 100, 150 and 200 ms were recorded at 293 and 303 K in H_2O or D_2O using the 0.8 mM unlabeled sample. TOCSY and double-quantum filtered correlation spectroscopy (DQF-COSY) spectra were recorded under similar conditions. ^1H-^{15}N HSQC, ^1H-^{13}C HSQC, and three-dimensional HCCH-TOCSY spectra were observed at 293 K in H_2O using a 0.2 mM uniformly $^{13}C^{15}N$-labeled sample. The labeled sample was prepared using a primer extension method[25]. All samples were dialyzed against a 50 mM sodium phosphate buffer (pH 6.5), 100 mM NaCl and 0.1 mM EDTA before measurements.

Structure determination. Interproton distance bounds were determined from the integrated peak intensities by the random error MARDIGRAS (RAND MARDI) procedure of the complete relaxation matrix analysis method[26]. Based on DQF-COSY, NOESY and phosphorus spectra, sugar puckers and backbone torsion angles were restrained to maintain an S-type sugar conformation and right-handed helix, respectively. Hydrogen-bonding restraints were imposed on the Watson-Crick base pairs and the NA-DNA hydrogen-bonding pairs. 422 distance constraints, including 58 sequential and 67 intermolecular distances and 180 dihedral angle constraints, were collected. The complex structure was calculated with a simulated annealing protocol using Crystallography & NMR System (CNS)[27]. Thirty structures without a distance violation greater than 0.5 Å were selected.

SPR imaging measurements. 5′-Thiol-modified oligomers HS-$d(T_{15}(CAG)_{10})$ and HS-$d(T_{15}(CTG)_{10})$, containing a dT_{15} spacer, were immobilized on the gold surface through the use of a hydrophilic heterobifunctional cross-linker[28]. The surface was exposed to NA (1 mM) for 200 s and then to buffer (10 mM phosphate, 300 mM NaCl, pH 7.4) in the SPR imaging instrument (Multi-SPRinter, Toyobo). The immobilized sites of the CAG and CTG repeats were confirmed by hybridization with corresponding complementary strands. The image was obtained by subtracting the data before and after interaction with NA.

SPR analyses for $(CAG)_n$ repeat number. A dimeric form of NA was synthesized by coupling of two molecules of NA and *N*-Boc imino-3,3′-bis(pentafluorophenylpropionate). After deprotection of the Boc group, the resulting secondary nitrogen was coupled with 4-*N*-Boc-amino-*N*-(3-oxopropyl)-butyramide. The terminal Boc group was removed to generate a

primary amine for the immobilization to the surface. A standard method recommended by Biacore was used for the immobilization to the carboxymethyl dextran surface of CM-5 sensor (Biacore). The amount of immobilized ligand was determined by the difference of SPR intensity before and after ligand immobilization. The SPR sensorgrams were obtained with a Biacore 2000 instrument (Biacore).

Accession codes. Atomic coordinates have been deposited in the Protein Data Bank under accession number 1X26.

Note: Supplementary information is available on the Nature Chemical Biology website.

ACKNOWLEDGMENTS

We thank T.L. James and M. Shimizu for valuable discussions. This work was partially supported by a Grant in Aid for Scientific Research (A) from the Japan Society for the Promotion of Science to K.N., Health and Labour Sciences Research Grants for Research on Advanced Medical Technology from the Ministry of Health, Labour and Welfare to K.N. and C.K., and CREST, Japan Science and Technology Agency to K.N.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 7 January; accepted 21 April 2005
Published online at http://www.nature.com/nchembio/

1. Ashley, C.T. & Warren, S.T. Trinucleotide repeat expansion and human disease. *Annu. Rev. Genet.* **29**, 703–728 (1995).
2. Paulson, H.L. & Fishbeck, K.H. Trinucleotide repeats in neurogenetic disorders. *Annu. Rev. Neurosci.* **19**, 79–107 (1996).
3. Wells, R.D. & Warren, S.T. (eds). (1998) *Genetic Instabilities and Hereditary Neurological Diseases.* Academic Press, San Diego.
4. McMurray, C.T. DNA secondary structure: A common and causative factor for expansion in human disease. *Proc. Natl. Acad. Sci. USA* **96**, 1823–1825 (1999).
5. Duyao, M.P. *et al.* Trinucleotide repeat length: instability and age of onset in Huntington's disease. *Nat. Genet.* **4**, 387–392 (1993).
6. Mangiarini, L. *et al.* Instability of highly expanded CAG repeats in mice transgenic for the Huntington's disease mutation. *Nat. Genet.* **15**, 197–200 (1997).
7. Gacy, A.M., Goellner, G., Juranic, N., Macura, S. & McMurray, C.T. Trinucleotide repeats that expand in human disease form hairpin structures in vitro. *Cell* **81**, 533–540 (1995).
8. Mitas, M. *et al.* Hairpin properties of single-stranded-DNA containing a GC-rich triplet repeat: $(CTG)_{15}$. *Nucleic Acids Res.* **23**, 1050–1059 (1995).
9. Petruska, J., Arnheim, N. & Goodman, M.F. Stability of intrastrand hairpin structures formed by the CAG/CTG class of DNA triplet repeats associated with neurological diseases. *Nucleic Acids Res.* **24**, 1992–1998 (1996).
10. Ohshima, K. & Wells, R.D. Hairpin formation during DNA synthesis primer realignment in vitro triplet repeat sequences from human hereditary disease genes. *J. Biol. Chem.* **272**, 16798–16806 (1997).
11. Freudenreich, C.H., Stavenhagen, J.B. & Zakian, V.A. Stability of a CTG/CAG trinucleotide repeat in yeast is dependent on its orientation in the genome. *Mol. Cell. Biol.* **17**, 2090–2098 (1997).
12. Nakatani, K., Sando, S. & Saito, I. Scanning of guanine-guanine mismatches in DNA by synthetic ligands using surface plasmon resonance. *Nat. Biotechnol.* **19**, 51–55 (2001).
13. Nakatani, K., Sando, S., Kumasawa, H., Kikuchi, J. & Saito, I. Recognition of guanine-guanine mismatches by the dimeric form of 2-amino-1,8-naphthyridine. *J. Am. Chem. Soc.* **123**, 12650–12657 (2001).
14. Hagihara, S. *et al.* Detection of guanine-adenine mismatches by surface plasmon resonance sensor carrying naphthyridine-azaquinolone hybrid on the surface. *Nucleic Acids Res.* **32**, 278–286 (2004).
15. Kobori, A., Horie, S., Suda, H., Saito, I. & Nakatani, K. The SPR sensor detecting cytosine-cytosine mismatches. *J. Am. Chem. Soc.* **126**, 557–562 (2004).
16. Syvänen, A-C. Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nat. Rev. Genet.* **2**, 930–942 (2001).
17. Nakatani, K. Chemistry challenge in SNP typing. *ChemBioChem* **5**, 1623–1633 (2004).
18. Han, X. & Gao, X. Sequence specific recognition of ligand-DNA complexes studied by NMR. *Curr. Med. Chem.* **8**, 551–581 (2001).
19. Wuthrich, K. NMR of proteins and nucleic acids. (John Wiley & Sons, Inc., New York, 1986).
20. Wijmenga, S.S. & van Buuren, B.N. The use of NMR methods for conformational studies of nucleic acids. *Prog. Nucl. Magn. Reson. Spectrosc.* **32**, 287–387 (1998).
21. Roberts, R.J. & Cheng, X. Base flipping. *Annu. Rev. Biochem.* **67**, 181–198 (1998).
22. Brockman, J.M., Nelson, B.P. & Corn, R.M. Surface plasmon resonance imaging measurements of ultrathin organic films. *Annu. Rev. Phys. Chem.* **51**, 41–63 (2000).
23. Smith, E.A. *et al.* Chemically induced hairpin formation in DNA monolayers. *J. Am. Chem. Soc.* **124**, 6810–6811 (2002).
24. Hashem, V.I. *et al.* Chemotherapeutic deletion of CTG repeats in lymphoblast cells from DM1 patients. *Nucleic Acids Res.* **32**, 6334–6346 (2004).
25. Zimmer, D.P. & Crothers, D.M. NMR of enzymatically synthesized uniformly $^{13}C^{15}N$-labeled DNA oligonucleotides. *Proc. Natl. Acad. Sci. USA* **92**, 3091–3095 (1995).
26. Liu, H., Spielmann, H.P., Ulyanov, N.B., Wemmer, D.E. & James, T.L. Interproton distance bounds from 2D NOE intensities: effect of experimental noise and peak integration errors. *J. Biomol. NMR* **6**, 390–402 (1995).
27. Brunger, A.T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
28. Kyo, M. *et al.* Evaluation of MafG interaction with Maf recognition element arrays by surface plasmon resonance imaging technique. *Genes Cells* **9**, 153–164 (2004).

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry
ELSEVIER
Bioorganic & Medicinal Chemistry 13 (2005) 4507–4512

N,N'-Bis(3-aminopropyl)-2,7-diamino-1,8-naphthyridine stabilized a single pyrimidine bulge in duplex DNA

Hitoshi Suda, Akio Kobori, Jinhua Zhang, Gosuke Hayashi and Kazuhiko Nakatani*

Department of Synthetic Chemistry and Biological Chemistry, Faculty of Engineering, Kyoto University, Kyoto 615-8510, Japan

Received 2 March 2005; revised 15 April 2005; accepted 15 April 2005
Available online 23 May 2005

Abstract—We here show the first identified ligand 2,7-diamino-1,8-naphthyridine (DANP) that strongly and specifically binds to the single cytosine and thymine bulges with exclusively 1:1 stoichiometry.

Small molecules that bind to bulges and mismatches in duplex DNA and those covalently attached to DNA are important probes for the dynamics of unusual DNA structures.[1,2] With a combination of bulge- and mismatch-forming hybridization, these molecules are also useful for the detection of base mutation and deletion.[3,4] Toward this end, we have investigated a series of compounds that have hydrogen-bonding surface fully complementary to that of the unpaired or partially paired nucleotide bases, and found that 2-amino-1,8-naphthyridine and 8-azaquinolone effectively functioned as a molecular element for the recognition of guanine and adenine bases, respectively.[5,6] For the cytosine, Teramae and co-workers used 2-amino-1,8-naphthyridine and proposed formation of two hydrogen bonds to the cytosine.[7] We have independently reported that a dimeric form of 2-amino-1,8-naphthyridine strongly stabilized C–C mismatches.[8] Because these ligands bound to the bulge and mismatched structure in duplex DNA, the equilibrium between single strands and a duplex shifted toward the duplex state, making the apparent stability of the duplex increase. In the study, we proposed a formation of three hydrogen bonds between a protonated form of 2-amino-1,8-naphthyridine and the cytosine. As protonation of the nitrogen in the heterocycles effectively modulated the hydrogen-bonding surface of the molecule, further studies were carried out to see the scope and limitation of the idea. We here describe a new molecular probe *N,N'*-bis(3-aminopropyl)-2,7-diamino-1,8-naphthyridine (DANP) that could stabilize not only a single cytosine but also the thymine bulge in duplex DNA (Chart 1). Cold spray ionization time-of-flight mass spectrometry showed that DANP bound to the cytosine and thymine bulges with a 1:1 stoichiometry. The pH dependency of UV spectra of 2,7-diamino-1,8-naphthyridine supported the hydrogen bonding of DANP to the cytosine and thymine through the protonated form.

We have focused on the 2,7-diamino-1,8-naphthyridine, because it has an alignment of hydrogen-bonding groups in the order of a donor (D), acceptor (A), acceptor, and donor. We anticipated that rich hydrogen-bonding groups in one edge of a planar aromatic ring provide a good chance to form stable complex with bulged nucleotides in duplex DNA. DANP was obtained by nucleophilic substitution of 2,7-dichloro-1,8-naphthyridine[9] with 1,3-propandiamine (Scheme 1). Since the initially formed DANP was difficult to isolate from 1,3-propandiamine used as a solvent, the crude

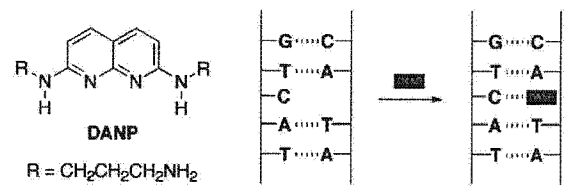


Chart 1. Structure of DANP and an illustration of DANP mediated stabilization of C-bulge.

Keywords: Cytosine; Bulge; CSI-TOF; DNA.
* Corresponding author at present address: Department of Regulatory Bioorganic Chemistry, The Institute of Scientific and Industrial Research, Osaka University, Ibaraki 567-0047, Japan. Tel.: +81 668 798 455; fax: +81 668 798 459; e-mail: nakatani@sanken.osaka-u.ac.jp

0968-0896/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2005.04.035