Doctoral Dissertation

Development and Brain Imaging of Repeat DNA-binding Molecules

Eitaro Murakami

Department of Chemistry

Graduate School of Science

Osaka University

Preface

The research described in this dissertation was conducted under the instruction of Professor Kazuhiko Nakatani at the Department of Chemistry in Osaka University from April, 2017 to March, 2022. The study concerned about development of new small molecule binding to GGCCCC repeat DNA related to ALS/FTD and brain imaging of binding molecules binding to CAG repeat DNA. All of the results presented in this dissertation in this paper were achieved by the author himself.

Index

Abstract.....- 4 -

General Introduction.....- 7 -

Chapter 1 Development of GGCCCC Repeat DNA Binding Ligand- 16 -

Chapter 1-1 Searching for Binding Information and New Binding Molecules to G2C4 Repeat.....- 20 -

Chapter 1-2 Redesign of Binding Molecules Based on Binding Information and Evaluation of Binding to G2C4/G2C4 DNA.....- 49 -

Chapter 1-3 Effects of G2C4 Repeat Binding Molecule on the transcription of G4C2/G2C4 Repeat .....- 71 -

Chapter 2 Brain Profiling of CAG Repeat DNA Binding Ligand.....- 80 -

Chapter 2-1 Construction of Protocol for Brain Profiling.....- 82 -

Chapter 2-2 Intracerebroventricular and Intranasal Administration .....- 99 -

Chapter 2-3 Detection of NA and NA-like Ligand in Brain using MS Imaging and PET .....- 119 -

List of Publications.....- 132 -

List of Presentations .....- 133 -

Acknowledgement .....- 135 -

Abstract

Chapter 1

In 2011, it was reported that the most frequent cause of amyotrophic lateral sclerosis with frontotemporal dementia (ALS/FTD) was the aberrant expansion of GGGGCC (G4C2) repeat in the first intron of the *C9orf72* gene. The mechanism of G4C2 repeat-mediated pathogenesis and the molecular mechanism of the abnormally expanded repeat sequence are fundamental problems common to all repeat diseases and have not been unravel yet. G4C2 repeat is known to be able to form the G-quadruplex and many molecules that bind to this structure have been reported, whereas there are few studies on small molecules binding to G2C4 repeat, the complementary sequence of G4C2 repeat.

We reported DQG as a small molecule that bound to G2C4/G2C4 motif DNA and G2C4 repeat DNA. In this thesis, *N,N'*-diheteroaryl guanidine derivatives which were expected to hydrogen bond with cytosine were designed to clarify the binding mode in the DQG-G2C4 complex. Comparing the binding evaluation using DQG and DQG analogs, D3iQG, D1iQG, and Q3iQG, led the structural factors binding to G2C4/G2C4.

Based on the extracted binding factors, 1,3-di(quinazolin-2-yl)guanidine (DQzG) was designed. The binding affinity of DQzG for G2C4/G2C4 was higher than that of DQG and other DQG analogs. NMR analysis of the DQzG-G2C4/G2C4 complex suggested that DQzG could form hydrogen bonds with cytosine.

G4C2/G2C4 repeat on *in vitro* transcription with DQzG was performed to investigate the biological effects of the G2C4 repeat-binding molecule on transcription. It showed that the

amount of RNA produced by the transcription reaction was not changed in the presence or absence of DQzG. It suggested that DQzG could not affect the transcription of the G4C2/G2C4 repeats, although further investigation was necessary.

Chapter 2

Huntington's disease (HD) is caused by the aberrant expansion of CAG repeats in the *HTT* gene, which results in the loss of neurons in the striatum of the brain. The mechanism of the aberrant expansion was proposed as a model which the CAG/CAG hairpin structure formed by the CAG repeats was incompletely repaired by mismatch repair enzymes. The small molecule, naphthyridine-azaquinolone (NA) was shown to cooperatively bind to CAG/CAG motif DNA with two molecules forming the hydrogen bonds with naphthyridine-guanine and azaquinolone-adenine and flipping two cytosine bases out of the duplex. In 2020, it was reported that the administration of NA directly to the brain striatum of mouse model for HD shortened CAG repeats.

Since NA is unable to penetrate blood-brain barrier (BBB), it was administered directly into the striatum of HD model mice. However, this administration is highly invasive and not suitable for long-term administration. We could not confirm the therapeutic effects of NA, such as improvement of motor functions. In case of intranasal administration, which is a delivery method to the brain without penetrating via the BBB, it has been a challenge to develop a technique to analyze the drug distribution in the brain. Therefore, to solve the problems of efficient delivery of small molecules to the brain and analysis of their distribution in the brain, the intranasal administration-tissue transparency whole brain imaging system using fluorescently labeled small

molecules was developed, combining intranasal administration with tissue clearing/3D imaging technology.

The designed fluorescence labeling NA, NA-Alexa594 was administrated directly into the brain parenchyma by stereotaxic injection to clarify whether NA-Alexa594 was detectable by tissue clearing/whole-brain imaging. Light sheet fluorescence microscopy (LSFM) revealed that NA-Alexa594 was detected in the cleared mouse brain and can provide the distribution of NA-Alexa594 in 3D whole brain.

Furthermore, intranasal delivery of NA-Alexa594 to the brain was tested using the designed protocol. Immediately after administration, NA-Alexa594 was present on blood vessels and olfactory bulb. On the other hand, one day after administration, detected in the olfactory bulb. These distribution of NA-Alexa594 in the brain indicated that intranasal delivery to the brain was possible. The temporal changes from immediately to 6 h after intranasal administration revealed the existence of two or more delivery pathways in Nose-to-Brain.

It was necessary to detect the delivery of unmodified NA to the brain by intranasal administration. MS imaging was determined for the purpose of detecting unmodified NA. Although the administration directly to the brain parenchyma was detectable using MS imaging, the delivery to the brain by using intranasal administration was not.

General Introduction

Repeat Disease

In the human genome, there are repeat sequences that are repeated multiple times with a few bases as a unit. These repeated sequences vary from person to person and have been used to identify individuals. In repeat sequences, there are the repeat sequences that consist of a set of 2-10 bases as a unit. When these sequences generated abnormally and exceed a certain threshold, serious neurodegenerative diseases are affected, known as repeat disease. In 1991, it was reported to discover abnormal CGG repeat expansion in the 5′-untranslated region (UTR) of the Fragile X Mental Retardation 1 (FMR1) gene in Fragile X Syndrome (FXS)[1] and abnormal CAG repeat expansion in the coding sequence of X-linked spinal cord in spinal and bulbar muscular atrophy (Kennedy's disease).[2] In the following year, abnormal extension of CTG repeat in the 3′-UTR of the muscular dystrophy protein kinase (DMPK) reportedly caused muscular dystrophy type 1 (DM1).[3] Aberrantly expanded repeat sequences were identified in coding and non-coding regions of the gene, including exons, introns, 5′-UTR, and 3′-UTR. At present, more than 40 diseases caused by the extension of specific repetitive DNA microsatellite sequences beyond a critical threshold repeat length have been reported.[4] (Table. 1) Longer repeat expansions tend to be associated with more severe and earlier-onset diseases; in case of CAG/polyglutamine, there was a strong inverse correlation between age of onset and repeat length.[5] Expanded repeat DNA is known to often change in size as they are passed on to the next generation.

Table. 1 Unstable repeat tracts and human disease and the induced neurological disorders

Disorder	Location	Repeat	Normal	Disease
Huntington's disease (HD)	Exon	CAG	10 - 34	35 <
Myotonic dystrophy type 1 (DM1)	3′-UTR	CTG	5 - 37	90 - 6500
Fragile X tremor/ataxia syndrome (FXTAS)	5′-UTR	CGG	6 - 52	59 - 230
Fragile X syndrome (FXS)				230 - 2000
Spinocerebellar ataxia 31 (SCA31)	Intron	TGGAA	0	500 - 760
Amyotrophic lateral sclerosis, Frontotemporal lobar degeneration (ALS/FTD)	Intron	GGGGCC	20	250 <

Pathology

It was suggested that the aberrant expanded repeat sequence can cause neurodegeneration from a variety of causes.[6,7] (Fig. 1) For example, the aberrant expansion of the repeat sequence in genome led haploinsufficiency, decrease in the expression of proteins that the genome naturally expresses, leading to neurological disease. Also, it was transcribed to produce repeat RNA and was bound by RNA-binding proteins, resulting in loss of function of the protein or abnormal aggregation of the generated cytotoxic complex. In addition, repeat-associated non-AUG

translation (RAN translation), a phenomenon in which repeat regions are translated regardless of coding or non-coding regions, has been reported. The translated and generated repeat peptides aggregated and become cytotoxic. The functional proteins lost that function or acquired toxicity when they contain repeat region. These and many other factors reportedly led to neurodegeneration and diseases.

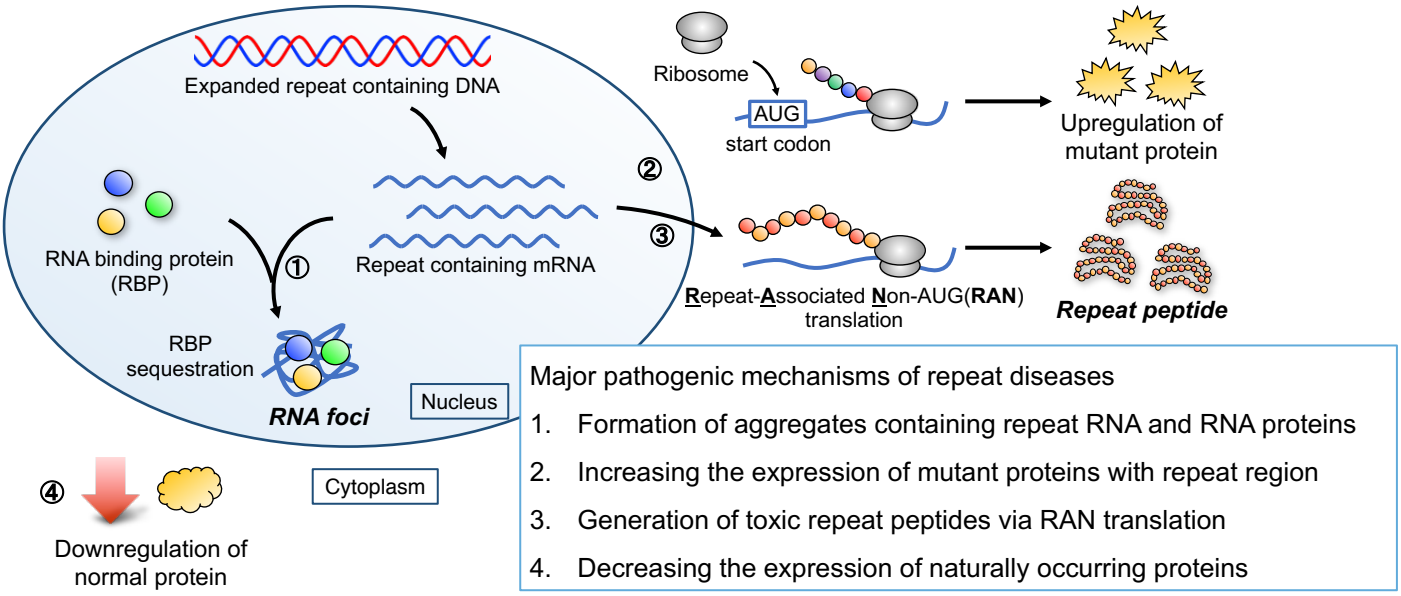


Fig. 1 Illustration of major pathogenic mechanisms in repeat diseases.

Therapeutic Approach to Repeat Disease

Many studies have been addressed as therapeutic strategies for diseases caused by expanded repeat sequence. Here were some of the reported therapeutic strategies.

Approaches using antisense oligonucleotides (ASOs) or siRNA have proven effective in downregulating the expression of the expanded allele in HD and clinical trials were underway.[8,9] Also, in C9-ALS, the ASOs binding to *C9orf72* transcript or repeat expansion reduced nuclear foci containing GGGGCC, reducing RAN translated products without altering overall *C9orf72* RNA expression levels.[10,11]

CRISPR-based gene-editing therapies have also been reported. It suggested that CRISPR/Cas9-based gene-editing has the potential to efficiently and permanently eliminate neuronal toxicity via 140 poly-Q repeat expansion in the adult brain.[12] In addition, the features in DM1, DM2, C9-ALS, and polyglutamine disease got better, using RNA target Cas9 (RCas9) removing RNA forks, reducing polyglutamine protein products, re-localizing repeat-binding proteins similar to healthy controls, and effective reversing abnormal splicing.[13]

Therapeutic studies using small molecules targeted the aberrant conformation of poly-Q protein encoded by CAG repeat and showed that arginine inhibited aberrant aggregation in animal models.[14] Small molecules binding to G4C2 repeat RNA has been reported to become biomarkers and inhibit RAN translation in C9-ALS.[15–17]

High Order Structure of Repeat Sequence

Expanded repeat DNA and RNA have been reported to form high order structures.[4,6,7,18] It is known to form due to the aberrant R-loop formed by the transcription in repeat expansion area. It suggested that the longer the repeat, the easier forming higher-order structures in repeat DNA and protein recruitment in repeat RNA. It reported that there were various higher-order structures formed by repeat DNA/RNA, such as hairpin structures including mismatches, G-quadruplex in G-rich sequence, i-motif in C-rich sequence, and so on. Small molecules that can bind to the higher-order structures formed by these expanded repeat sequences can lead to biological changes.

Binding Molecules to Repeat DNA/RNA

Our research group has developed many molecules that can bind to the hairpin structure, which is a metastable structure containing mismatched base pairs, one of the higher-order structures formed by repeat sequences. We dedicated on the formation of stable complexes with DNA or RNA repeats and small molecules to change their biological functions. In previous reports, naphthyridine carbamate dimer (NCD) was synthesized that bound to 5′-CGG-3′/5′-CGG-3′. Two NCD moieties recognized four guanine bases of CGG/CGG by hydrogen bonds.[19] In addition, naphthyridine-azaquinolone (NA) that bound to 5′-CAG-3′/5′-CAG-3′ was developed and the azaquinolone moiety was able to recognize adenine. Furthermore, structural analysis by NMR suggested the NA-CAG/CAG complex structure with cytosine flipped out of the duplex.[20] Recent studies reported 2,9-diaminoalkyl-1,10-phenanthroline (DAP) and PPDIQ (JM608) as candidates

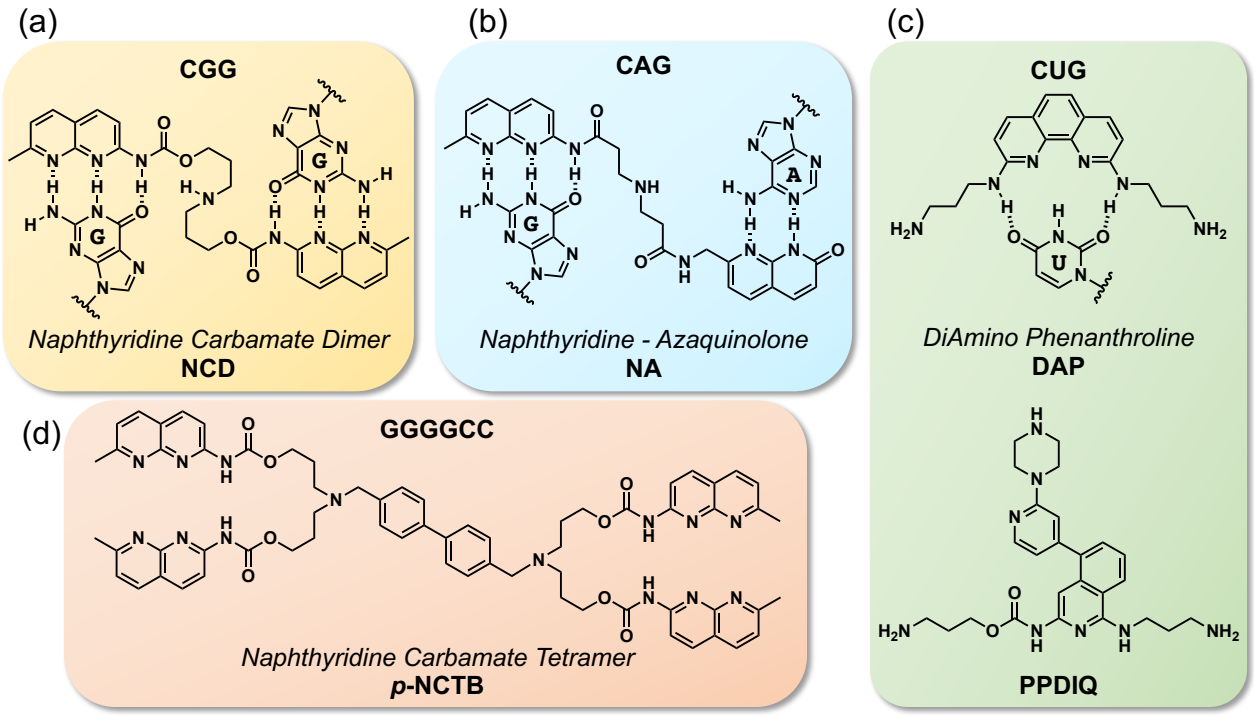


Fig. 2 Chemical structure of molecules binding to repeat DNA/RNA

for CUG binding.[21,22] The dimeric derivatives of DAP (JM608), DDAP (JM642) interfered with the sequestration of MBNL1 protein by RNA CUG repeat, a pathological feature of DM1 caused by CUG repeat, and showed promising binding properties to RNA CUG repeat. In 2020, our group also revealed that NCD bound to UGGAA repeats causing SCA31; molecular binding of NCD to UGGAA/UGGAA and, surprisingly, NMR complex analysis showed that adenine was flip-out.[23] Furthermore, when SCA31 model drosophila with expanded UGGAA repeats were fed NCDs from the embryonic stage, the compound eyes in absence of NCD were degenerated, while in presence of NCD become closer to normal compound eyes.

Reference

1. Verkerk, A. J. M. H. *et al.* Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65**, 905–914 (1991).
2. Fischbeck, K. H., Spada, A. R. La, Wilson, E. M., Lubahn, D. B. & Harding, A. E. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **352**, 77–79 (1991).
3. Brook, J. D. *et al.* Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family member. *Cell* **68**, 799–808 (1992).
4. Pearson, C. E., Edamura, K. N. & Cleary, J. D. Repeat instability: Mechanisms of dynamic mutations. *Nat. Rev. Genet.* **6**, 729–742 (2005).
5. Langbehn, D. R. *et al.* CAG-repeat length and the age of onset in Huntington Disease (HD): A review and validation study of statistical approaches. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **153**, 397–408 (2010).
6. Malik, I., Kelley, C. P., Wang, E. T. & Todd, P. K. Molecular mechanisms underlying nucleotide repeat expansion disorders. *Nat. Rev. Mol. Cell Biol.* **22**, 589–607 (2021).
7. La Spada, A. R. & Taylor, J. P. Repeat expansion disease: Progress and puzzles in disease pathogenesis. *Nat. Rev. Genet.* **11**, 247–258 (2010).
8. Wild, E. J. & Tabrizi, S. J. Therapies targeting DNA and RNA in Huntington's disease. *Lancet Neurol.* **16**, 837–847 (2017).

9. Estevez-Fraga, C., Flower, M. D. & Tabrizi, S. J. Therapeutic strategies for Huntington's disease. *Curr. Opin. Neurol.* **33**, 508–518 (2020).

10. Donnelly, C. J. *et al.* RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. *Neuron* **80**, 415–428 (2013).

11. Lagier-Tourenne, C. *et al.* Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc. Natl. Acad. Sci. U. S. A.* **110**, (2013).

12. Yang, S. *et al.* CRISPR / Cas9-mediated gene editing ameliorates neurotoxicity in mouse model of Huntington ' s disease Find the latest version : CRISPR / Cas9-mediated gene editing ameliorates neurotoxicity in mouse model of Huntington ' s disease. *J. Clin. Invest.* **127**, 2719–2724 (2017).

13. Batra, R. *et al.* Elimination of Toxic Microsatellite Repeat Expansion RNA by RNA-Targeting Cas9. *Cell* **170**, 899-912.e10 (2017).

14. Minakawa, E. N. *et al.* Arginine is a disease modifier for polyQ disease models that stabilizes polyQ protein conformation. *Brain* **143**, 1811–1825 (2020).

15. Su, Z. *et al.* Discovery of a Biomarker and Lead Small Molecules to Target r(GGGGCC)-Associated Defects in c9FTD/ALS. *Neuron* **83**, 1043–1050 (2014).

16. Wang, Z. F. *et al.* The Hairpin Form of r(G 4 C 2) exp in c9ALS/FTD Is Repeat-Associated Non-ATG Translated and a Target for Bioactive Small Molecules. *Cell Chem. Biol.* **26**, 179-190.e12 (2019).

17. Ursu, A. *et al.* Structural Features of Small Molecules Targeting the RNA Repeat Expansion That Causes Genetically Defined ALS/FTD. *ACS Chem. Biol.* **15**, 3112–3123 (2020).

18. Castel, A. L., Cleary, J. D. & Pearson, C. E. Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat. Rev. Mol. Cell Biol.* **11**, 165–170 (2010).

19. Peng, T., Murase, T., Goto, Y., Kobori, A. & Nakatani, K. A new ligand binding to G-G mismatch having improved thermal and alkaline stability. *Bioorganic Med. Chem. Lett.* **15**, 259–262 (2005).

20. Nakatani, K. *et al.* Small-Molecule Ligand Induces Nucleotide Flipping in (Cag)n Trinucleotide Repeats. *Nat. Chem. Biol.* **1**, 39–43 (2005).

21. Li, J. *et al.* A Dimeric 2,9-Diamino-1,10-phenanthroline Derivative Improves Alternative Splicing in Myotonic Dystrophy Type 1 Cell and Mouse Models. *Chem. - A Eur. J.* **24**, 18115–18122 (2018).

22. Matsumoto, J. *et al.* The Dimeric Form of 1,3-Diaminoisoquinoline Derivative Rescued the Mis-splicing of Atp2a1 and Clcn1 Genes in Myotonic Dystrophy Type 1 Mouse Model. *Chem. - A Eur. J.* **26**, 14305–14309 (2020).

23. Shibata, T. *et al.* Small molecule targeting r(UGGAA)n disrupts RNA foci and alleviates disease phenotype in Drosophila model. *Nat. Commun.* **12**, 1–13 (2021).

Chapter 1

Development of

GGCCCC Repeat DNA Binding Ligand

Introduction

・ALS/FTD G4C2 Repeat

In 2011, it was reported that aberrant expansion of GGGGCC (G4C2) repeat in the first intron flanked by exons 1a and 1b in the *C9orf72* gene was the most frequent cause of amyotrophic lateral sclerosis with frontotemporal dementia (ALS/FTD).[1,2] It has been found in more than 40% of patients with familial and more than 8% of patients with sporadic ALS/FTD patients. The number of G4C2 repeats has been found to be expanded to more than 250 in ALS/FTD patients, while it is less than 20 in non-affected individuals.[3,4] The pathogenic mechanism of the aberrantly expanded G4C2 repeat sequence was the formation of abnormal aggregates of RNA-binding proteins trapped by the repeat RNA generated after transcription (RNA foci)[5,6] and the formation of abnormal peptide repeat protein aggregates by repeat-associated non-AUG translation (RAN translation), a phenomenon in which repeat regions are translated regardless of whether they are coding or non-coding regions.[7] (Fig. 1) RNA foci induced loss of function in splicing and nuclear-cytoplasmic transport by trapping various RNA-binding proteins. A pathological mechanism was also reported that the expression of the protein which was supposed to be expressed was repressed due to the inclusion of expanded repeat DNA (Haploinsufficiency).[8] However, the molecular

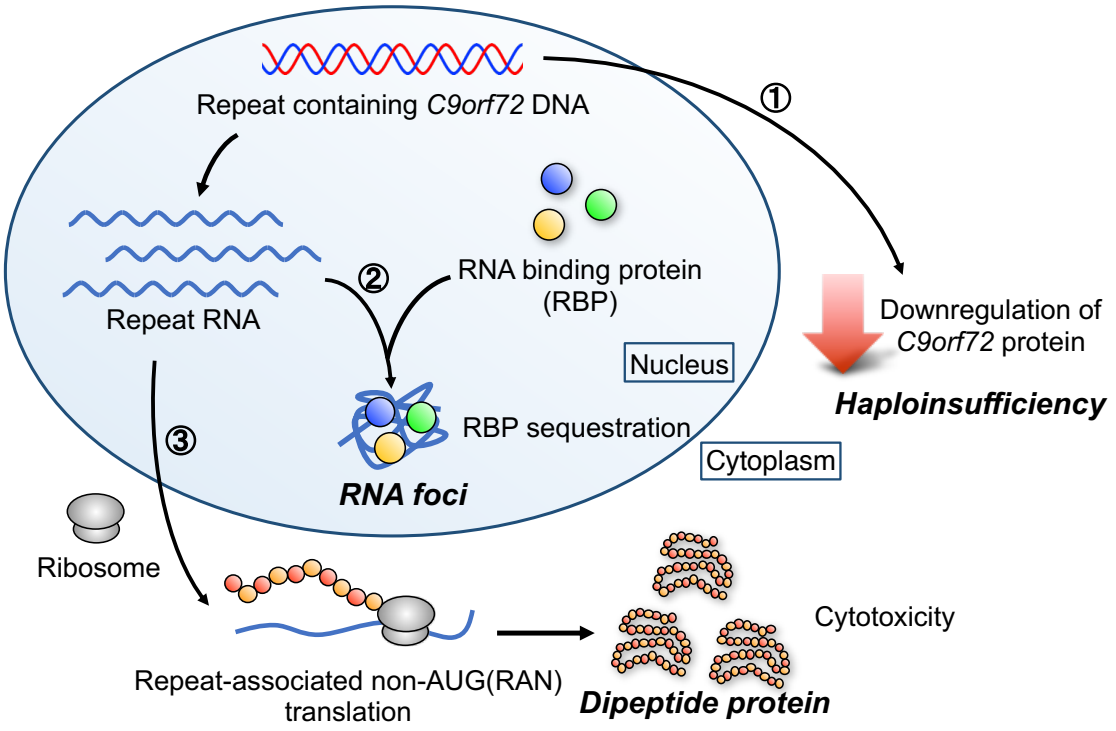


Fig. 1 The potential mechanisms of neurodegeneration in C9-ALS/FTD.

mechanism of aberrant expansion, which is the underlying cause of repeat disease, remained unclear.

There were several descriptions of the methylation status of G4C2/G2C4 repeats. Previously encoded G4C2 repeats have been examined. Expansion-specific hypermethylation of 5′CpG-island near G4C2 repeat region was identified in some of the carriers.[9] Subsequently, methylation of G4C2 repeat itself was detected in blood, brain, and spinal cord tissues of the same individuals.[10] High methylation levels were suggested to correlate with G4C2 repeat expansion and reduction of total transcript levels of *C9orf72*.[11] On the other hand, *in vitro* experiments with methylated G2C4 repeat, the antisense of G4C2 repeat reported the formation of G-C-G-C-tetrads and C-C-C-C-tetrads.[12]

In addition to the B-type double helix that forms in normal DNA, other higher-order structures such as hairpins, triple-stranded and quadruple-stranded structures are known to increase repeat instability.[13–15] It suggested that repeat instability was an important factor in the replication, recombination, transcription and DNA repair processing.

• High-order Structure of G4C2/G2C4 Repeat

The abnormally expanded G4C2 repeats can form hairpin structures, G-quadruplex structures, and iCD-DNA, in addition to the regular B-form duplexes.[16–21] It suggested that the higher-order structures formed by aberrantly expanded G2C4 repeats can have multiple conformations. In 2018, the DNA sequence containing G2C4/G2C4 motif with two consecutive C-C mismatches was reported to from an extrahelical structure (e-motif) in molecular dynamics calculations.[17,22] (Fig.

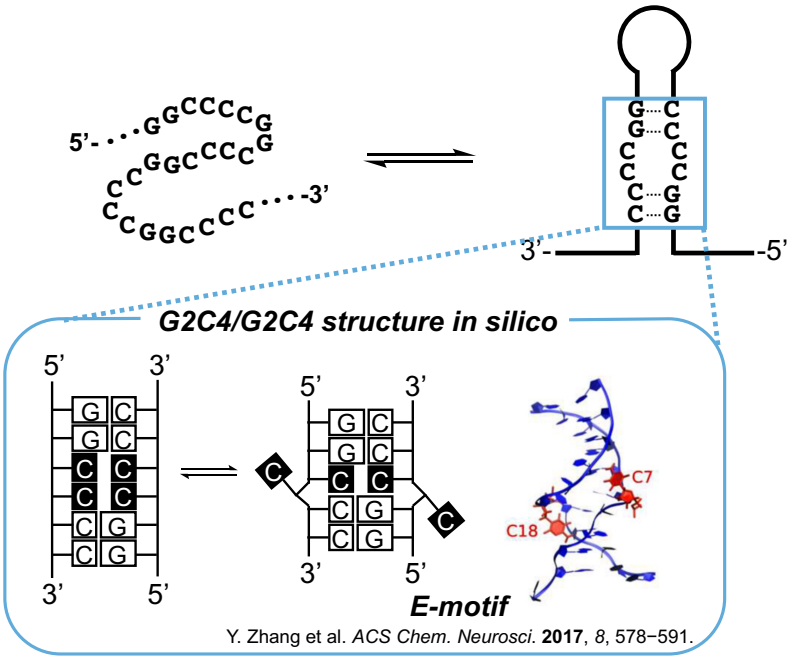


Fig. 2 Potential structure, E-motif in G2C4 repeat DNA

2) E-motif structure had a pair of C-C mismatch flipped out to the minor groove side outside DNA duplex.

These non-canonical structures in G4C2/G2C4 repeat are potential targets for small molecules.

・G4C2/G2C4 Repeat Binding Molecules

Small molecules that target the higher-order structure of the extended repeat sequence are expected to be useful tools for studying and further controlling repeat instability. Guanine quadruplex structure has been reported as a higher-order structure formed by the G4C2 repeat. (Fig. 3) On the other hand, there were reportedly the molecule binding to G4C2 repeat RNA obtained by screening that can bind to G2C4 repeat, the complementary strand of G4C2 repeat.[23] In addition, it was reported small molecules could stabilize i-motif structure that G2C4 repeats may form, without G2C4 repeat sequence, as well as antibodies could bind to.[24–27] In this study, I aimed to develop the molecule that could bind to G2C4 repeat DNA and induce the formation of higher-order structures.

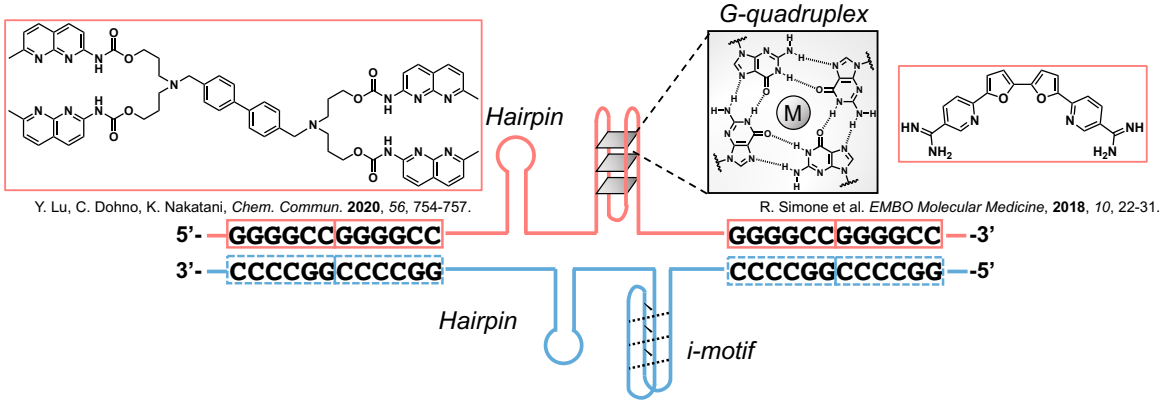


Fig. 3 High-order structures of G4C2/G2C4 repeat DNA and binding molecules to its structures.

Chapter 1-1

Searching for Binding Information and New Binding Molecules to G2C4 Repeat

Molecular Design

・G2C4/G2C4 Binding Molecule – DQG

The binding molecule was designed, taking into account in the higher-order structure formed by G2C4 repeat DNA. The unstable DNA duplex containing G2C4/G2C4 motif shown in the previous section was a potential target binding site for small molecules. Therefore, we designed a molecule, 1,3-di(quinolin-2-yl)guanidine (DQG), which consisted of two quinoline rings linked by a guanidine.[28] (Fig. 4) The positively charged nitrogen atom at position 2 of the guanidyl group was assumed to stabilize the molecule by conformational selection in which multiple planar structures are formed by intramolecular hydrogen bonding with the nitrogen atom of quinoline and an appropriate conformation is selected for binding to G2C4/G2C4 motif. The planer conformations were the structures with two hydrogens in the guanidyl group as two hydrogen bond donors, which was expected to bind to cytosine, two hydrogen bond acceptors.

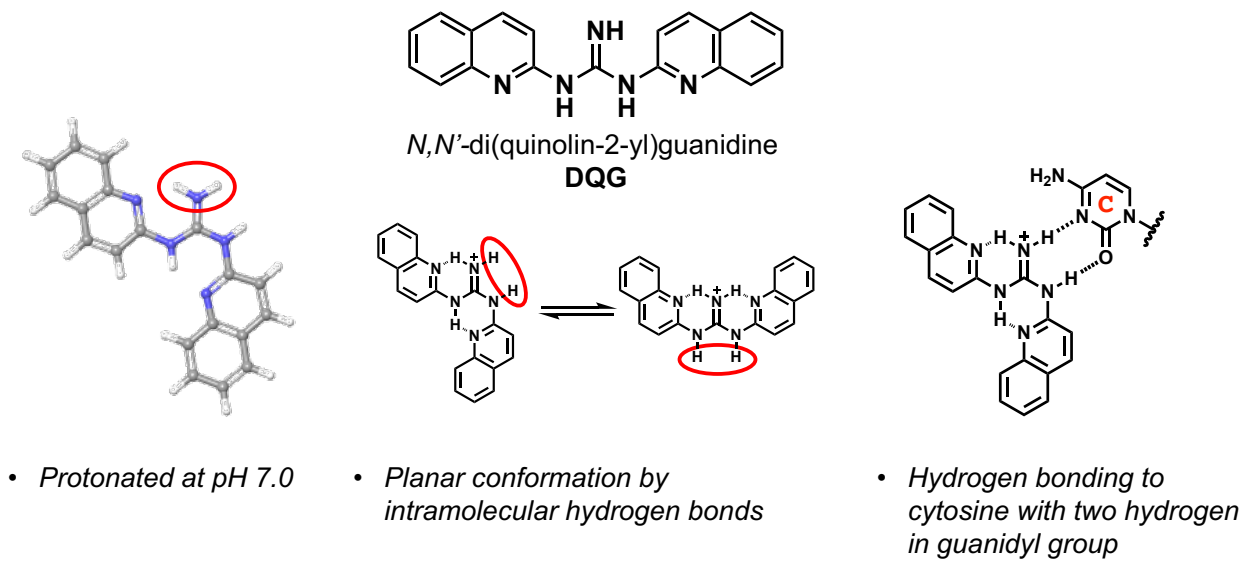


Fig. 4 Proposed schematic illustrations of hydrogen bonding patterns with DQG and cytosine.

The designed molecule was found to bind to G2C4/G2C4 by binding evaluation such as T_m, CD, and CSI-MS.[28]

The melting temperatures (T_m) of DNA double strands containing GGXXCC/GGXXCC motifs (X = A, T, G or C) were measured in the presence and absence of the compounds. In addition, 5′-GCATGGCCCCTACG-3′/5′-CGTAGGGGCCATGC-3′ was used as the full complementary strand DNA without mismatch. The results of melting temperature measurements showed that when DQG and DNA containing G2C4/G2C4 motifs were used, the ΔT_m increased by 19.6 ºC, suggesting that the ligand bound to DNA containing G2C4/G2C4 motifs. (Fig. 5-a) This increase in T_m was not pronounced when other motifs (X = A, T, G) and full-match DNA were used as compared to G2C4/G2C4 motif DNA, indicating that DQG specifically stabilized the G2C4/G2C4 motif. DQG did not stabilize the unstable structure in non-specific binding. Circular dichroism (CD) results suggested that DQG and G2C4 motif DNA can form stable complexes. And cold-spray ionization mass spectrometry (CSI-MS) measurements showed the stoichiometric ratio DNA:DQG =1:1 and 1:2 in the complex with DQG and G2C4/G2C4 motif DNA. Therefore, it suggested that one or two DQG molecules could bind to the G2C4/G2C4 motif DNA without binding to the full match.

In addition, surface plasmon resonance (SPR) measurements evaluated the binding affinity of DQG for G2C4 repeat. In SPR assay, the binding of the ligand to (G2C4)9 repeat sequence was evaluated using single-cycle kinetics analysis. It was measured at the concentrations scale (1.25, 2.5, 5.0, 10.0, 20.0 μM). (Fig. 5-b) This binding assay suggested that DQG could bind to G2C4 repeat.

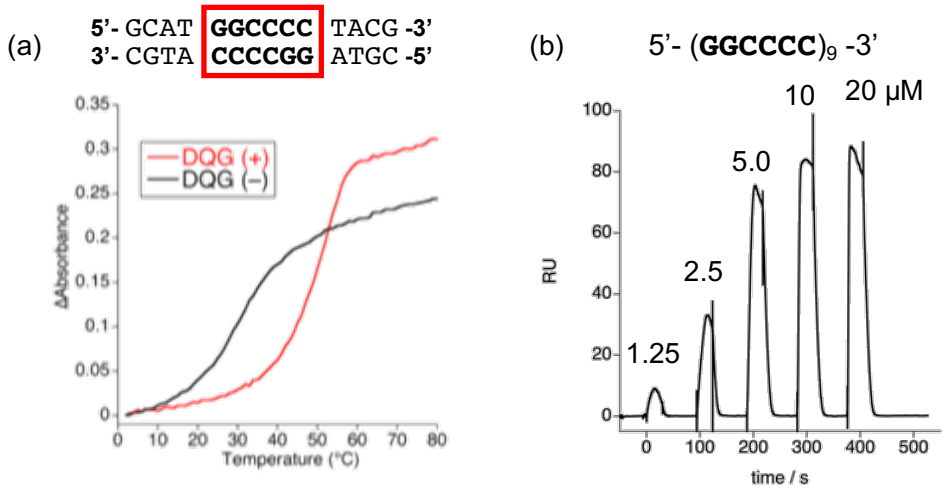


Fig. 5 (a) Thermal melting profiles of DNA duplexes in the presence (black) and absence (red) of ligand. (b) SPR data for ligands binding to surface-immobilized d(G2C4)9. The ligands were sequentially added at 1.25, 2.5, 5.0, 10.0 and 20.0 μM.

• DQG Derivatives

The binding information with the T-substituted sequences of G2C4/G2C4 motif DNA were evaluated to identify the sites recognized by DQG.[28] On the other hand, the validity of the molecular design and the binding mode have not been elucidated yet. Therefore, in order to validate the concept of G2C4/G2C4 binding molecules, the control compounds and DQG analogs with similar properties to DQG were designed and their binding to DNA G2C4/G2C4 were evaluated compared with DQG binding. This evaluation with the DQG analogs also was intended to search for the molecules with more binding affinity than DQG.

First, the control compounds, 1-(quinolin-2-yl)guanidine (QG), *N,N'*-di(naphthalen-2-yl)guanidine (DNpG), and *N,N'*-di(quinolin-2-yl)urea (DQU) were designed. QG has one

quinoline ring and one guanidine.[28] (Fig. 6) DNpG has two naphthalene rings and one guanidine, the nitrogen in rings required for intramolecular hydrogen bonding. The hydrogen in the guanidine was necessary for DQG putative binding to cytosine and the nitrogen in the quinoline involved in intramolecular hydrogen bonds were to confirm the necessity of binding to G2C4/G2C4. DQU was also added to obtain the binding information. It has two quinoline rings and one urea moiety to examine the necessity of intramolecular hydrogen bonding for binding to DNA G2C4/G2C4.

Next, DQG analogs, *N,N'*-di(isoquinolin-3-yl)guanidine (D3iQG) and *N,N'*-di(isoquinolin-1-yl)guanidine (D1iQG) were designed. DQG and DQG analogs differ in the position of the benzene in the two aromatic rings. It was assumed that the difference in the position of the benzene could reveal the optimal structure of the ligand when they bind to cytosine. If DQG analogs could have more optimal structures than DQG, they would be molecules with higher binding than DQG.

The possibility of forming planar structures in DQG and DQG analogs was verified with computational methods and molecular calculations. The first step was to perform conformational search in molecular modeling, Maestro using the molecular mechanics (MM) method to find stable conformations that could be formed without preconceptions by covering planar and non-planar conformations. In this study, a computational chemical analysis of DQG and DQG analogs (D1iQG and D3iQG) was performed to clarify how the stability differences among the various configurations of DQG analogs differ among the compounds and whether these differences affect the binding to the target nucleic acids. The force field, OPLS4 was used to search for configurations in aqueous solution. The MM calculation generated 10,000 conformations and their conformations with energies within 5 kcal/mol from the most stable conformation were

extracted. As a result, the energy of DQG; 21, D3iQG; 19, and D1iQG; 32 were found. The calculation obtained three types of conformations; intramolecular hydrogen bonds and formed planar structures, and one intramolecular hydrogen without planar structures.

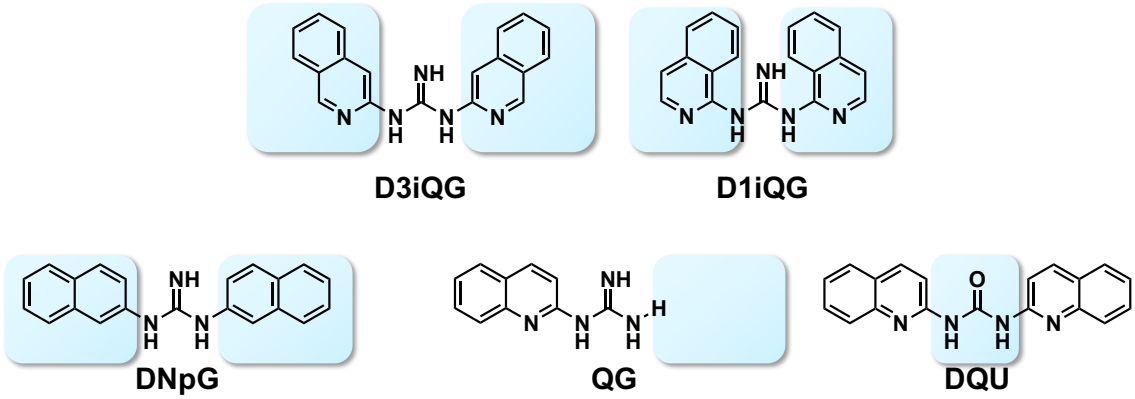


Fig. 6 Chemical structures of DQG derivatives.

Next, quantum mechanics (QM) calculations using all the obtained configurations were performed and compared the energies. The method, DFT wb97xd/6-31G^{+} (d, p) was used to calculate the optimization of ligands using Gaussian 09. The results showed that for DQG and D3iQG, within 3 kcal/mol of the restable conformation, there were only planar conformations,

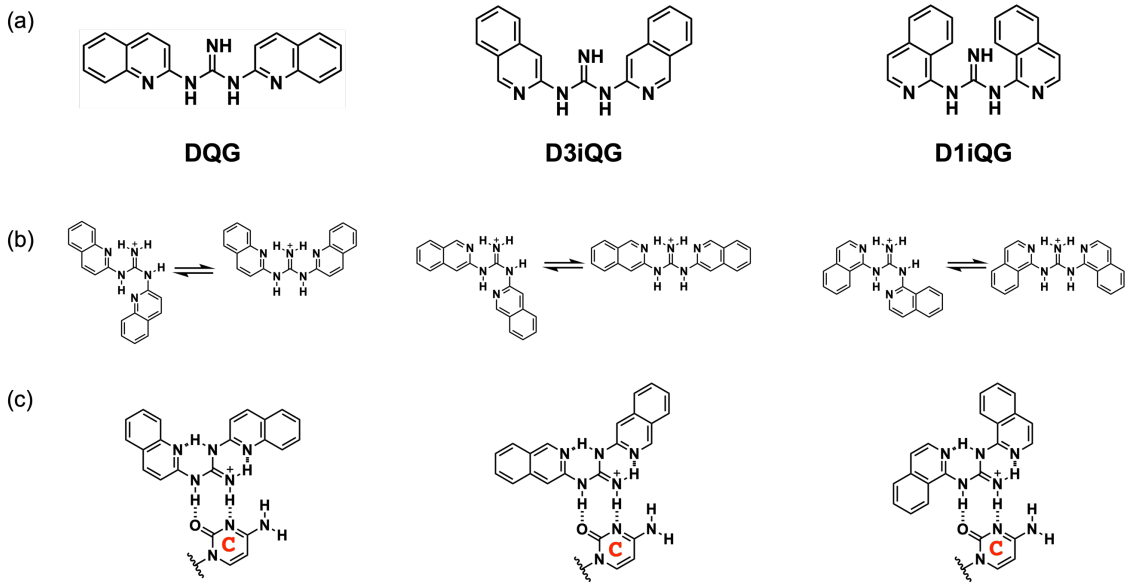


Fig. 7 The features of DQG, D3iQG, and D1iQG (a) chemical structures, (b) planer conformations, and (c) the model of hydrogen bonds with cytosine.

plus all possible planar conformations. The other conformations were at energies higher than 7 kcal/mol from the restable conformation. In other words, DQG, D3iQG and D1iQG significantly can form planar configurations, as revealed by computational methods. The most stable conformation obtained by molecular calculations for each compound is shown in Fig. 8.

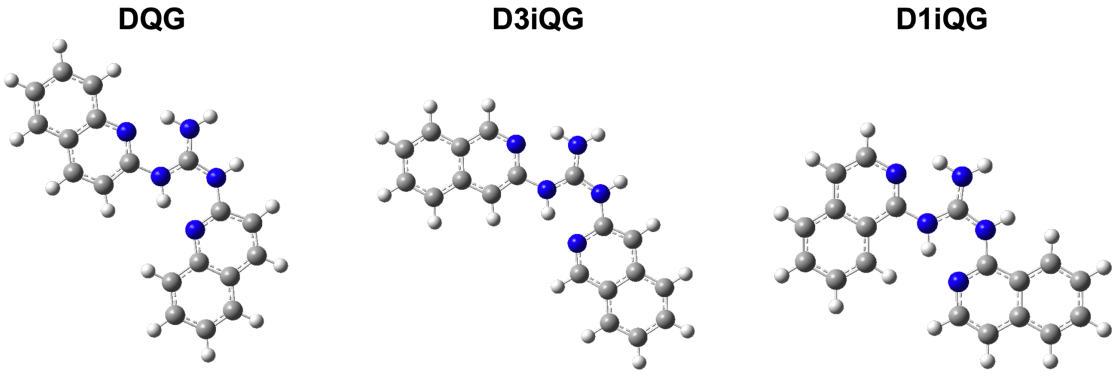


Fig. 8 The most stable conformation obtained by molecular calculations in DQG, D3iQG, and D1iQG.

The designed compounds were synthesized. D3iQG, D1iQG, and DNpG were synthesized by using purchasable 3-chloroisoquinoline, 1-chloroquinoline, and 2-bromonaphthalene as starting materials, coupled with guanidine reagent by Buchwald-Hartwig reaction in one step, and purified by HPLC. (Material & Method)

Results & Discussion

• Melting Temperature

T_m of the DNA double strand containing GGCCCC/GGCCCC motif was measured in the presence and absence of the compounds. In a mixture of sodium phosphate buffer (10 mM, pH 7.0), sodium chloride solution (100 mM) and DMSO (5%), DQG analogs (20 μM) and DNA double strand containing GGCCCC/GGCCCC motif DNA, 5′-GCATGGCCCCTACG-3′/5′-CGTAGGCCCCATGC-3′ duplex DNA (5.0 μM) was dissolved in the mixture solution for the measurement. The prepared sample was heated at a rate of 1 ºC/ min and the change in the absorption spectra at 260 nm was observed and T_m was calculated by median method. The binding ability was compared by calculating ΔT_m, which is the difference between the T_m in the presence and absence of the ligand. (Fig. 9)

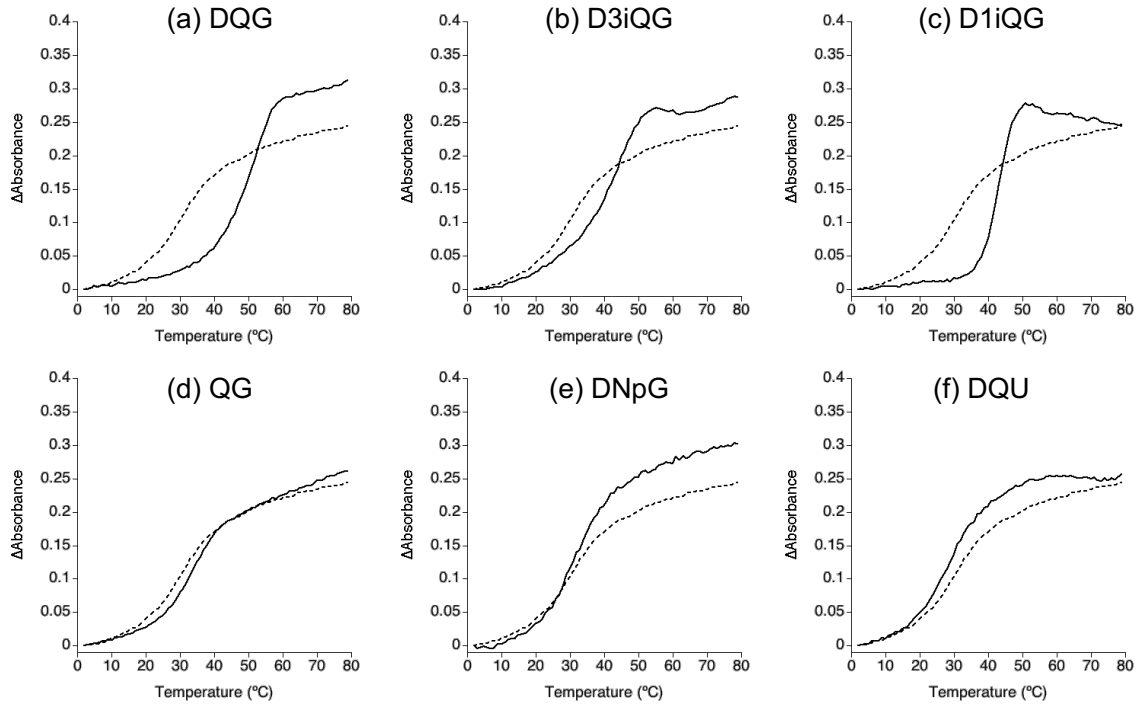
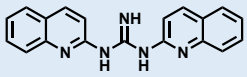
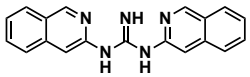
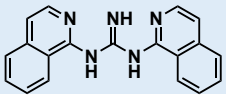
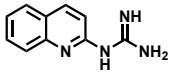
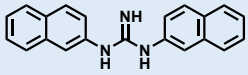
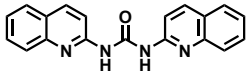


Fig. 9 Thermal melting profiles of DNA duplexes in the presence and absence of ligands (a) DQG, (b) D3iQG, (c) D1iQG, (d) QG, (e) DNpG, and (f) DQU.

Table. 1 Melting temperature is calculated by median method. ΔT_m (ºC) is calculated as difference of T_m (ºC) in the absence T_m (–) and presence T_m (+) of ligands. All measurements were made three times or more.

Entry	Ligand	T_m (+) (*S.D.*)		ΔT_m
1	**DQG** 	49.4	(0.3)	19.5
2	**D3iQG** 	40.7	(1.4)	10.9
3	**D1iQG** 	43.2	(0.7)	13.4
4	**QG** 	33.2	(0.8)	3.4
5	**DNpG** 	31.1	(0.4)	1.3
6	**DQU** 	30.3	(1.0)	0.4

As a result, T_m of the control compounds and DQG analogs were not higher than that of DQG. (Table. 1) The control compounds, DNpG and DQU, showed very little binding. T_m value of QG increased 3 degrees higher than DNpG and DQU, indicating that two quinolines are required for DQG to bind to G2C4/G2C4 compared to QG. The reason for not binding to G2C4/G2C4 was DNpG and DQU could not form a planar structure in the MM calculation results. In contrast, D3iQG and D1iQG differ from DQG in the position of the benzene, but their T_m values were lower than T_m of DQG. The position of the benzene was found to be an important factor for binding to G2C4/G2C4.

• Circular Dichroism

In order to obtain the binding information of D3iQG and D1iQG, circular dichroism (CD) was measured. The interaction between DQG and DQG analogs was evaluated by circularly polarized light dichroism spectroscopy using DNA containing G2C4/G2C4 motifs. (Fig. 10). In a mixture of potassium phosphate buffer (10 mM, pH 7.0), aqueous sodium chloride solution (100 mM), and DMSO (5%), Ligand and DNA double-stranded containing GGCCCC/GGCCCC motif 5′-GCATGGCCCCTACG-3′/5′-CGTAGGCCCCATGC-3′ duplex DNA (5.0 μM) was dissolved and measured. The concentrations of ligands were added in the ratio of 0, 2, 4, 6, 8, and 10 to the DNA. The results in the CD spectra of DQG and D3iQG showed that the positive 280 nm and negative 255 nm bands appeared without ligand but with nucleic acid changed with increasing ligand concentration. Furthermore, a positive band around 255 nm was commonly observed.

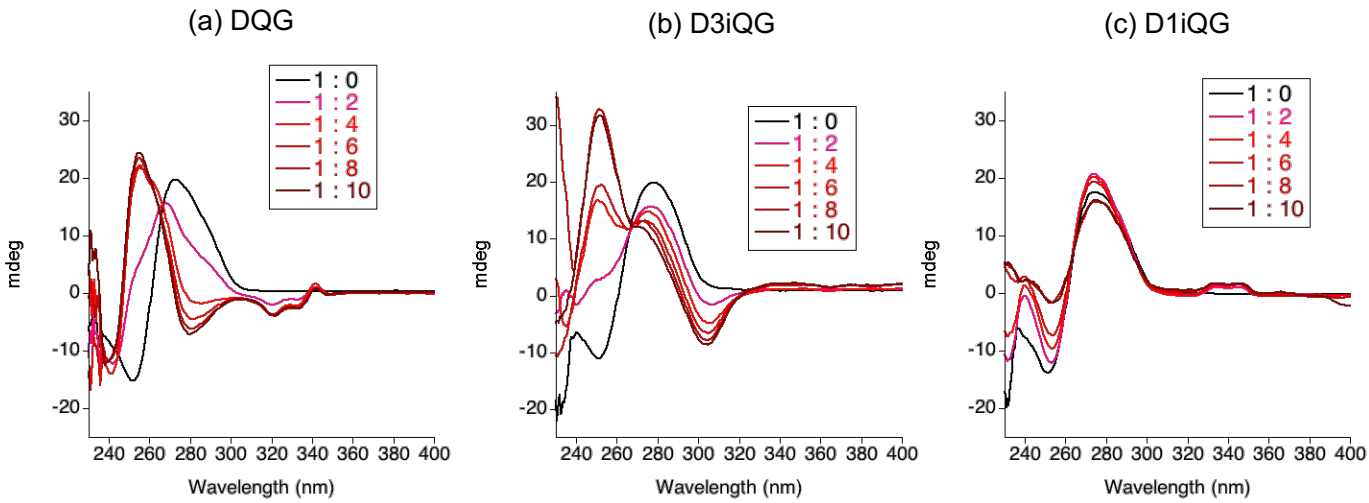


Fig. 10 CD spectra of G2C4/G2C4 motif containing DNA duplexes (5 μM), in the absence (black line) and higher concentration presence (red line) of the ligand (10, 20, 30, 40, 50 μM) in 10 mM phosphate buffer (pH 7.0), 100 mM NaCl and 5% DMSO. Key: (a) DQG, (b) D3iQG, (c) D1iQG.

On the other hand, the CD profile of D1iQG showed ligand-derived induced CD bands in the 330-360 nm region, but there were changes in the nucleic acid region around 255 nm and 280 nm, but these changes were smaller than those of DQG and D3iQG. The CD results suggested that the nucleic acid structure was changed by DQG and D3iQG binding to G2C4/G2C4, while not changed by D1iQG binding.

・Design and Synthesis of Q3iQG

In the previous section, it suggested that DQG and D3iQG have similarities in their binding mode. Thus, in order to further clarify the boundary between the two and explore the binding factors, a hetero compound, *N*-(isoquinolin-3-yl)-*N'*-(quinolin-2-yl)guanidine (Q3iQG) including the moieties of DQG and D3iQG was designed. (Fig. 11) Q3iQG has the potential to form intramolecular hydrogen bonds and the planar coordination formed was expected to be hydrogen bonding with cytosine such as DQG and the other DQG analogs. In addition, the three planar structures formed with the intramolecular hydrogen bond, which is different from DQG, D3iQG, and D1iQG. Q3iQG was calculated to confirm that the planar conformations were significantly formed.

Q3iQG was synthesized QG using Buchwald-Hartwig reaction with QG and 3-chloroisoquinoline. It was purified by HPLC.

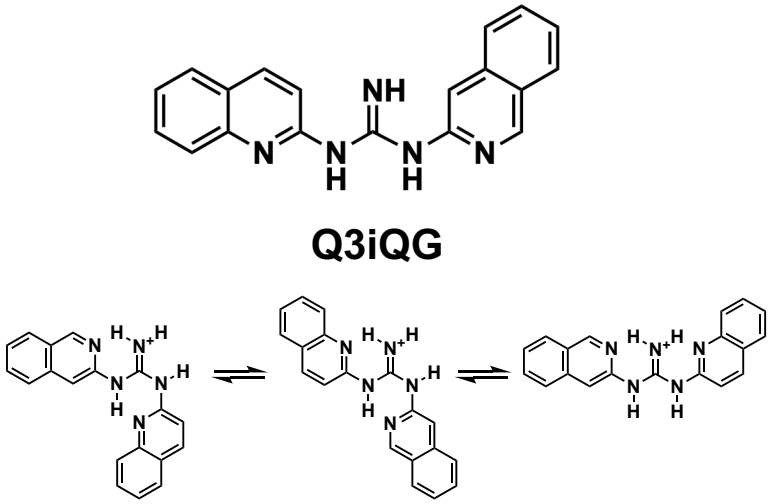


Fig. 11 Chemical structure of Q3iQG.

T_m measurement showed the increase of 19 ºC in Q3iQG, comparable to that of DQG, suggesting that 3-isoquinoline guanidine do not provide the same stabilizing effect on G2C4/G2C4 as quinoline guanidine. (Fig. 12-a)

CD measurement was performed using Q3iQG and showed that the nucleic acid structure was significantly changed at 255 and 280 nm at the ligand's concentration of 20 μM. (Fig. 12-b)

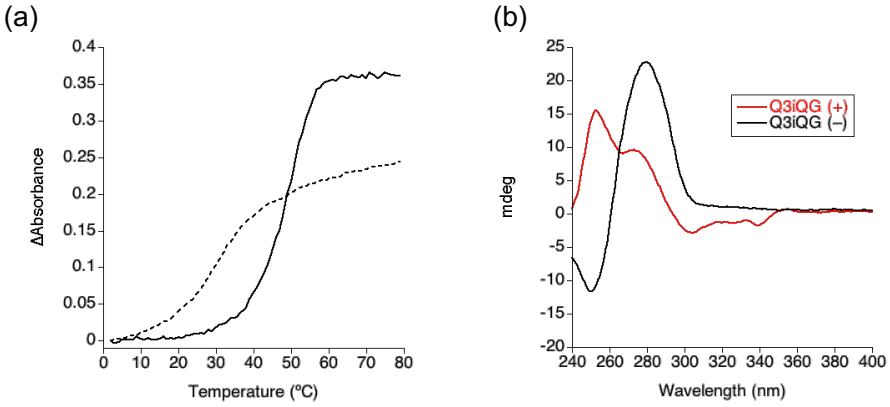


Fig. 12 (a) Thermal melting profiles of DNA duplexes in the presence (black line) and absence (dot line) of Q3iQG. (b) CD spectra of G2C4/G2C4 motif containing DNA duplexes (5 μM), in the absence (black line) and presence (red line) of Q3iQG (20 μM).

・SPR

In addition, surface plasmon resonance (SPR) measurements were performed to evaluate the binding affinity for G2C4/G2C4. Single-cycle kinetics analysis in SPR was used to evaluate the binding of the ligand to hairpin sequences with G2C4/G2C4 motifs. First, all ligands were measured at the same concentrations (0.25, 0.5, 1.0, 2.0, 4.0 μM). DQG control compounds, DNpG and DQU, showed not binding to G2C4/G2C4. (Fig. 13) QG interacted with G2C4/G2C4 more than DNpG and DQU, which was correlated with the results in T_m measurements. The apparent K_D of DQG and Q3iQG were found to be 1.0 μM and 1.4 μM, indicating similar binding affinity for G2C4/G2C4. (Fig. 14) D1iQG had lower affinity than D3iQG. In addition, DQG and all DQG analogs showed similar association and dissociation shape in the sensorgram, indicating that the binding patterns of these ligands to GGCCCC repeats were similar. Consistently, the sensorgram rose with binding and drops from the top to the dissociation phase. Presumably, the complex was changing after contacting. In the sensor chip, after the ligand initially contacted the nucleic acid, the complex of nucleic acid and ligand formed the stable structure. This phenomenon was often observed with poorly water-soluble ligands, but the details of this have not yet been clarified.[29]

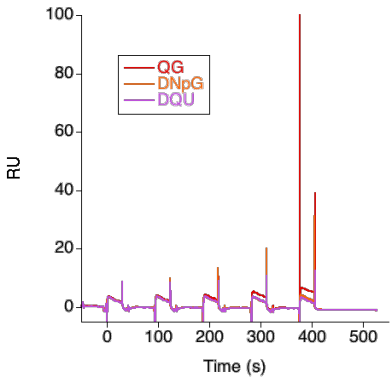


Fig. 13 SPR data for ligands binding to surface-immobilized G2C4/G2C4 hairpin structure. The ligands were sequentially added at 0.25, 0.50, 1.0, 2.0, 4.0 μM (QG, DNpG, DQU).

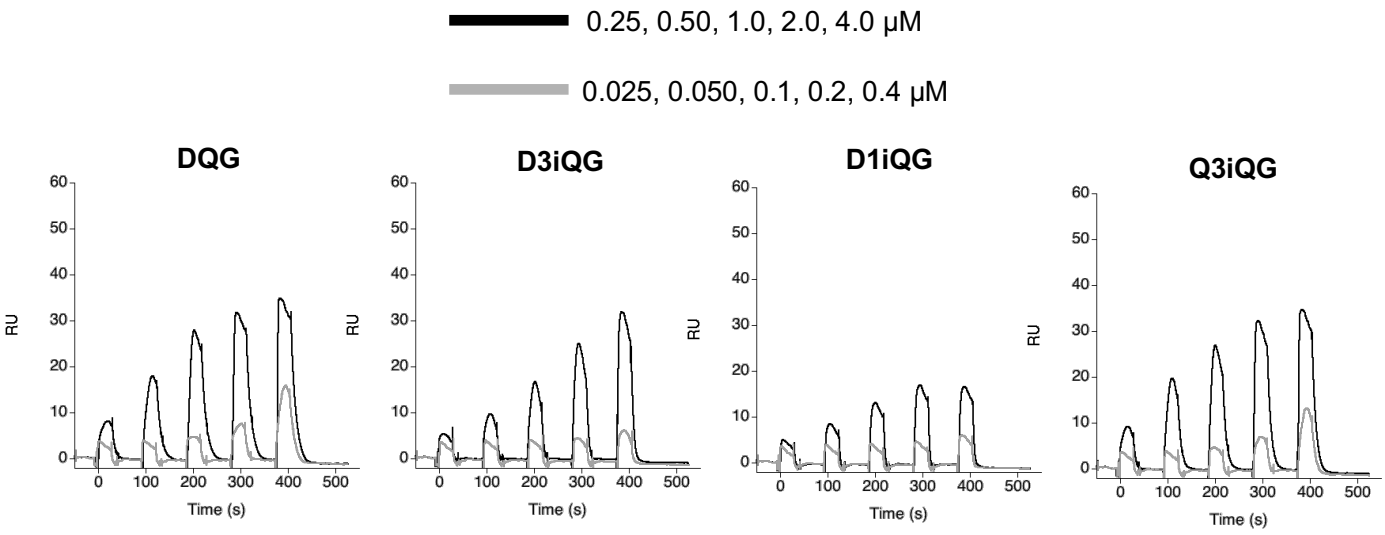


Fig. 14 SPR data for ligands binding to surface-immobilized G2C4/G2C4 hairpin structure. The ligands were sequentially added at 0.25, 0.50, 1.0, 2.0, 4.0 μM and 0.025, 0.050, 0.1, 0.2, 0.4 μM (DQG, D3iQG, D1iQG, Q3iQG).

・Quantitative Structure-Activity Relationship

These evaluations showed that the position of benzene in the heterocyclic ring of DQG was important and that the quinoline and guanidine moieties were necessary for binding, as DQG showed the same degree of binding in Q3iQG with one of the moieties changed to isoquinoline. On the other hand, QG, the compound from which the common structure was extracted, bound to G2C4/G2C4 motif DNA but did not stabilize significantly. In other words, the QG moiety required an additional heteroaromatic ring to stabilize G2C4/G2C4.

Assuming the formation of hydrogen bonds with cytosine, in the molecular modeling, the ligand showed a single molecule bound complex to the 5′-CCC-3′/5′-GG-3′ bulge of cytosine, forming several planar structures in DQG, Q3iQG, D3iQG and D1iQG. (Fig. 15) Among the planar conformations, those that are hydrogen bonded to cytosine and those that do not have a proton of the heterocyclic ring close to the amine of cytosine were selected and modeled. The conformations obtained from the molecular calculations, CS1 for DQG, CS1 and CS2 for Q3iQG, CS1 for D3iQG, and CS5 for D1iQG were selected. After arranging cytosine and ligand to form hydrogen bonds, the figures showed the results after MM calculation-minimization by constraining the sugars bases, and phosphates without the ligand, the sugar and phosphate connected to the bulge cytosine, and the bulge cytosine. (Fig. 16, Fig. 17, Fig. 18) In the modeling figures, the two constituent heterocycles of DQG, Q3iQG, and D3iQG protruded into the minor

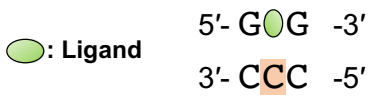


Fig. 15 Model structure: GG/CCC and DQG analogs

and major grooves. Furthermore, the heterocyclic ring on the major groove side was suggested to have a stacking interaction with the 3′-terminal guanine. In T_m and SPR, DQG and Q3iQG, which provided higher stabilization to G2C4/G2C4 than the other compounds, had a quinoline ring sticking out on the major groove side and overlapped with adjacent bases when they formed DQG-CS1 and Q3iQG-CS1. And the quinoline rings and isoquinoline rings appeared on the minor groove side. Presumably, this action on the minor groove side was hydrophobic interaction resulting from the increased entropy of water. On the other hand, the stabilization and binding affinity of D3iQG was lower than that of Q3iQG because it was not stabilized by the interaction with the neighboring bases. In addition, when D1iQG was attached, it went out of the duplex. This was due to the repulsion of isoquinoline from the phosphate backbone when it formed hydrogen bonds with cytosine. The stabilizing effect of QG on G2C4/G2C4 was significantly lower than that of DQG and Q3iQG. It was found that QG backbone and the whole molecule needed more planar than QG. The role of the structure on DQG binding to G2C4/G2C4 was

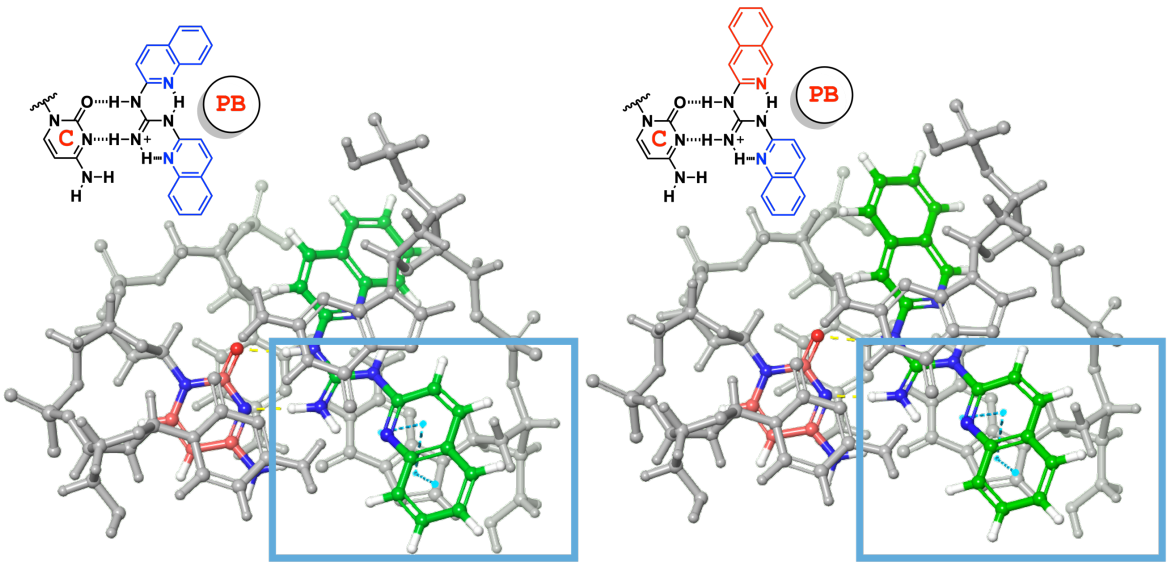


Fig. 16 Model structure: GG/CCC and (a) DQG or (b)Q3iQG – hydrogen bond

summarized. (Fig. 19) These are suggestions and further detailed studies will require elaborate molecular design, synthesis, and evaluation of the binding. In addition, information on hydrogen bonding with cytosine has not yet been obtained and needed to be clarified.

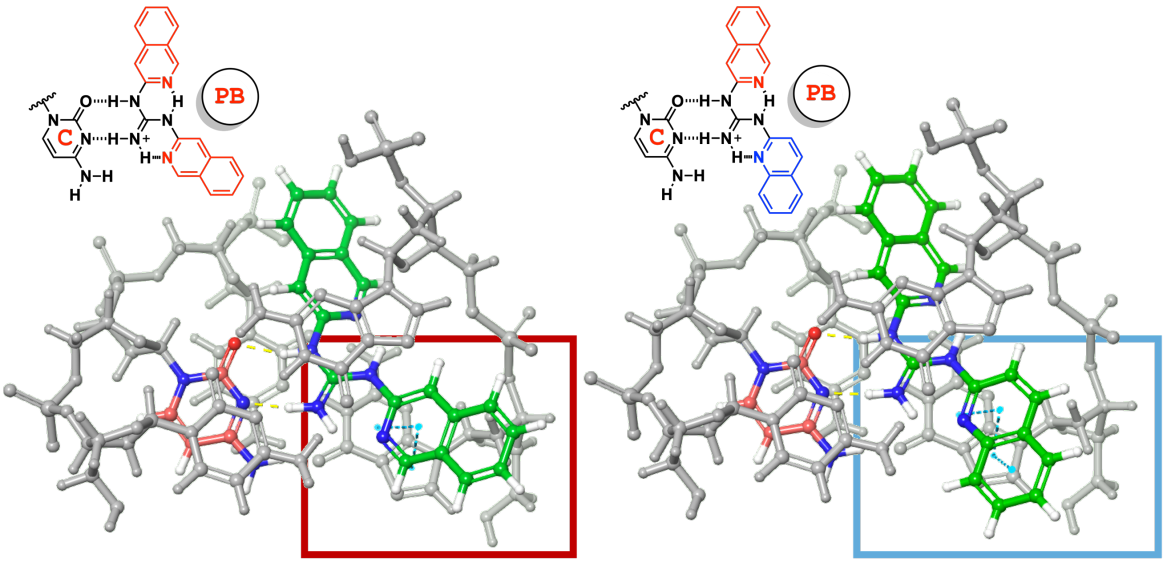


Fig. 17 Model structure: GG/CCC and (c) D3iQG or (b) Q3iQG – hydrogen bond

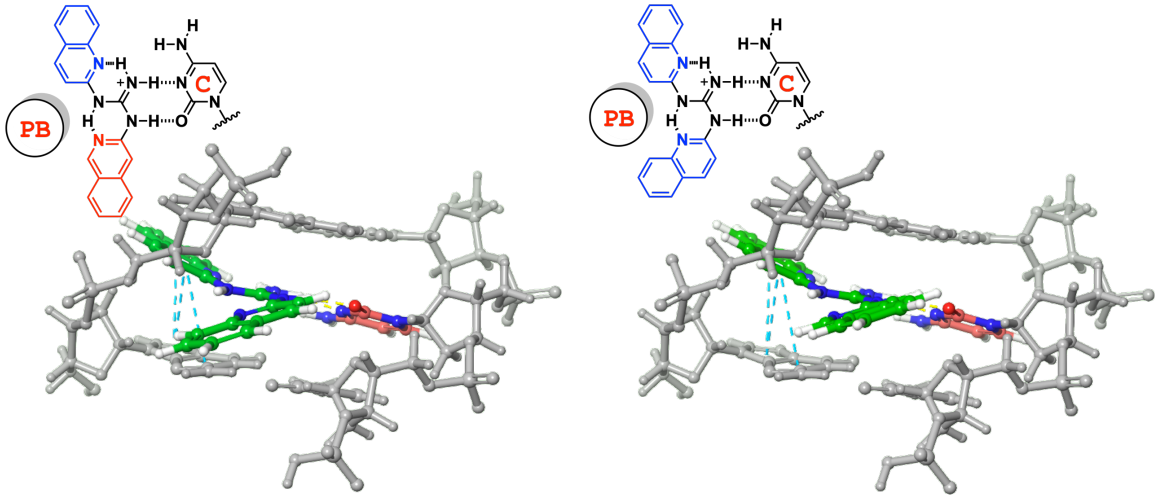


Fig. 18 Model structure: GG/CCC and (d) DQG or (e) Q3iQG – minor groove

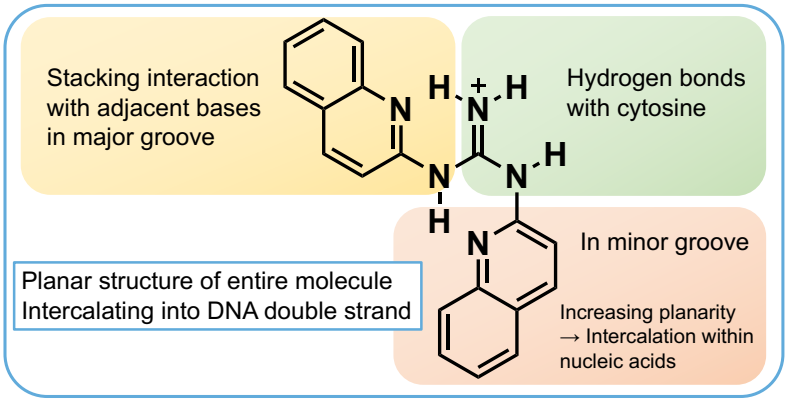


Fig. 19 DQG-binding structural factor

Conclusion

DQG, a molecule that bound to G2C4 repeats, was used as the lead compound and its analogs and control compounds were designed and synthesized. These compounds were evaluated for binding to G2C4/G2C4 to identify structural factors of DQG binding to G2C4/G2C4 and new compounds with more binding to G2C4/G2C4. The control compounds, DNpG, QG and DQU, were found to decrease binding ability in the T_m and SPR results. DQG analogs showed similar or lower stabilization than DQG, although binding to G2C4 was still observed. By the quantitative structure-activity relationship using the binding results, the structural binding factors for DQG were extracted. However, this study was not able to completely identify the structural factors of binding to G2C4/G2C4 DNA.

Materials and Methods

General

Reagents and solvents were purchased from standard suppliers and used without further purification. Reactions were monitored with TLC plates precoated with Merck silica gel 60 F254. Wako gel C-200 was used for silica gel flash chromatography. ^{1}H-NMR spectra were measured with BRUKER Avance Ⅲ 700, JEOL JNM-LA400 and JEOL LA600. Coupling constants (J values) are represented in hertz. The chemical shifts are expressed in ppm relative to residual solvent as an internal standard. ^{13}C-NMR spectra were measured with JEOL LA600 and BRUKER Avance Ⅲ 700. DNAs were purchased from Life Technologies.

Synthesis

_N,N'_-di(isoquinolin-3-yl)guanidine – D3iQG

A mixture of guanidine nitrate (373.1 mg, 3.06 mmol), $Pd(OAc)_2$ (0.61 mmol, 137.2 mg), SPhos (1.22 mmol, 501.7 mg) and Cs_2CO_3 (9.17 mmol, 2.99 g) in 1,4-dioxane (25.0 mL) was heated at 100 °C for 1 h under argon atmosphere. And 3-chloroisoquinoline (500.0 mg, 3.06 mmol) in 1,4-dioxane (25.0 mL) was added and heated at 100 ºC for 45 h. The crude product was purified by silica gel chromatography (AcOEt and $CHCl_3$/MeOH = 1:0~10:1). The crude product was removed the solvent, diluted with H_2O and purified by HPLC (eluent: CH_3CN and 0.1% AcOH) to yield pure D3iQG (7.2 mg, 2%). ^{1}H-NMR (CD_3OD, 700 MHz) δ = 9.16 (s, 2H), 8.07 (d, J = 8.2 Hz, 2H), 7.87 (d, J = 8.3 Hz, 2H), 7.74 (d, J = 7.1 Hz, 2H), 7.58-7.55 (m, 4H); ^{13}C-

NMR (CD3OD, 175 MHz) δ = 154.7, 151.5, 150.0, 139.6, 132.8, 129.1, 127.5, 127.3, 127.1, 109.1; HRMS (ESI) *m/z*: calcd. for $[C_{19}H_{15}N_5+H]^+$ 314.1400; found 314.1390.

***N,N'*-di(isoquinolin-1-yl)guanidine – D1iQG**

A mixture of guanidine nitrate (373.1 mg, 3.06 mmol), $Pd(OAc)_2$ (0.61 mmol, 137.2 mg), XPhos (1.22 mmol, 582.8 mg) and Cs_2CO_3 (9.17 mmol, 2.99 g) in 1,4-dioxane (25.0 mL) was heated at 100 °C for 1 h under argon atmosphere. And 1-chloroisoquinoline (500.0 mg, 3.06 mmol) in 1,4-dioxane (25.0 mL) was added and heated at 100 ºC for 21 h. The crude product was purified by silica gel chromatography (AcOEt and $CHCl_3$/MeOH = 1:0~10:1). The crude product was removed the solvent, diluted with H_2O and purified by HPLC (eluent: CH_3CN and 0.1% AcOH) to yield pure D1iQG (3.0 mg, 0.3%). ^{1}H-NMR (CD_3OD, 600 MHz) δ = 8.58 (d, J = 8.2 Hz, 1H), 8.18 (d, J = 5.8 Hz, 1H), 7.81 (d, J = 8.2 Hz, 1H), 7.72 (t, J = 7.6 Hz, 1H), 7.66 (t, J = 7.2 Hz, 1H), 7.31 (d, J = 5.5 Hz, 1H); HRMS (ESI) *m/z*: calcd. for $[C_{19}H_{15}N_5+H]^+$ 314.1400; found 314.1391.

***N*-(isoquinolin-3-yl)-*N'*-(quinolin-2-yl)guanidine – Q3iQG**

A mixture of 1-(quinolin-2-yl)guanidine (447.6 mg, 2.40 mmol), $Pd(OAc)_2$ (0.96 mmol, 216.0 mg), XPhos (1.92 mmol, 917.2 mg) and Cs_2CO_3 (7.21 mmol, 2.35 g) in 1,4-dioxane (20.0 mL) was heated at 110 °C for 1 h under argon atmosphere. And 3-chloroisoqinoline (393.2 mg, 2.40 mmol) in 1,4-dioxane (20.0 mL) was added and heated at 100 ºC for 45 h. The crude product was purified by silica gel chromatography (AcOEt/Hex = 1:1 and $CHCl_3$/MeOH = 1:0~10:1). The crude product was removed the solvent, diluted with H_2O and purified by HPLC (eluent: CH_3CN

and 0.1% AcOH) to yield pure Q3iQG (34.3 mg, 6%).^{1}H-NMR (CD_3OD, 700 MHz) δ = 9.16 (s, 1H), 8.23 (d, J = 8.7 Hz, 1H), 8.05 (d, J = 8.1 Hz, 1H), 7.95 (d, J = 8.2 Hz, 1H), 7.88 (d, J = 8.1 Hz, 1H), 7.83 (d, J = 7.8 Hz, 1H), 7.71 (mbr, 2H), 7.54 (t, J = 7.4 Hz, 1H), 7.51 (s, 1H), 7.48 (t, J = 7.3 Hz, 1H), 7.21 (d, J = 8.6 Hz, 1H); ^{13}C-NMR (CD_3OD, 175 MHz) δ = 155.8, 151.5, 146.8, 139.9, 139.7, 132.6, 131.4, 129.0, 128.8, 127.8, 127.2, 127.1, 126.4, 126.1, 117.2, 109.2; HRMS (ESI) *m/z*: calcd. for $[C_{19}H_{15}N_5+H]^+$ 314.1400; found 314.1400.

***N,N'*-di(quinazolin-2-yl)guanidine – DQzG**

A mixture of guanidine nitrate (148.3 mg, 1.22 mmol), 2-chloroquinazoline (200.0 mg, 1.22 mmol), $Pd(OAc)_2$ (0.24 mmol, 54.6 mg), XPhos (0.49 mmol, 231.7 mg) and Cs_2CO_3 (3.65 mmol, 1.19 g) in 1,4-dioxane (6.0 mL) was heated at 100 °C for 20 h under argon atmosphere. The reaction mixture was added to water and the solvent was evaporated. The resulting mixture was added to chloroform and washed with water and brine. The organic layer was dried over $MgSO_4$ and evaporated to dryness. The crude product was removed the solvent, diluted with H_2O and purified by HPLC (eluent: CH_3CN and 0.1% AcOH) to yield pure DQzG (11.8 mg, 6%). ^{1}H-NMR (CD_3OD, 700 MHz) δ = 9.38 (s, 2H), 7.99 (d, J = 7.9 Hz, 2H), 7.92 (dt, J = 7.5, 1.2 Hz, 2H), 7.88 (d, J = 8.3 Hz 2H), 7.55 (t, J = 7.0 Hz 2H); ^{13}C-NMR (CD_3OD, 175 MHz) δ = 164.0, 160.0, 157.2, 151.2, 136.3, 129.0, 127.1, 126.7, 122.7; HRMS (ESI) *m/z*: calcd. for $[C_{17}H_{13}N_7+H]^+$ 316.1305; found 316.1296.

UV-melting analysis

Thermal denaturation profiles were recorded on a UV-2700 spectrophotometer (Shimadzu) equipped with a TMSPC-8 temperature controller and a 10 mm path-length cell. The absorbance of DNAs (5 μM DNA duplex) with ligand (20 μM) in phosphate buffer (10 mM, pH 7.0) containing NaCl (100 mM) and 5% DMSO was monitored at 260 nm from 2 to 80 ºC or 100 °C (1 °C min^{-1}). T_m was calculated by using the median method.

Circular Dichroism (CD) Measurements

CD experiments were performed with a J-725 CD spectrometer (JASCO) using a 10 mm path length cell. CD spectra of DNAs (5 μM DNA duplex) in the absence and presence of ligand (0-50 μM) were measured in phosphate buffer (10 mM, pH 7.0) containing NaCl (100 mM) and 5% DMSO.

Surface Plasmon Resonance (SPR)

SPR single cycle kinetics assay was performed using Biacore T200 platform (GE Healthcare, Life Science). Immobilization of the oligomers on the Series S sensor chip SA surface was carried out using avidin–biotin coupling in HBS-EP^+ buffer (10 mM pH 7.4 HEPES, 150 mM NaCl, 3 mM EDTA, 0.005% surfactant P20). Concentration of 5′-biotinylated oligonucleotides were adjusted in 10 mM HEPES-500 mM NaCl. Biotinylated 5′-GCATGGCCCCTACGTTTTCGTAGGCCCCATGC-3′ was immobilized on SA surface via avidin–biotin. Blank immobilization was performed in the flow cell 1 to permit reference subtraction. Ligand solution was diluted using HBS-EP+ buffer. Sensorgrams were obtained in

the ligand concentrations of 0.25, 0.50, 1.0, 2.0, 4.0 μM and 0.025, 0.050, 0.1, 0.2, 0.4 μM in single-cycle mode (contact time 30 s, Dissociation time 120 s, Flow rate 60 μM / min). All sensorgrams were corrected by reference subtraction of blank flow cell response and buffer injection response.

References

1. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **72**, 245–256 (2011).
2. Renton, A. E. *et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257–268 (2011).
3. van Blitterswijk, M. *et al.* Association between repeat sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): A cross-sectional cohort study. *Lancet Neurol.* **12**, 978–988 (2013).
4. Swinnen, B., Robberecht, W. & Van Den Bosch, L. RNA toxicity in non-coding repeat expansion disorders. *EMBO J.* **39**, 1–23 (2020).
5. Hideyama, T. *et al.* Profound downregulation of the RNA editing enzyme ADAR2 in ALS spinal motor neurons. *Neurobiol. Dis.* **45**, 1121–1128 (2012).
6. Zhang, N. & Ashizawa, T. RNA toxicity and foci formation in microsatellite expansion diseases. *Curr. Opin. Genet. Dev.* **44**, 17–29 (2017).
7. Mori, K. *et al.* The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science (80-.).* **339**, 1335–1338 (2013).
8. Shi, Y. *et al.* Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat. Med.* **24**, 313–325 (2018).
9. Xi, Z. *et al.* Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. *Am. J. Hum. Genet.* **92**, 981–989 (2013).

10. Xi, Z. *et al.* The C9orf72 repeat expansion itself is methylated in ALS and FTLD patients. *Acta Neuropathol.* **129**, 715–727 (2015).

11. Jackson, J. L. *et al.* Elevated methylation levels, reduced expression levels, and frequent contractions in a clinical cohort of C9orf72 expansion carriers. *Mol. Neurodegener.* **15**, 1–11 (2020).

12. Zamiri, B., Mirceta, M., Bomsztyk, K., Macgregor, R. B. & Pearson, C. E. Quadruplex formation by both G-rich and C-rich DNA strands of the C9orf72 (GGGGCC)8•(GGCCCC)8 repeat: Effect of CpG methylation. *Nucleic Acids Res.* **43**, 10055–10064 (2015).

13. Castel, A. L., Cleary, J. D. & Pearson, C. E. Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat. Rev. Mol. Cell Biol.* **11**, 165–170 (2010).

14. Schmidt, M. H. M. & Pearson, C. E. Disease-associated repeat instability and mismatch repair. *DNA Repair (Amst).* **38**, 117–126 (2016).

15. Iyer, R. R., Pluciennik, A., Napierala, M. & Wells, R. D. DNA triplet repeat expansion and mismatch repair. *Annu. Rev. Biochem.* **84**, 199–226 (2015).

16. Kumar, V., kashav, T., Islam, A., Ahmad, F. & Hassan, M. I. Structural insight into C9orf72 hexanucleotide repeat expansions: Towards new therapeutic targets in FTD-ALS. *Neurochem. Int.* **100**, 11–20 (2016).

17. Zhang, Y., Roland, C. & Sagui, C. Structure and Dynamics of DNA and RNA Double Helices Obtained from the GGGGCC and CCCCGG Hexanucleotide Repeats That Are the Hallmark of C9FTD/ALS Diseases. *ACS Chem. Neurosci.* **8**, 578–591 (2017).

18. Brčić, J. & Plavec, J. ALS and FTD linked GGGGCC-repeat containing DNA oligonucleotide folds into two distinct G-quadruplexes. *Biochim. Biophys. Acta - Gen. Subj.* **1861**, 1237–1245 (2017).

19. Zhang, Y., Roland, C. & Sagui, C. Structural and Dynamical Characterization of DNA and RNA Quadruplexes Obtained from the GGGGCC and GGGCCT Hexanucleotide Repeats Associated with C9FTD/ALS and SCA36 Diseases. *ACS Chem. Neurosci.* **9**, 1104–1117 (2018).

20. Zhao, J., Bacolla, A., Wang, G. & Vasquez, K. M. Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci.* **67**, 43–62 (2010).

21. Lat, P. K. & Sen, D. (C2G4)n repeat expansion sequences from the C9orf72 gene form an unusual DNA higher-order structure in the pH range of 5-6. *PLoS One* **13**, 1–20 (2018).

22. Pan, F., Zhang, Y., Man, V. H., Roland, C. & Sagui, C. E-motif formed by extrahelical cytosine bases in DNA homoduplexes of trinucleotide and hexanucleotide repeats. *Nucleic Acids Res.* **46**, 942–955 (2018).

23. Simone, R. *et al.* G-quadruplex-binding small molecules ameliorate C9orf72 FTD / ALS pathology in vitro and in vivo . *EMBO Mol. Med.* **10**, 22–31 (2018).

24. Brown, S. L. & Kendrick, S. The I-motif as a molecular target: More than a complementary DNA secondary structure. *Pharmaceuticals* **14**, 1–25 (2021).

25. Cooper-Knock, J. *et al.* Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain* **137**, 2040–2051 (2014).

26. Takahashi, S., Bhattacharjee, S., Ghosh, S., Sugimoto, N. & Bhowmik, S. Preferential targeting cancer-related i-motif DNAs by the plant flavonol fisetin for theranostics applications. *Sci. Rep.* **10**, 1–13 (2020).

27. Zeraati, M. *et al.* I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.* **10**, 631–637 (2018).

28. Shibata, T., Murakami, E. & Nakatani, K. 1,3-Di(quinolin-2-yl)guanidine binds to GGCCCC hexanucleotide repeat DNA in C9ORF72. *Bioorganic Med. Chem. Lett.* **28**, 2364–2368 (2018).

29. Matsumoto, J. *et al.* The Dimeric Form of 1,3-Diaminoisoquinoline Derivative Rescued the Mis-splicing of Atp2a1 and Clcn1 Genes in Myotonic Dystrophy Type 1 Mouse Model. *Chem. - A Eur. J.* **26**, 14305–14309 (2020).

Chapter 1-2

Redesign of Binding Molecules Based on Binding Information and Evaluation of Binding to G2C4/G2C4 DNA

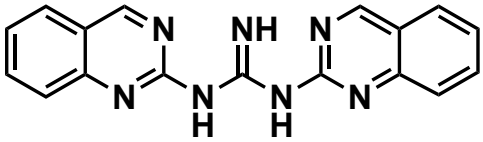
Molecular Design

In chapter 1-1, the binding evaluation with the control compound and DQG analogs revealed that the planar conformation and the hydrogen-bonding surface that can be formed with cytosine were important for binding to G2C4/G2C4 DNA. The compound was designed and synthesized to further support this nucleic acid binding information and to improve the binding ability.

・Design of New Small Molecule Binding to G2C4/G2C4 DNA

The structure required for binding to G2C4 was found to be guanidine and a heteroaromatic ring with intramolecular hydrogen bonding to guanidine. It also suggested that the position of the benzene in the heteroaromatic ring of quinolines was important for binding. *N,N'*-di(quinazolin-2-yl)guanidine (DQzG) was designed and has guanidine-linked two quinazoline rings, the C at position 3 of quinoline converted to nitrogen. (Fig. 1) In addition to the continuous hydrogen-bond donor of DQG, the newly added nitrogen of quinazoline become one hydrogen-bond acceptor and the intramolecular hydrogen bonding generated the ADD hydrogen-bonding surface. (Fig. 2) As a result, DQzG was assumed to form stronger hydrogen bonds than DQG with cytosine,

which has the DAA hydrogen-bonding surface as well as guanine-cytosine base pair. Furthermore, DQzG can form the similarly planar conformer formed by DQG, D3iQG, and Q3iQG. This multiple conformational formation has a potential to stabilize G2C4/G2C4 DNA, since it was able to adopt the appropriate conformation for cytosine in DNA duplex.



DQzG

Fig. 1 Chemical structure of DQzG.

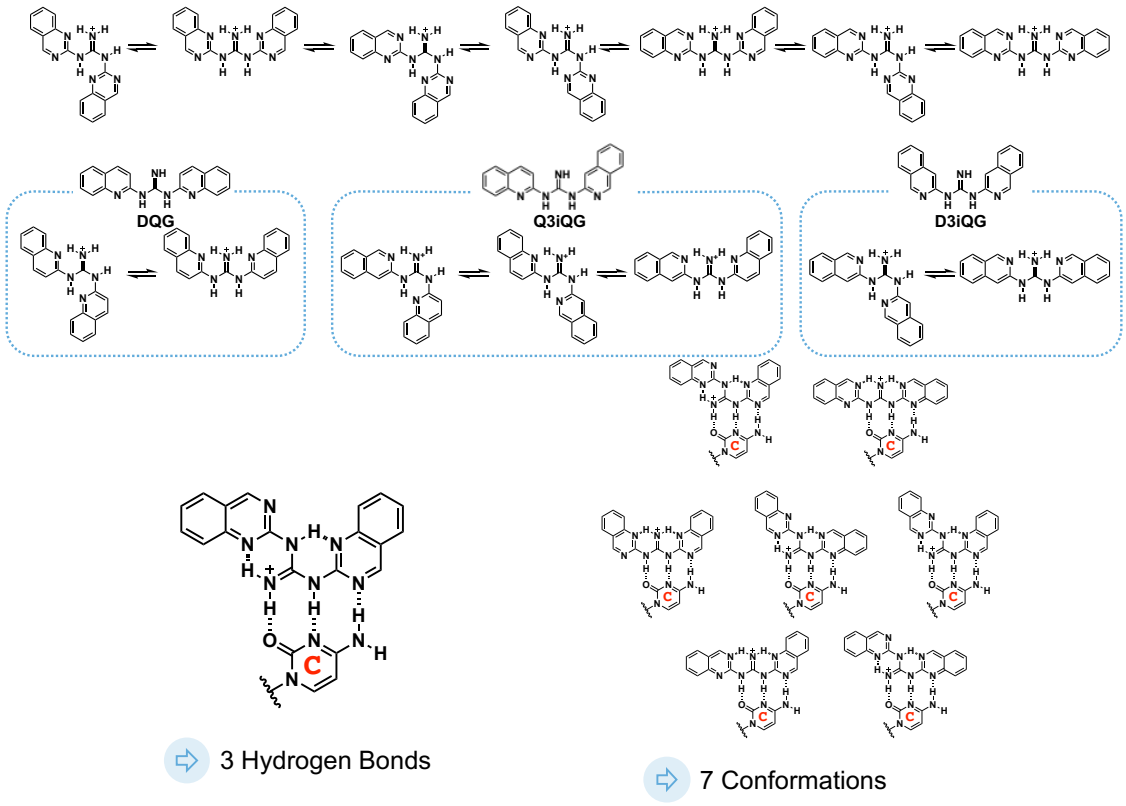


Fig. 2 The features of DQzG and proposed schematic illustrations of hydrogen bonding patterns with DQzG and cytosine.

As well as DQG analogs study, it was necessary to clarify whether DQzG can form planar conformations and whether the conformations were energetically more stable than non-planar structures using computational method. MM-conformational search obtained 11 conformations, of which 7 were planar and all the assumed planar conformations were exhausted. In addition, QM-optimization calculations were performed for all 11 conformations of DQzG were also performed as well as the DQG analogs study and it was found that all planar conformations can form within 3 kcal/mol of energy. The energy of the non-planar conformations was 16 kcal/mol. The results of this calculation suggested that DQzG could form planar conformations. The most stable conformation obtained by molecular calculations for DQzG is shown in Fig. 3.

DQzG was synthesized in one step by Buchwald-Hartwig reaction starting from purchasable 2-chloroquinazoline.

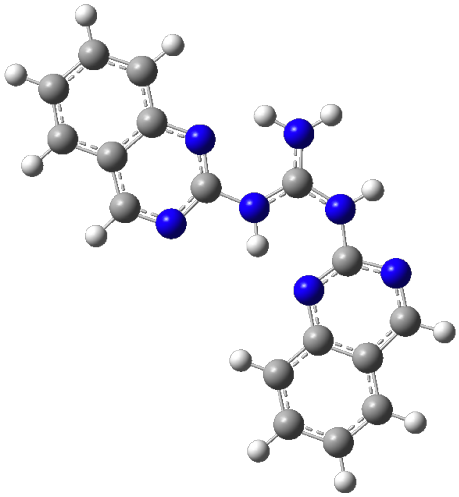


Fig. 3 The most stable conformation obtained by molecular calculations in DQzG.

Results & Discussion

• Melting Temperature

DQzG binding to G2C4 was evaluated by measuring the melting temperature (T_m) of DNA double strands containing the GGCCCC/GGCCCC motif in the presence and absence of the compound. (Fig. 4) In a mixture of sodium phosphate buffer (10 mM, pH 7.0), sodium chloride solution (100 mM) and DMSO (5%), DQzG (20 μM) and DNA double strand containing GGCCCC/GGCCCC motif 5'-GCATGGCCCCTACG-3'/5'-CGTAGGCCCCATGC-3' duplex DNA (5.0 μM) was dissolved in the solution for the measurement. The prepared sample was heated at a rate of 1 ºC/ min and the change in the absorption spectrum at 260 nm was observed

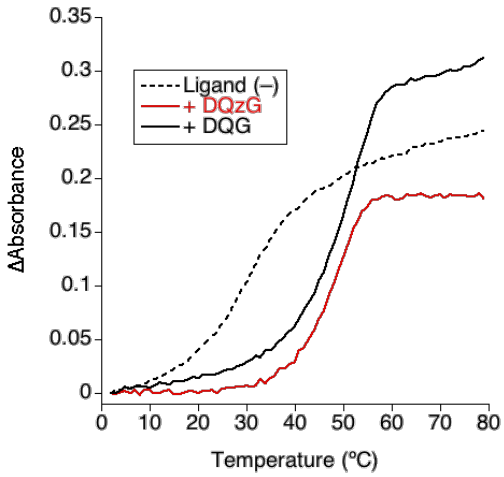


Fig. 4 Thermal melting profiles of GGCCCC/GGCCCC motif containing DNA duplexes (5 μM), without ligand (black dot line) and with DQG (black line), DQzG (red line) (20 μM) in 10 mM phosphate buffer, 100 mM NaCl and 5% DMSO. Detection at 260 nm.

and T_m was calculated by median method. The binding capacity was compared by calculating ΔT_m, the difference in T_m in the presence and absence of the ligand.

The calculated ΔT_m of DQzG was 20.2. T_m measurement was found to stabilize G2C4/G2C4 on the same degree as DQG.

・SPR

Surface plasmon resonance (SPR) measurements were performed to evaluate the binding affinity of DQzG to G2C4/G2C4 as well as DQG and DQG analogs. SPR assay was used to evaluate the binding of ligands to hairpin sequences with G2C4/G2C4 motifs. Single-cycle kinetics analysis was used with immobilized chips of DQG analogs measurements.

First, DQzG was measured on the concentration scale (0.25, 0.5, 1.0, 2.0, 4.0 μM). The results showed that DQzG bound to the G2C4 hairpin structure. (Fig. 5) Since the binding was saturated at this concentration scale, DQzG was also measured at the concentration scale (0.025, 0.05, 0.1, 0.2, 0.4 μM). The results showed that the DQG analogs had higher affinity than DQG and DQG analogs. Using the two concentration scales, the apparent K_D was calculated, 63 nM, which was lower than that of DQG and DQG analogs. Furthermore, focusing on the shape of the sensorgram, in the previous section, focusing on the shape of the sensorgram, there were two phases of rising and falling in the binding phase in DQG, DQzG was found as well. On the other hand, it was not found at lower concentration scale. The reason why this phenomenon in sensorgram was not observed at lower concentration scale was assumed to be related to the solubility of small

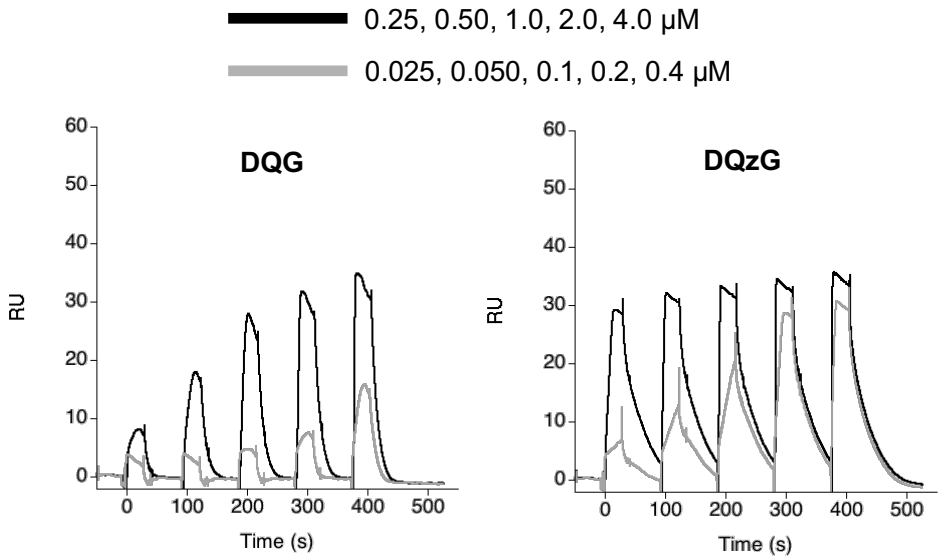


Fig. 5 SPR data for ligands binding to surface-immobilized G2C4/G2C4 hairpin structure. The ligands were sequentially added at 0.25, 0.50, 1.0, 2.0, 4.0 μM and 0.025, 0.050, 0.1, 0.2, 0.4 μM (DQG, DQzG).

molecules. Compared to DQG, DQzG dissociated more slowly in the dissociation phase, indicating that it was less likely to dissociate from G2C4/G2C4.

SPR measurements showed that the more higher binding affinities of DQzG than DQG. On the other hand, the change in the value of T_m in DQzG was the same degree as DQG, the difference of 0.7 degree.

There are two possible reasons for this effect.

In Fig. 4, the UV change with temperature was smaller for DQzG than for only nucleic acid. The temperature at 0.1 absorbance was 48 ºC for DQzG and 45 ºC for DQG. (Fig. 6) To investigate further, UV measurements were performed on DQzG and DQG in the temperature-dependent

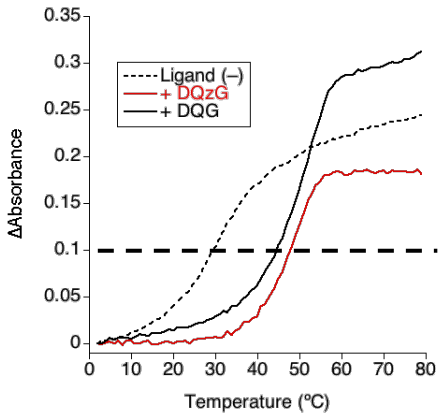


Fig. 6 Thermal melting profiles of GGCCCC/GGCCCC motif containing DNA duplexes (5 μM), without ligand (black dot line) and with DQG (black line), DQzG (red line) (20 μM) in 10 mM phosphate buffer, 100 mM NaCl and 5% DMSO. Detection at 260 nm. Abs. 0.1(broken line).

range of 200-500 nm. (Fig. 7) DQzG showed a decrease in absorption around 260 nm in the T_m measurement with increasing temperature. The existence of temperature-dependent UV changes in the ligands of DQzG and DQG makes it difficult to compare the changes in T_m.

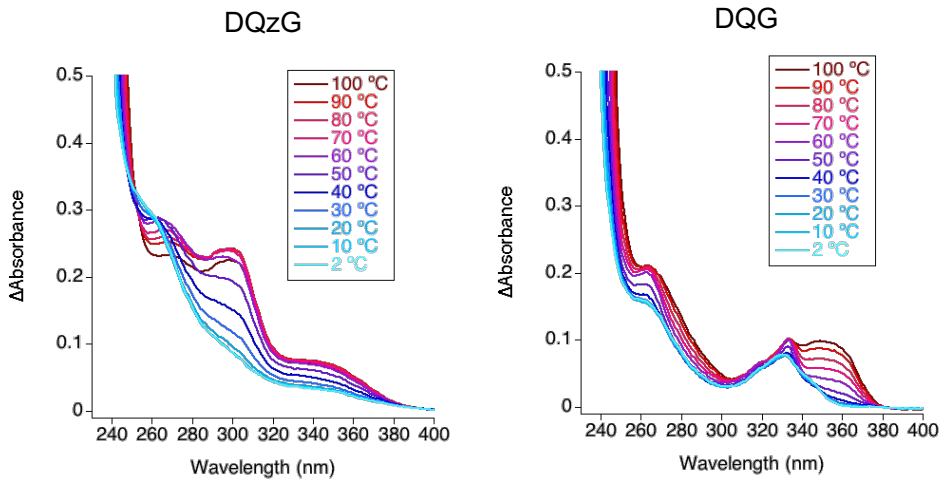


Fig. 7 Absorption spectra of DQG, DQzG (20 μM) recorded in 10 mM phosphate buffer, 100 mM NaCl and 5% DMSO at 2-100 ºC.

Second reason was the ligand characteristics of binding to G2C4/G2C4 DNA. DQzG can form more planar configurations than DQG, which accompanies the concept of stabilizing G2C4/G2C4 in the most suitable conformations. (Fig. 8) However, it is possible to lose the binding energy in selecting the conformation due to a lot of conformations. The T_m measurement profiles temperature-dependent UV change in the dissociation reaction. In T_m, the increase in temperature may have affected the dissociation reaction by increasing the energy of the conformational entropy. In SPR, the temperature was constant at ambient temperature and the association and dissociation reactions were monitored, the effect of the conformational entropy was smaller than T_m measurement.

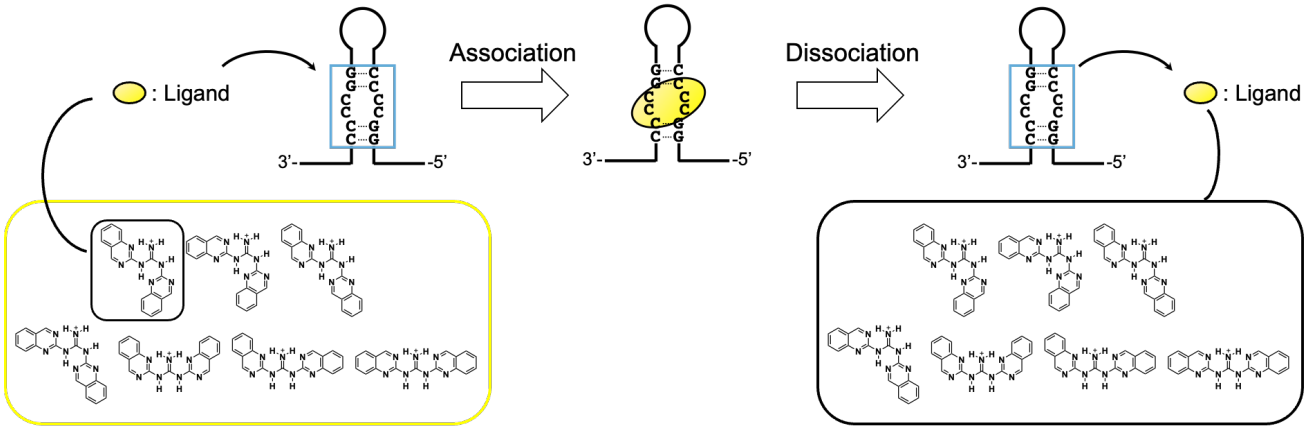


Fig. 8 Schmatic illustration of entropy loss in binding to G2C4/G2C4

・NMR

DQzG was shown to have the highest binding affinity among the ligands synthesized. In order to elucidate the binding mode, NMR measurements were performed. In particular, it was required to show the fact that DQzG could form hydrogen bonds with cytosine. There have been several reports of NMR-based binding analysis of nucleic acids and small molecules. Our research group has clarified the structures of naphthyridine-azaquinolone (NA) complexed with DNA CAG/CAG and naphthyridine carbamate dimer (NCD) complexed with RNA UGGAA/UGGAA.[1,2] Based on this information, we analyzed the complexes formed.

Since the imino-proton peaks of the nucleic acid duplex structure exist at upfields and is easy to distinguish with other peaks. Therefore, DQzG was added to the G2C4/G2C4 DNA duplex and the changes in the imino-proton region were analyzed. The samples were measured in water at 700 Hz at concentrations of 0, 1, 2, 3, and 4 ratios of ligand to DNA duplex. (Fig. 9) The peak of A-T base pair appeared at 13-14 ppm and the peak of G-C base pair at 12-13 ppm. As the concentration of the ligand was increased, the signal in the 12-14 ppm region increased. When the concentration was increased to 1:4, the change slowed down and the signal of DNA only disappeared and became saturated. This change suggested that the ligand bound to G2C4/G2C4 at 1:4 without non-specific binding to the other part of the duplex. Furthermore, at 1:1, the newly emerged signal was similar to the signal at 1:4 ratio. This result suggested that DNA-4DQzG complex have already formed at 1:1. In addition, when the concentration of the ligand was increased, several new peaks appeared around 10-12 ppm. Although there were several peaks, the two peaks at 11.8 and 10.3 ppm were particularly intense.

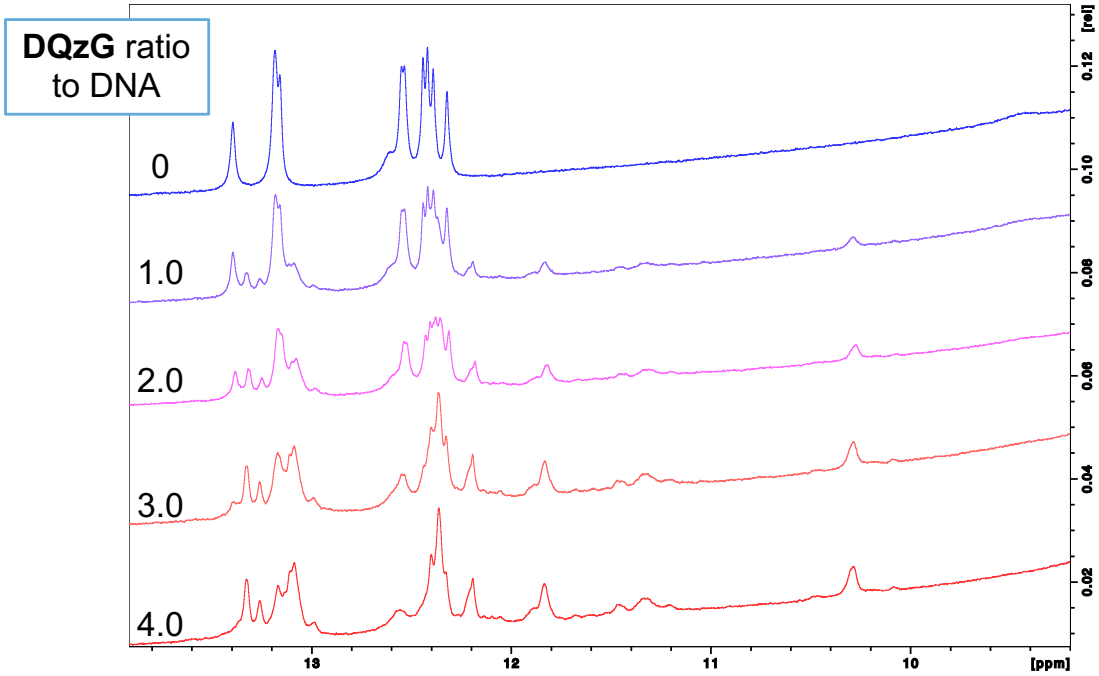


Fig. 9 ^{1}H NMR of GGCCCC/GGCCCC motif containing DNA duplexes (300 μM), without ligand (blue line) and DQzG (300, 600, 900, 1200 μM) in 20 mM phosphate buffer pH 6.5, 50 mM NaCl and 10% D_2O.

The NOE measurement with the 1:4 sample revealed that the newly emerged signals at 11.8 and 10.3 ppm were due to the presence of signals at 5.8 and 7.5 ppm. (Fig. 10) The newly emerged signals indicated that DQzG and the amino group at position 4 of cytosine formed hydrogen bond. Furthermore, the nucleotide region and NOE of this DQzG-C hydrogen bond pair were present in both the G-C and A-T base pairs. This result suggested that the DQzG-C hydrogen bond pair existed between Watson-Crick type G-C and A-T base pairs. The binding information revealed using NMR analysis was shown below.

1. The G2C4/G2C4-DQzG complex has a stoichiometric ratio of 1:4
2. DQzG forms a hydrogen bond with cytosine
3. The DQzG-cytosine pair is sandwiched between G-C and A-T base pairs

These results provided the tentative structure for the G2C4-DQzG complex. Further NMR analysis and computational method should be used to obtain more detailed information about the complex. (Fig. 11)

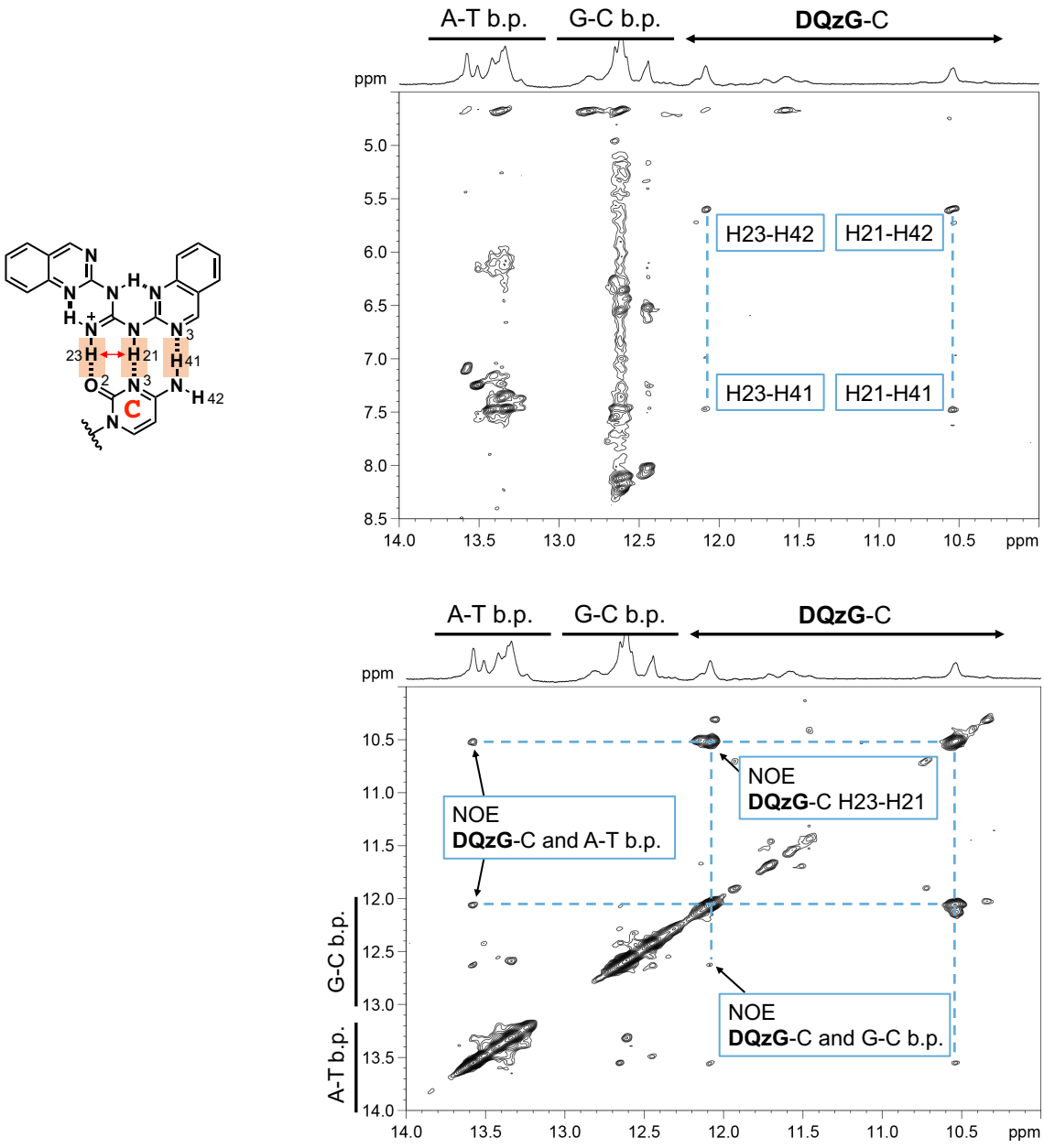


Fig. 10 NOESY of GGCCCC/GGCCCC motif containing DNA duplexes (300 µM), with DQzG (1200 µM) in 20 mM phosphate buffer pH 6.5, 50 mM NaCl and 10% D_2O.

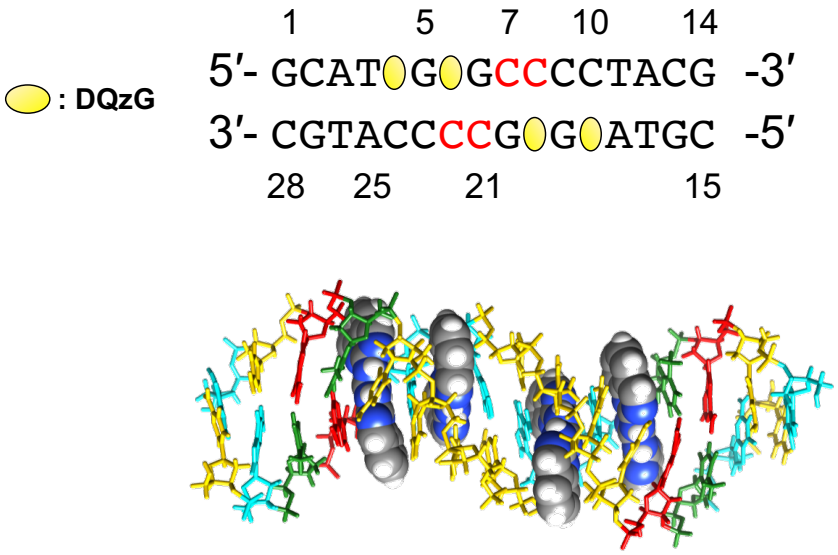


Fig. 11 Tentative structure of the complex (GGCCCC/GGCCCC motif containing DNA duplexes and DQzG)

Conclusion

In chapter 1-1, the structural binding factors of DQG was extracted. In chapter 1-2, based on this information, a new molecule binding to G2C4/G2C4 DNA, DQzG was designed. DQzG was a structure with more nitrogen in the quinoline ring than DQG and formed many planar conformations by intramolecular hydrogen bonds. In addition, all planar conformations have the DDA surface, which can match the AAD surface of cytosine. Binding evaluation with G2C4/G2C4 motif DNA revealed that DQzG had the higher binding affinity than DQG. Furthermore, NMR complex analysis showed that DQzG could form hydrogen bonds with cytosine. Binding information obtained by NMR revealed the tentative structure of the G2C4/G2C4-4DQzG complex.

Perspective

Based on the previous binding information, molecules that binds to G2C4/G2C4 more strongly were designed. It suggested that four molecules bound to G2C4/G2C4 due to the planarity of the entire molecule, the stacking interaction with the 3'-end base on the major groove side, and the hydrophobic interaction on the minor groove side. Based on these findings, the following are some ideas for the design of new molecules to improve the binding ability.

1. Stacking and intercalation ability
2. Electrostatic interaction with phosphoric acid
3. Dimerization
4. Steric anchoring of ligands
5. Covalent attachment of ligand

The following are specific examples.

1. Stacking and intercalation ability

A common feature among DQzG and other binding molecules is the planarity of the entire molecule, suggesting that it can intercalate into nucleic acids and be stabilized by adjacent nucleobases. To improve its ability, we proposed the addition of a benzene ring to quinazoline.[3,4] By extending the benzene ring at positions 5-8 of the quinazoline, the stacking area with bases on the major groove was improved. Intercalation to non-specific nucleic acids is a concern and may decrease G2C4/G2C4 selectivity.[5]

2. Electrostatic interaction with phosphoric acid

When DQzG form hydrogen bonds with cytosine, there is a phosphate backbone near the aromatic rings of the ligand. Functional groups that can form electrostatic interactions with phosphate are amines and guanidine. In the past, they have been added to the ends of molecules in order to improve the binding strength of small molecules.[6,7] The addition of cationic functional groups is expected to improve the water solubility of the molecule.[8]

3. Dimerization

Dimerization of the ligand may reduce the loss of entropy of the ligand upon binding and may also improve repeat selectivity by using an appropriate linker length.[7,9,10] DQzG can form multiple planar structures, suggesting that DQzG can bind to cytosine in the appropriate planar conformation. In other words, the type of linker could be tolerated due to flexible recognition site on DQzG. Considering the modeling and synthesis, extending the linker from position 4 was considered to be the least obstructive. It is also possible to add water-soluble functional groups or functional groups that interact electrostatically with the phosphate backbone to the linker.

4. Steric anchoring of ligands

DQzG can form multiple planar structures and all of the conformations can bind to cytosine thorough hydrogen bonds. With multiple planar conformations, it was assumed that the suitable conformation will be selected for binding to G2C4/G2C4 via conformational selection. However, it was not identified which conformations can bind to the four cytosines of G2C4/G2C4. Fixing the possible conformations which DQzG could form was considered to conserve the energy of

conformation selection before binding to nucleic acid.[11] Linking two quinazolines in a molecule with a linker physically restricts the formation of multiple planar conformations. When two quinazoline rings are connected by a linker, the formation of multiple planar configurations will be restricted, but the linker connected to cytosine during hydrogen bonding may be in close proximity to the phosphate backbone, which requires reconsideration.

5. Covalent attachment of ligands

There have been reports that some repeat-binding molecules were multimerized after binding, through the linker from the recognition molecule forming covalent bonds with the adjacent recognition molecule in order to bind more tightly to repeat DNA/RNA.[12,13] DQzG was able to be extend by the linker. This molecule protruded in the major and minor grooves in binding to G2C4/G2C4. If the linker is covalently attached to the two quinazolines so that it surrounds the phosphate after hydrogen bonding with cytosine, dissociation will be suppressed. However, in this case, after the covalent bond is formed, it cannot dissociate and will affect the target genomic DNA.

There are other proposed improvements such as improving the hydrogen-binding capacity of cytosine and increasing the binding point of cytosine to amines.

Experiments have also shown that this ligand binds to DNA-G2C4/G2C4 but not to RNA-G2C4/G2C4. The above proposed improvement is a useful molecular design guideline for binding to repeat RNA in the future. I hope that the molecular design will be useful for the future development of repeat-binding molecules.

Materials and Methods

General

Reagents and solvents were purchased from standard suppliers and used without further purification. Reactions were monitored with TLC plates precoated with Merck silica gel 60 F254. Wako gel C-200 was used for silica gel flash chromatography. ^{1}H-NMR spectra were measured with BRUKER Avance Ⅲ 700, JEOL JNM-LA400 and JEOL LA600. Coupling constants (J values) are represented in hertz. The chemical shifts are expressed in ppm relative to residual solvent as an internal standard. ^{13}C-NMR spectra were measured with JEOL LA600 and BRUKER Avance Ⅲ 700. DNAs were purchased from Life Technologies.

UV-melting analysis

Thermal denaturation profiles were recorded on a UV-2700 spectrophotometer (Shimadzu) equipped with a TMSPC-8 temperature controller and a 10 mm path-length cell. The absorbance of DNAs (5 μM DNA duplex) with ligand (20 μM) in phosphate buffer (10 mM, pH 7.0) containing NaCl (100 mM) and 5% DMSO was monitored at 260 nm from 2 to 80 ºC or 100 °C (1 °C min^{-1}). T_m was calculated by using the median method.

Circular Dichroism (CD) Measurements

CD experiments were performed with a J-725 CD spectrometer (JASCO) using a 10 mm path length cell. CD spectra of DNAs (5 μM DNA duplex) in the absence and presence of ligand (20

μM) were measured in phosphate buffer (10 mM, pH 7.0) containing NaCl (100 mM) and 5% DMSO.

Surface Plasmon Resonance (SPR)

SPR single cycle kinetics assay was performed using Biacore T200 platform (GE Healthcare, Life Science). Immobilization of the oligomers on the Series S sensor chip SA surface was carried out using avidin–biotin coupling in HBS-EP^{+} buffer (10 mM pH 7.4 HEPES, 150 mM NaCl, 3 mM EDTA, 0.005% surfactant P20). Concentration of 5′-biotinylated oligonucleotides were adjusted in 10 mM HEPES-500 mM NaCl. Biotinylated 5′-GCATGGCCCCTACGTTTTCGTAGGCCCCATGC-3′ was immobilized on SA surface via avidin–biotin. Blank immobilization was performed in the flow cell 1 to permit reference subtraction. Ligand solution was diluted using HBS-EP+ buffer. Sensorgrams were obtained in the ligand concentrations of 0.25, 0.50, 1.0, 2.0, 4.0 μM and 0.025, 0.050, 0.1, 0.2, 0.4 μM in single-cycle mode (contact time 30 s, Dissociation time 120 s, Flow rate 60 μM / min). All sensorgrams were corrected by reference subtraction of blank flow cell response and buffer injection response.

References

1. Nakatani, K. *et al.* Small-Molecule Ligand Induces Nucleotide Flipping in (Cag)n Trinucleotide Repeats. *Nat. Chem. Biol.* **1**, 39–43 (2005).
2. Shibata, T. *et al.* Small molecule targeting r(UGGAA)n disrupts RNA foci and alleviates disease phenotype in Drosophila model. *Nat. Commun.* **12**, 1–13 (2021).
3. Li, J. *et al.* Naphthyridine-Benzoazaquinolone: Evaluation of a Tricyclic System for the Binding to (CAG)n Repeat DNA and RNA. *Chem. - An Asian J.* **11**, 1971–1981 (2016).
4. Murata, A., Otabe, T., Zhang, J. & Nakatani, K. BzDANP, a Small-Molecule Modulator of Pre-miR-29a Maturation by Dicer. *ACS Chem. Biol.* **11**, 2790–2796 (2016).
5. Almaqwashi, A. A., Paramanathan, T., Rouzina, I. & Williams, M. C. Mechanisms of small molecule-DNA interactions probed by single-molecule force spectroscopy. *Nucleic Acids Res.* **44**, 3971–3988 (2016).
6. Eubanks, C. S., Forte, J. E., Kapral, G. J. & Hargrove, A. E. Small molecule-based pattern recognition to classify RNA structure. *J. Am. Chem. Soc.* **139**, 409–416 (2017).
7. Matsumoto, J. *et al.* The Dimeric Form of 1,3-Diaminoisoquinoline Derivative Rescued the Mis-splicing of Atp2a1 and Clcn1 Genes in Myotonic Dystrophy Type 1 Mouse Model. *Chem. - A Eur. J.* **26**, 14305–14309 (2020).
8. Umemoto, S. *et al.* Structure-activity studies on the fluorescent indicator in a displacement assay for the screening of small molecules binding to RNA. *Chem. - A Eur. J.* **18**, 9999–10008 (2012).

9. Li, J. *et al.* A Dimeric 2,9-Diamino-1,10-phenanthroline Derivative Improves Alternative Splicing in Myotonic Dystrophy Type 1 Cell and Mouse Models. *Chem. - A Eur. J.* **24**, 18115–18122 (2018).

10. Peng, T., Murase, T., Goto, Y., Kobori, A. & Nakatani, K. A new ligand binding to G-G mismatch having improved thermal and alkaline stability. *Bioorganic Med. Chem. Lett.* **15**, 259–262 (2005).

11. Mukherjee, S., Dohno, C., Asano, K. & Nakatani, K. Cyclic mismatch binding ligand CMBL4 binds to the 5′-T-3′/5′-GG-3′ site by inducing the flipping out of thymine base. *Nucleic Acids Res.* **44**, 7090–7099 (2016).

12. Yamada, T., Miki, S., Ul'Husna, A., Michikawa, A. & Nakatani, K. Synthesis of Naphthyridine Carbamate Dimer (NCD) Derivatives Modified with Alkanethiol and Binding Properties of G-G Mismatch DNA. *Org. Lett.* **19**, 4163–4166 (2017).

13. Rzuczek, S. G., Park, H. & Disney, M. D. A toxic RNA catalyzes the in cellulo synthesis of its own inhibitor. *Angew. Chemie - Int. Ed.* **53**, 10956–10959 (2014).

Chapter 1-3

Effects of G2C4 Repeat Binding Molecule on the transcription of G4C2/G2C4 Repeat

Introduction

Aberrant expansion of the non-coding GGGGCC(G4C2) repeat sequence in the first intron of the *C9orf72* gene was reported as the cause of amyotrophic lateral sclerosis with frontotemporal dementia (ALS/FTD).[1,2] The expanded repetitive sequence associated with ALS/FTD from 100 to 4500 with the majority of normal individuals having only 30 or fewer bases.[3] G4C2 repeat in the intron, which was supposed to be truncated by splicing process, was transcribed and the repeat dipeptide was produced via RAN translation.[4] It was also reported that *C9orf72*-mRNA retaining the G4C2 repeat-containing intron was present after splicing processing in lymphoblasts from G4C2 repeat expanded *C9orf72* carriers.[5] Altering the biological processes of transcription or splicing by small molecules may provide new insights.

・Transcription in G4C2/G2C4 Repeat

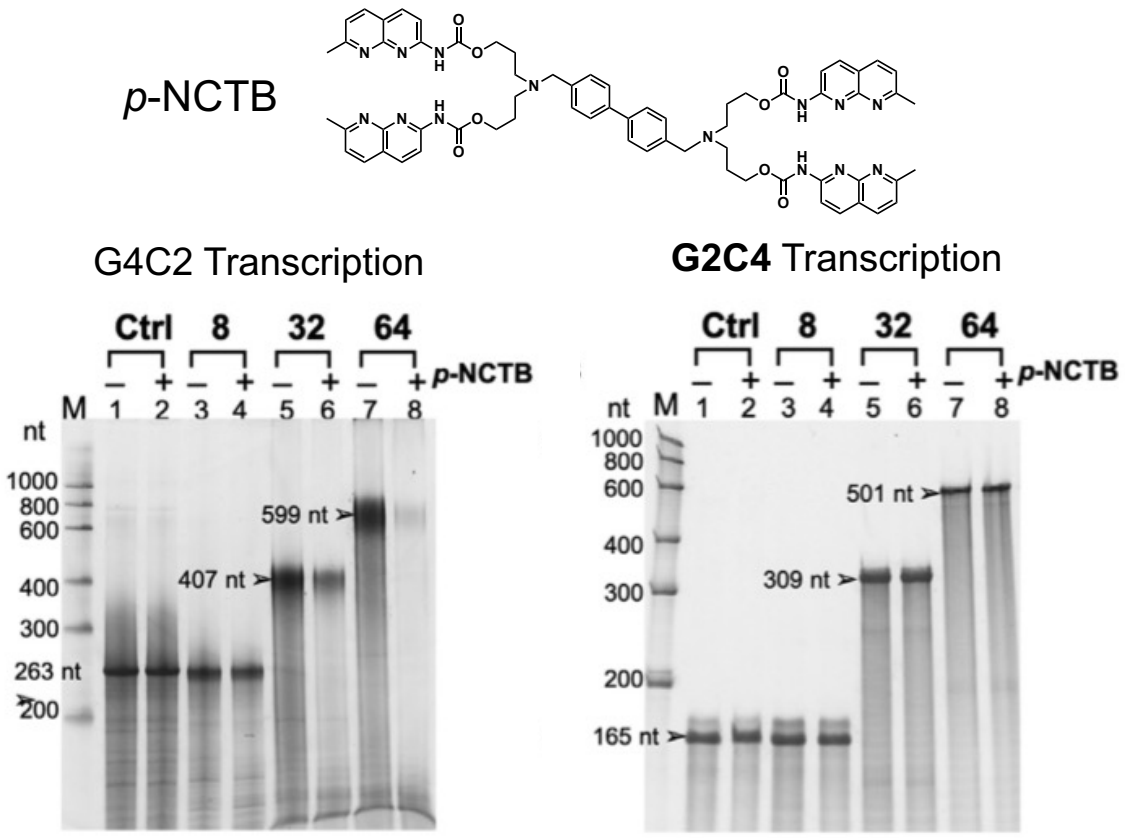
Introns containing these large expansions were transcribed and RNA containing their repetitive sequences were identified to cause cytotoxicity in two pathways. First pathway was that the protein transcribed sense (G4C2)n and/or antisense (G2C4)n RNA can form RNA foci to alter original protein function. Second was the production of toxic dipeptide repeat proteins (DPRs) from the repeat-containing RNA via repeat-associated non-AUG (RAN) translation.[6–8]

Consequently, neutralization/degradation of toxic RNA G2C4 repeats will become a promising direct therapeutic strategy for the control of ALS/FTD. Inhibiting the transcription of repeat DNA was a beneficial option for suppressing the formation of harmful RNA repeats from their origin.

・Effects of G4C2 Repeat-binding Molecules on Transcription

The naphthyridine tetramers (NCTX), *p*-NCTB, was developed by dimerizing the G4C2 repeat-binding molecule naphthyridine carbamate dimer (NCD).[9,10] *In vitro* transcription experiments were performed using double-stranded templates with 64, 32, and 8 repeats of G4C2/G2C4. In case of transcription producing RNA G4C2 repeats using G4C2 templates as the sense strand, *in vitro* transcription was significantly inhibited using *p*-NCTB in the 64-repeat template compared to the template sequence containing short repeats or not. (Fig. 1) The results of G4C2 transcription inhibition suggested a model for inhibition. However, these molecules can bind to both DNA and RNA of G4C2 repeats, which complicated the identification of the inhibition mechanism.

In contrast to previously reported molecules, DQzG, a molecule that can bind to the opposite side chain, G2C4 repeat, to observe the effects of transcription. If the relationship between the higher-order structure of the repeats and the function of the repeats by the repeat-binding molecules by examining the effects on the G2C4 repeats are clarified, it will contribute to elucidate the function of the sense and antisense chains in the pathogenesis of ALS/FTD. In particular, there are few systems that can control the higher-order structure of bilateral chains by G4C2 and G2C4 repeat-binding molecules, which will be useful for elucidating new molecular mechanisms for repeat diseases.



Y. Lu, C. Dohno, K. Nakatani, *Chem. Commun.* **2020**, *56*, 754-757.

Fig. 1 Denaturing PAGE (5%) analysis of *p*-NCTB's effect on G4C2 transcription. Transcription tube (10 μL) containing 5 ng PCR product, and *p*-NCTB: 4 μM was kept at 37 °C for 2 h using T7-Scribe Standard RNA IVT kit (CellScript) followed by DNase I treatment. Denaturing PAGE (5%) was run at 200 V for 60 min at room temperature. Lane M: RiboRuler low range RNA ladder (ThermoFisher).

Results & Discussion

・ Preparation of Transcription Templates

The template used for *in vitro* transcription was previously reported templates.[9] The DNA templates were prepared from plasmids with different repeat lengths, (G4C2)8, (G4C2)32, and (G4C2)64, and a short fragment (48 b.p.) derived from GFP sequence was used as control DNA. This template contained the T7 promoter sequence, which was recognized by T7 RNA polymerase to initiate transcription *in vitro*. A sequence near the G4C2 repeat in *C9orf72* gene was also inserted to reproduce the *C9orf72* gene. The PCR products for transcriptional evaluation were four types of templates: control template (without G4C2 repeats) with T7 promoter sequence upstream, template 8×, 32×, 64× with (G4C2)8, 32, 64, and four types of templates were used: control template (without G2C4 repeats), template 8×, 32×, 64× with (G2C4)8, 32, 64.

・Effect of DQzG on *in vitro* Transcription

Next, the biological effect of the ligands was evaluated on the *in vitro* transcription of repeat-containing DNA. PCR products containing G4C2 units of different lengths and control sequences were used as templates. In this case, the (G4C2)n strand was the coding strand and the (C4G2)n strand was used as the template strand to produce RNA (G4C2)n (sense transcription). The T7-Scribe Standard RNA IVT Kit contained T7 RNA polymerase (RNAP), NTPs, and buffer solution, and other components necessary for transcription were used.

DQzG was added to the transcription solution to 10 μM and the reaction was performed at 37 ºC for 2 h. Then DNaseI was added and heated to digest the template DNA. The transcription was performed through PAGE 5% and the transcript levels were analyzed. (Fig. 2) In the transcription reaction of G4C2 repeats, there was no change in the concentration of the bands in the control, 8×, and 32× with or without ligand. 64× showed low transcription levels with and without ligand. This may be attributed to the small amount of template DNA products. In the transcription reaction of G2C4 repeats, there were also no change in the amount of transcription in the presence or absence of the ligand, as well as in the transcription reaction of G4C2 repeats. The results suggested that DQzG did not affect the transcription of G4C2/G2C4 repeat DNA. The experiments with higher concentrations of the ligand in the reaction will be the final determination.

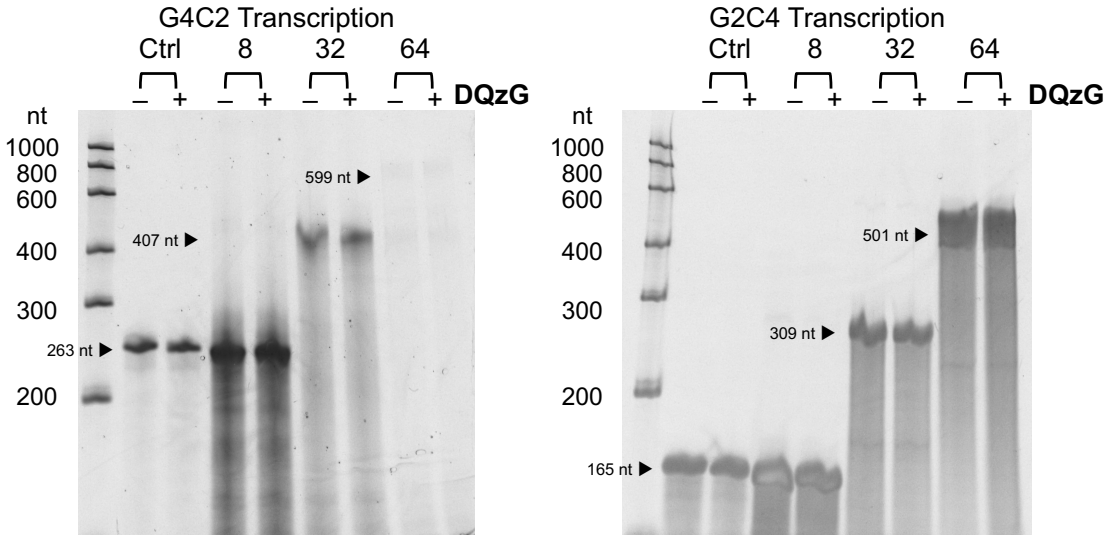


Fig. 2 Denaturing PAGE (5%) analysis of DQzG's effect on G4C2 transcription. Transcription tube (10 μL) containing 5 ng PCR product, and DQzG: 10 μM was kept at 37 °C for 2 h using T7-Scribe Standard RNA IVT kit (CellScript) followed by DNase I treatment. Denaturing PAGE (5%) was run at 200 V for 45 min at room temperature. Lane M: RiboRuler low range RNA ladder (ThermoFisher).

The difference between *p*-NCTB, which inhibited transcription, and DQzG, which did not, was that DQzG binds to both repeat DNA and RNA, or only to repeat DNA. DQzG did not bind to RNA, but only to DNA. On the other hand, *p*-NCTB bound to both DNA and RNA and across both, suggesting that it inhibited transcription by binding to template DNA and transcript RNA.

Conclusion

The effect of G4C2/G2C4 repeats on transcription was performed to investigate the biological consequences of the G2C4 repeat binding molecule. This was performed using the constructed *in vitro* transcribable G4C2/G2C4 repeat template. DQzG was added to the reaction solution and the transcription reaction was checked on PAGE for the amount of RNA transcribed from the G4C2 control, 8, 32, 64 repeats, and the G2C4 control, 8, 32, 64 repeats. The PAGE results showed no change in the amount of RNA transcribed in the presence or absence of DQzG. Furthermore, it was clear that DQzG did not affect the transcription of the G4C2/G2C4 repeats, although additional experiments, such as concentration-dependent, would be necessary.

Materials and Methods

In vitro transcription

The transcription reaction was performed according to the protocol of T7-Scribe Standard RNA IVT kit (CellScript) in a scale of 10 μL. The reaction tube was incubated at 37 °C for 2 h, followed by addition of 0.5 μL DNase I. The DNase I digestion time was 20 min at 37 °C. Then add 10.5 μL 2× loading buffer, heat to 90 °C for 5 min, and quickly on ice. Denaturing PAGE (5%, 7 M urea) was then run at 55 ºC, 45 min. The obtained gel was stained with SYBR Gold.

References

1. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **72**, 245–256 (2011).
2. Renton, A. E. *et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257–268 (2011).
3. Swinnen, B., Robberecht, W. & Van Den Bosch, L. RNA toxicity in non-coding repeat expansion disorders. *EMBO J.* **39**, 1–23 (2020).
4. Mori, K. *et al.* The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science (80-.).* **339**, 1335–1338 (2013).
5. Niblock, M. *et al.* Retention of hexanucleotide repeat-containing intron in C9orf72 mRNA: implications for the pathogenesis of ALS/FTD. *Acta Neuropathol. Commun.* **4**, 18 (2016).
6. Cooper-Knock, J. *et al.* Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain* **137**, 2040–2051 (2014).
7. Lee, Y. B. *et al.* Hexanucleotide repeats in ALS/FTD form length-dependent RNA Foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep.* **5**, 1178–1186 (2013).
8. Ash, P. E. A. *et al.* Unconventional Translation of C9ORF72 GGGGCC Expansion Generates Insoluble Polypeptides Specific to c9FTD/ALS. *Neuron* **77**, 639–646 (2013).

9. Lu, Y., Dohno, C. & Nakatani, K. Recognition of expanded GGGGCC hexanucleotide repeat by synthetic ligand through interhelical binding. *Biochem. Biophys. Res. Commun.* **531**, 56–61 (2020).

10. Lu, Y., Dohno, C. & Nakatani, K. A novel naphthyridine tetramer that recognizes tandem G-G mismatches by the formation of an interhelical complex. *Chem. Commun.* **56**, 754–757 (2020).

Chapter 2

Brain Profiling of

CAG Repeat DNA Binding Ligand

Introduction

・CAG Repeat Disease – Huntington's Disease

Huntington's Disease (HD), spinocerebellar ataxia type 1 (SCA1), SCA2, SCA3, SCA7, and SCA17 were reported as neurodegenerative diseases with aberrant expansion of CAG repeats.[1,2] In addition, longer CAG expansion in HD was associated with an earlier age of onset and accelerated disease advance.[3] The contraction of CAG repeats can lead to fundamental treatment of repeat diseases.

・Previous NA Studies

The mechanism of aberrant expansion of repeat sequences in detail is unknown. Recently, abnormal higher-order structural slip-out DNA formed by repeat sequences has been reported.[4,5] In the case of CAG/CTG repeats, it suggested that the hairpin structure of CAG/CAG and CTG/CTG containing mismatches was formed after the slip-out and its structure could be further expanded or contracted by incomplete repair by mismatch repair enzymes. Naphthridine-azaquinolone (NA) was reported as a molecule that could bind to the CAG/CAG DNA.[4] NA can form hydrogen bonds with naphthyridine-guanine and azaquinolone-adenine and two molecules can bind to one CAG/CAG unit. In addition, NMR study has shown that two cytosine bases in

CAG/CAG were flipped out of the duplex. Furthermore, in 2020, we conducted *in vivo* experiment using mouse model of Huntington's disease (R6/2 mice) with the genome mutated to abnormally expanded CAG repeats.[5] NA was administered by stereotaxic injection into the striatum of R6/2 mice and induce shortening expanded CAG repeats in the striatum. Moreover, NA induced to decrease the aggregation of *mHTT* protein, a biomarker for Huntington's disease. The binding of NA has been shown not to affect the transcription and translation. Therefore, it suggested that the CAG contraction was due to the contracting effect of NA.

・Issue

NA has the property of specifically binding to aberrantly formed DNA structures and inducing the contracting effect of CAG repeats *in vivo*. This uniquely characterized molecule would make it a potential therapeutic candidate in repeat diseases that have been difficult to treat curative therapy. On the other hand, since the distribution of NA after administration has not been clarified, direct evidence has not been able to be provided. In addition, the direct administration of NA was highly invasive to mice and Huntington's disease model mice were damaged and die after NA administration. This made it difficult to prove that the contracting of CAG repeats by NA administration leads to recovery of the model mice.

Chapter 2-1

Construction of Protocol for Brain Profiling

Experimental Design

・Solution

One of the two problems needing to be solved is that the distribution of NA after administration was unknown. The distribution of drugs after administration *in vivo* is important to clarify the pharmacokinetics of other drug candidates as well as NA.

・Introduction

PET and SPECT have been used to clarify the distribution of compounds after administration. These methods are sensitive enough to detect even small amounts and can observe the samples using *in vivo* live imaging.[6]

Clearing protocol, clear, unobstructed brain imaging cocktails and computational analysis (CUBIC) was a chemically and theoretically based protocol that require chemical modifications such as immobilization of small molecules in the tissue.[7–10] Using tissue clearing approach, tissues can be analyzed at the cellular level and there are examples of reports on the construction of mouse brain atlases and pathological diagnosis in humans.[11,12] A series of protocols to reveal the distribution of compounds after administration using tissue clearing techniques can be combined with other stains and accumulated pathology as well as drug distribution and will provide a new technological basis for pharmacokinetic studies.

• Design of Protocol

A protocol from administration to three-dimensional analysis was developed. (Fig. 1) The method of administration to mice is shown in Material & Method. After administration, mice were woken up spontaneously from anesthesia and kept for the predetermined period of time (except for 0 h) in order to observe drug penetration of temporal patterns. Then, in perfusion fixation, PBS and PFA was circulated throughout the body to cross-link the primary amine of the administered NA-Alexa594 with the amine of the protein via formaldehyde, causing it to lose fluidity in the body and its small molecules to become immobilized. In clearing protocol, it was necessary to be dehydrated and homogenized with MeOH before immersing the immobilized and removed the brain since the aromatic clearing reagent is immiscible with water. In addition to the high RI hydrophobic clearing reagent, benzyl alcohol and benzyl benzoate mixed organic solvent (BABB), an amino alcohol, *N*-butyldiethanolamine, which was useful as delipidation and

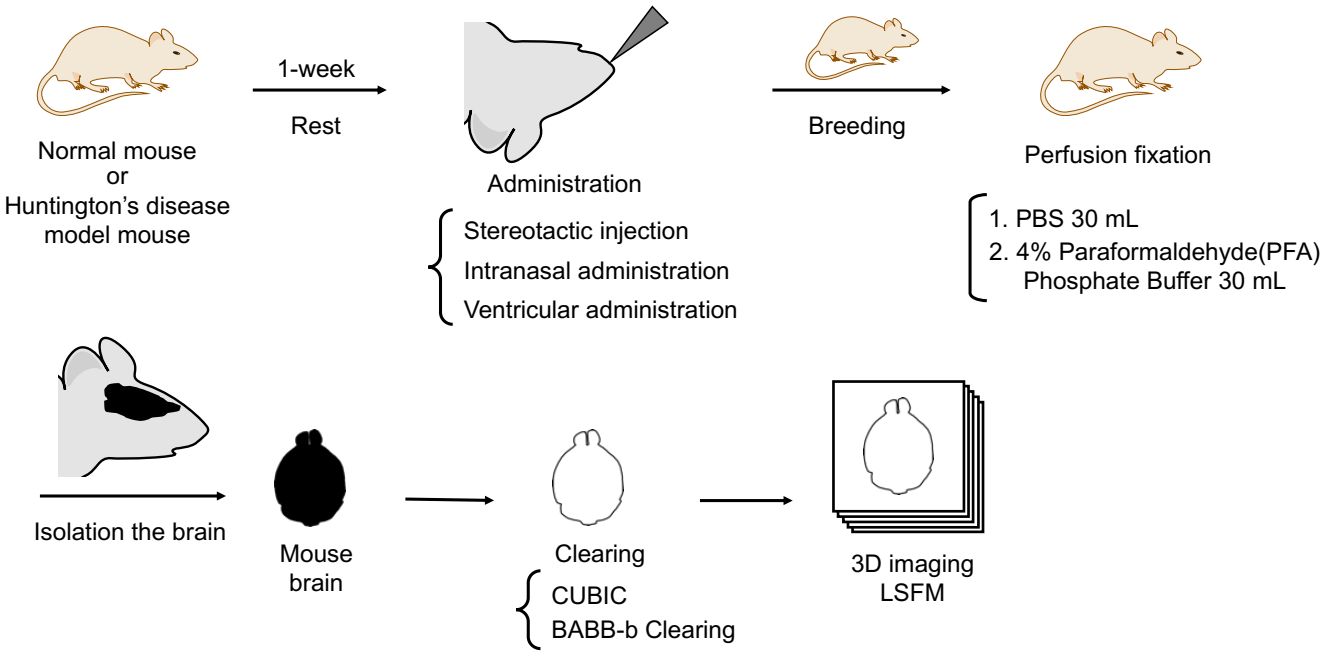


Fig. 1 The designed protocol

decolorization reagent to obtain high resolution images using LSFM, was added *N*-butyldiethanolamine to the mixed solution (BABB-b).[10] This protocol required a few reagents, high transparency, no expansion and contraction of the specimen due to infiltration of reagents and could be completed in 4 days after removal of the brain. (Fig. 2)

Fig. 2 Mouse brain image after clearing using BABB-b.

• Design of Molecule

In order to reveal the three-dimensional drug distribution in the brain using NA, a candidate drug for the treatment of HD, through the clearing technique – light sheet 3D microscopy, the rational molecular design of the administered compound was conducted. A compound, NA-Alexa594 was designed. NA-Alexa549 had 1) the NA moiety, 2) primary amine required for formaldehyde immobilization of the compound administered before clearing brain tissue, and 3) Alexa594, which has a different fluorescence wavelength from the autofluorescence of brain tissue in LSFM after clearing. (Fig. 3)

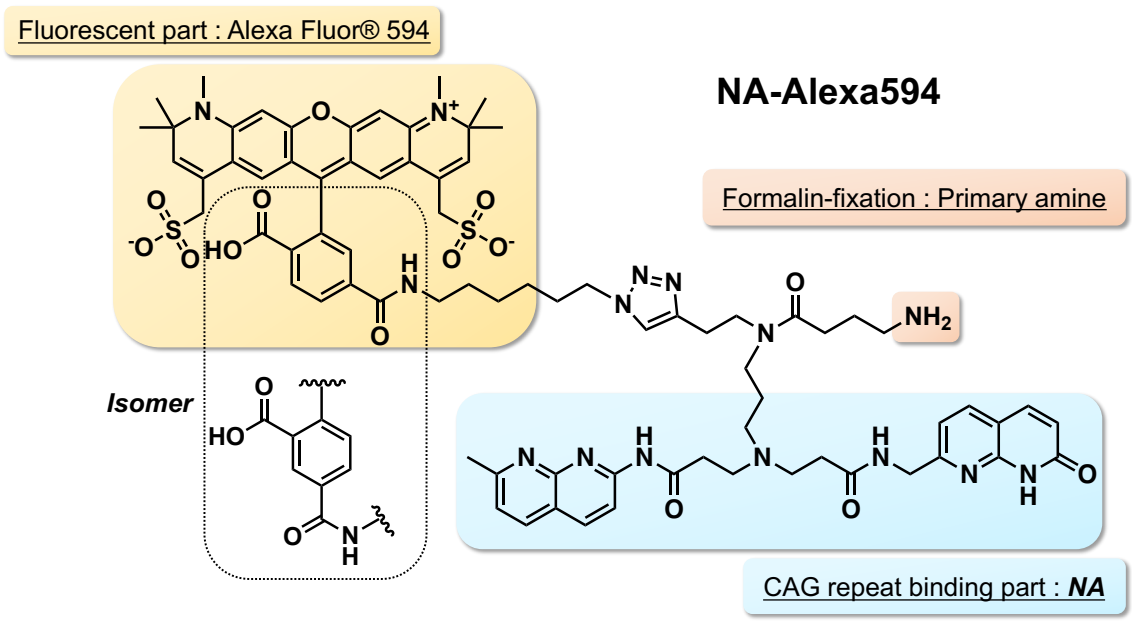


Fig. 3 Chemical structure of NA-Alexa594.

Results & Discussion

• Evaluation of Binding to CAG Repeats

The binding affinity of NA-Alexa594 to CAG repeats using surface plasmon resonance (SPR) measurement was decreased compared to unmodified NA. In order to further investigate the binding affinity, single cycle kinetics analysis was used to measure the binding affinity. (Fig. 4) Compared to the previous measurement, the binding ability was significantly lower than NA-Linker without Alexa594. The apparent K_D needs to be calculated after concentration calibration. Although lower binding affinity than unmodified NA, it was clear that NA-Alexa594 could bind to CAG repeats DNA.

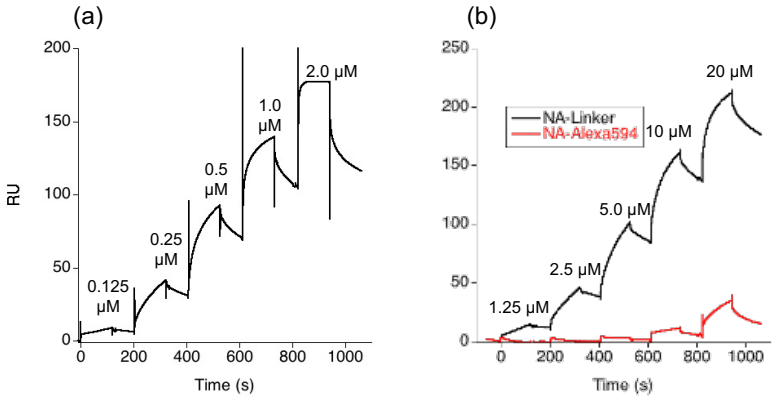


Fig. 4 SPR data for ligands binding to surface-immobilized CAG 9 repeat structure. The ligands were sequentially added at (a) NA 0.125, 0.25, 0.5, 1.0, 2.0 μM and (b) NA-Linker and NA-Alexa594 1.25, 2.5, 5.0, 10, 20 μM

・Administration to Brain Parenchyma

For the purpose of elucidating whether the ligand could be analyzed using the established protocol, NA-Alexa594 was administered to mice in the brain parenchyma anterior to the left ventricle, after which the mice were immediately immobilized and the brain was removed. Then, it was cleared. After the dehydration step with MeOH, BABB-b treatment was performed and the brain was found to be transparent. In LSFM analysis, apart from autofluorescence of the mouse brain at 532 nm, compound-derived fluorescence at 594 nm was detected at the injection point. (Fig. 5) These results indicated that NA-Alexa594 administered into the brain parenchyma can be

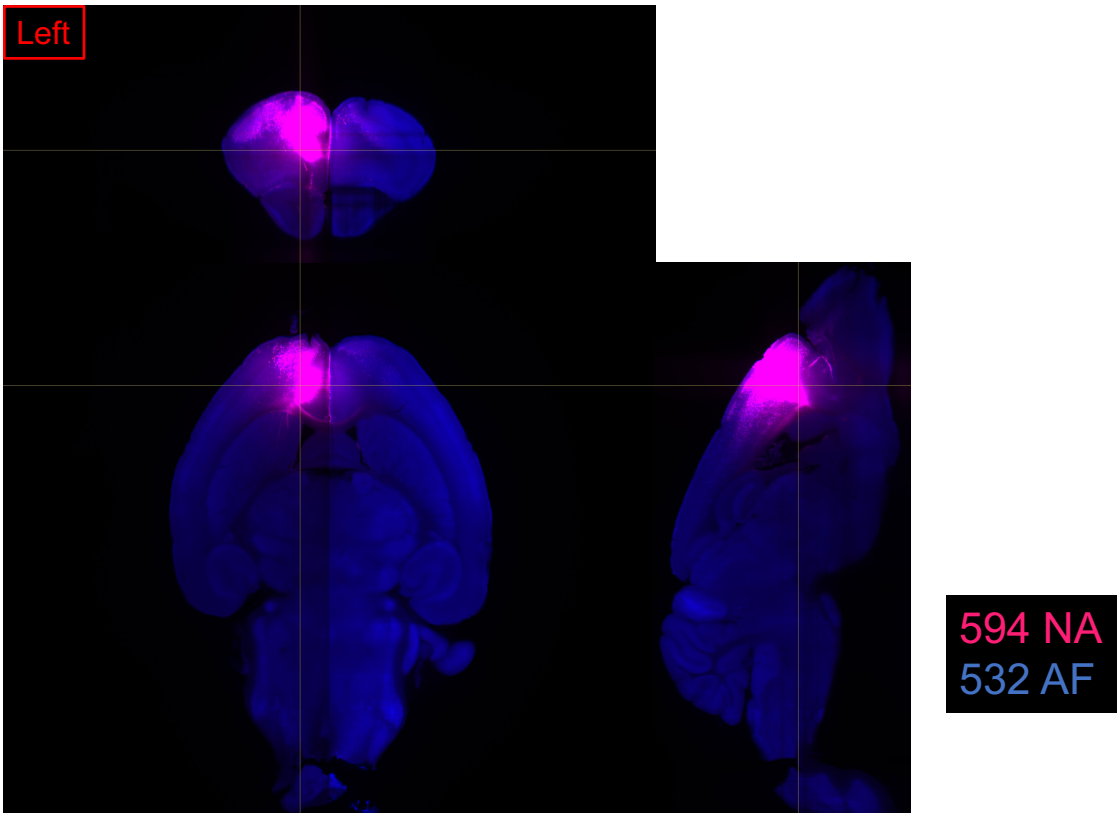


Fig. 5 LSFM image of a brain. Intraparenchymal injection of NA-Alexa594 solution, no breeding time after the injection.

detected by LSFM using the transparency protocol established in this study. In the brain parenchyma, NA-Alexa594 did not spread to a large extent in the brain.

Conclusion

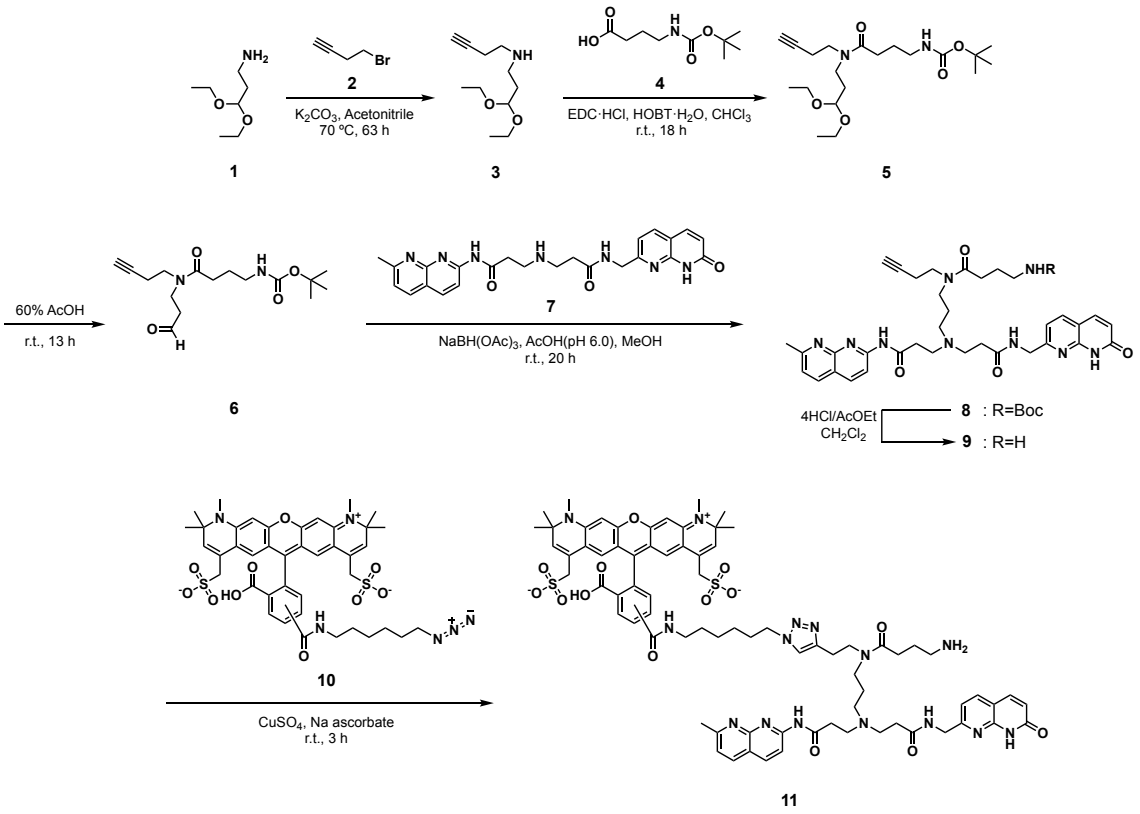
The designed NA-Alexa594 was found to be detectable using LSFM in the mouse brain cleared by BABB-b. The image obtained by LSFM showed that NA-Alexa594 was detected and remained in and around the injection point after left-intracerebral administration.

Materials and Methods

General

Reagents and solvents were purchased from standard suppliers and used without further purification. Reactions were monitored with TLC plates precoated with Merck silica gel 60 F254. Wako gel C-200 was used for silica gel flash chromatography. ^{1}H-NMR spectra were measured with BRUKER Avance Ⅲ 700, JEOL JNM-LA400 and JEOL LA600. Coupling constants (J values) are represented in hertz. The chemical shifts are expressed in ppm relative to residual solvent as an internal standard. ^{13}C-NMR spectra were measured with JEOL LA600 and BRUKER Avance Ⅲ 700. DNAs were purchased from Life Technologies.

Synthesis of NA-Alexa594



NA was synthesized by the previously reported method.

N-(3,3-Diethoxypropyl)but-3-yn-1-amine **3**

To a solution of K_2CO_3 (6.2 g, 45 mmol) in CH_3CN 30 mL was added 1-amino-3,3-diethoxypropane **1** (1.7 g, 11 mmol) and 4-bromo-1-butyne **2** (1.5 g, 11 mmol). The mixture was heated at 70 ºC for 63 h. The resulting solution was filtered to the solid by vacuo filtration with chloroform. The filtrated solution was extracted with chloroform and washed with water and brine. The organic layer was dried over $MgSO_4$ and evaporated to dryness. The crude compound was purified by column chromatography (Hex/AcOEt = 1:1~1:0) to yield pure **3** (1.2 g, 52%).

1H-NMR (600 MHz, CDC l ₃) δ = 4.59 (t, J = 5.6 Hz, 1H), 3.68-3.63 (m, 2H), 3.53-3.48 (m, 2H), 2.77 (t, J = 6.6 Hz, 2H), 2.71 (t, J = 7.0 Hz, 2H), 2.39 (td, J = 6.6, 2.7 Hz, 2H), 1.98 (t, J = 2.7 Hz, 1H), 1.83 (td, J = 6.9, 5.7 Hz, 2H), 1.22-1.19 (m, 6H)

13C-NMR (151 MHz, $CDCl_3$) δ = 101.9, 82.6, 69.4, 61.2, 48.1, 45.1, 33.9, 19.6, 15.3

HRMS (ESI): Calculated. for $C_{11}H_{21}NO_2$: 200.1645 $[M+H]^+$, Found: 200.1643 $[M+H]^+$.

5

A mixture of *N*-(tert-butoxycarbonyl)-4-aminobutyric acid **4** (1.0 g, 5.0 mmol) and 1-hydroxybenzotriazole monohydrate (2.3 g, 15 mmol) in 40 mL chloroform was stirred for 5 min. And **3** and 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride were added and the mixture stirred at ambient temperature for 18 h. 1M HCl (24 mL) was added, the mixture was extracted with chloroform and washed with saturated aqueous sodium bicarbonate and water. The organic layer was dried over $MgSO_4$ and evaporated to dryness. The crude compound was purified by column chromatography ($CHCl_3$/MeOH = 1:0~50:1) to yield pure **5** (0.69 mg, 36%).

1H-NMR (600 MHz, $CDCl_3$, mixture of rotamers) δ = 4.88 (br, 1H), 4.45 (t, J = 5.5 Hz, 1H), 3.60 (t, J = 7.0 Hz, 2H), 3.46-3.31 (m, 6H), 3.11 (q, J = 6.2 Hz, 2H), 2.41-2.32 (m, 4H), 1.96 (t, J = 2.7 Hz, 1H), 1.84-1.76 (m, 4H), 1.37 (s, 9H), 1.17-1.13 (dt, J = 6.9 Hz, 6H)

13C-NMR (151 MHz, $CDCl_3$) δ 172.5, 172.4, 156.1, 101.1, 100.4, 81.9, 80.5, 78.9, 71.0, 69.7, 61.6, 61.3, 46.5, 45.1, 44.4, 42.2, 40.2, 33.0, 31.8, 30.4, 30.3, 28.4, 25.2, 18.8, 17.5, 15.3

HRMS (ESI): Calculated. for $C_{20}H_{36}N_2O_5$: 385.2697 $[M+H]^+$, Found: 385.2696 $[M+H]^+$.

Linker-Boc **6**

A mixture of **5** (0.28 g, 0.73 mmol) and 60% acetic acid (3 mL) was stirred at ambient temperature for 13 h. Add saturated sodium carbonate to adjust to pH 11. The mixture was extracted with chloroform and washed with water and brine. The organic layer was dried over $MgSO_4$, evaporated to dryness and yield **6** (93.4 mg, 41%).

1H-NMR (600 MHz, $CDCl_3$, mixture of rotamers) δ = 9.81-9.79 (t, J = 1.2 Hz, 1H), 4.84-4.75 (s, 1H), 3.72-3.60 (t, J = 7.1 Hz, 1H), 3.52-3.44 (t, J = 6.9 Hz, 2H), 3.19-3.14 (m, 3H), 2.84-2.78 (td, J = 6.5, 1.1 Hz, 2H), 2.48-2.37 (m, 5H), 2.06-2.04 (t, J = 2.6 Hz, 1H), 1.86-1.80 (m, 2H), 1.66 (s, 2H), 1.43 (s, 9H)

HRMS (ESI): Calculated. for $C_{16}H_{26}N_2O_4$: 333.1785 $[M+Na]^+$, Found: 333.1785 $[M+Na]^+$.

NA-Linker **9**

Dissolved **6** and **NA 7** (93 mg, 0.20 mmol) in MeOH 6 mL. The solution was adjusted to pH 6 by adding acetic acid. And sodium triacetoxyborohydride (0.17 g 0.41 mmol) was added and the mixture stirred at ambient temperature for 20 h. The mixture was extracted with chloroform and washed with saturated aqueous sodium bicarbonate and water. The organic layer was dried over $MgSO_4$ and evaporated to dryness. To a solution of the crude compound in 4N-HCl/AcOEt was added in CH_2Cl_2 at ambient temperature. The resulting mixture was removed the solvent, diluted with H_2O and purified by HPLC (eluent: CH_3CN and 0.1% AcOH) to yield pure NA-Linker **9** (24 mg, 17%).

1H-NMR (400 MHz, $CDCl_3$, mixture of rotamers) δ = 11.25-11.47 (1H), 11.00-11.24 (0H), 8.57 (s, 0H), 8.27 (d, J = 9.2 Hz, 0H), 8.24 (dd, J = 27.2, 8.9 Hz, 1H), 8.20 (d, J = 8.7 Hz, 1H), 7.86-

7.78 (m, 2H), 7.72 (d, J = 8.7 Hz, 1H), 7.64 (dd, J = 65.5, 8.2 Hz, 1H), 7.56 (d, J = 7.8 Hz, 0H), 7.42-7.34 (m, 2H), 7.11 (dd, J = 8.0, 5.3 Hz, 1H), 6.87 (dd, J = 10.8, 8.0 Hz, 1H), 6.47 (d, J = 9.2 Hz, 1H), 6.42 (dd, J = 40.5, 9.4 Hz, 1H), 6.37 (d, J = 9.6 Hz, 0H), 4.38 (q, J = 5.5 Hz, 2H), 3.53 (q, J = 7.2 Hz, 2H), 3.31 (dd, J = 10.8, 6.2 Hz, 3H), 3.21 (t, J = 5.7 Hz, 1H), 2.94 (t, J = 6.4 Hz, 1H), 2.81-2.69 (m, 9H), 3.56-1.72 (m, 35H), 2.57-2.49 (m, 7H), 2.48-2.44 (m, 1H), 2.26 (td, J = 6.8, 2.4 Hz, 1H), 2.15 (t, J = 5.7 Hz, 1H), 2.04-2.02 (m, 6H), 1.88 (t, J = 6.9 Hz, 1H), 1.83 (t, J = 2.5 Hz, 1H), 1.72 (s, 1H)

HRMS (ESI): Calculated. for $C_{35}H_{43}N_9O_4$: 654.3511 $[M+H]^+$, Found: 654.3504 $[M+H]^+$.

NA-Alexa594 **11**

Four solutions, NA-Linker **6** 1 mM (16 mg, 22 μmol), AFDye™ 594 Azide (also known as Alexa Fluor® 594 Azide, Fluoroprobes, Scottsdale, Arizona) **10** 1mM (22 mg, 23 μmol), $CuSO_4$ 100 mM (0.27 mg, 1.1 mmol) and sodium L-ascorbate (0.26 mg, 1.3 mmol) were prepared. And 10 mL of the solution each were added and stirred at ambient temperature for 3 h. The mixture was added ethylenediamine-*N,N,N′,N′*-tetraacetic acid (0.88 g, 3.0 mmol) and filtered. The resulting solution was purified by HPLC (eluent: CH_3CN and 0.1% AcOH) to yield pure NA-Alexa594 **11** (1.7 mg).

HRMS (ESI): Calculated. for $C_{35}H_{43}N_9O_4$: Calc. 1498.6082 $[M-H]^+$, Found: 1498.6096 $[M-H]^+$.

Animal

Mouse handling and experimental procedures were performed following the guidelines in Osaka University for the animal welfare and were approved by the institutional review board.

C57BL/6NJcl male mice (7-8 weeks) from CLEA Japan, Inc. were housed in groups (maximum of 5 animals/group) with ad libitum feeding, free access to water and food.

Stereotaxic administration

The mice were anesthetized and stereotaxically injected with 3.0 μL of 500 μM NA-Alexa594 dissolved in PBS (left side). Stereotaxic injections were delivered to one site within the brain parenchyma with the following coordinates: (medial–lateral = 1.1, anterior–posterior = 0.5 mm from the intersection of the inferior cerebral vein under the frontal bone with the midline, dorsal–ventral = 2.5 mm) using a 10-μL Hamilton microsyringe within 30 sec.

Perfusion fixation

The mouse was injected with PBS buffer 30 mL and 4% PFA from its left ventricle (flow rate 5 mL/min). The brain tissue was isolated from the perfused mouse. Isolated the mouse brain was immersed in 4% PFA at 4 ºC for 24 h.

Clearing

Wash

The fixed brains were washed by PBS buffer at three times for 1 h each.

Dehydration

The washed brain samples were immersed in 60% MeOH with shaking at room temperature for 4 h, followed by in 80% MeOH with shaking at room temperature for 4 h. Finally, the samples were immersed in 100% MeOH with shaking at ambient temperature overnight.

Delipidation, decolorization and RI matching

The solution exchange new 100% MeOH for 8 h. The samples were immersed in a 100:200:9 mixture of benzyl alcohol, benzyl benzoate and *N*-butyldiethanolamine (BABB-b) with shaking at ambient temperature overnight. The samples were preserved in new BABB-b solution.

References

1. Lieberman, A. P., Shakkottai, V. G. & Albin, R. L. Polyglutamine Repeats in Neurodegenerative Diseases. *Annu. Rev. Pathol. Mech. Dis.* **14**, 1–27 (2019).
2. McColgan, P. & Tabrizi, S. J. Huntington's disease: a clinical review. *Eur. J. Neurol.* **25**, 24–34 (2018).
3. Langbehn, D. R. *et al.* CAG-repeat length and the age of onset in Huntington Disease (HD): A review and validation study of statistical approaches. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **153**, 397–408 (2010).
4. Nakatani, K. *et al.* Small-Molecule Ligand Induces Nucleotide Flipping in (Cag)n Trinucleotide Repeats. *Nat. Chem. Biol.* **1**, 39–43 (2005).
5. Nakamori, M. *et al.* A slipped-CAG DNA-binding small molecule induces trinucleotide-repeat contractions in vivo. *Nat. Genet.* **52**, 146–159 (2020).
6. Shields, A. F. *et al.* Imaging proliferation in vivo with [F-18]FLT and positron emission tomography. *Nat. Med.* **4**, 1334–1336 (1998).
7. Susaki, E. A. *et al.* Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis. *Cell* **157**, 726–739 (2014).
8. Tainaka, K. *et al.* Whole-body imaging with single-cell resolution by tissue decolorization. *Cell* **159**, 911–924 (2014).
9. Kubota, S. I. *et al.* Whole-Body Profiling of Cancer Metastasis with Single-Cell Resolution. *Cell Rep.* **20**, 236–250 (2017).

10. Tainaka, K. *et al.* Chemical Landscape for Tissue Clearing Based on Hydrophilic Reagents. *Cell Rep.* **24**, 2196-2210.e9 (2018).

11. Murakami, T. C. *et al.* A three-dimensional single-cell-resolution whole-brain atlas using CUBIC-X expansion microscopy and tissue clearing. *Nat. Neurosci.* **21**, 625–637 (2018).

12. Nojima, S. *et al.* CUBIC pathology: Three-dimensional imaging for pathological diagnosis. *Sci. Rep.* **7**, 1–14 (2017).

Chapter 2-2

Intracerebroventricular and Intranasal Administration

Experimental Design

・Solution

The second problem to be solved was the ability of NA not to penetrate the blood-brain barrier (BBB), which existed in the capillaries of the brain. The presence of BBB has hindered the development of NA as well as other therapeutic candidates targeting central nervous system (CNS).[1]

・Introduction

Drug delivery to the brain has been challenging due to the presence of BBB. As usual, NA is also hard to deliver to the brain and we have demonstrated its efficacy in the striatum by brain stereotaxic administration with surgery.[2] There has been much interest in BBB permeation strategies for drugs, such as transient BBB opening,[3] chemical modification of drugs,[4] and nanoparticle-mediated delivery of therapeutically active molecules into the brain via the BBB.[5]

In this study, I focused on intranasal administration, which has been reported as an effective non-invasive administration method.[6–9] The nose-to-brain mechanism has been reported for intranasal administration and there were intracellular and extracellular pathways for delivery to the brain.[10] In addition, there are several reports on the delivery of drugs to the brain via the nose-

to-brain mechanism, including studies on the delivery of drugs to the brain by coadministration or micelles with peptides that improve cell permeability or compounds that change the structure of olfactory epithelial cells to facilitate absorption.[11,12] Intranasal administration is a promising way to deliver drugs to the brain without BBB. It has the potential to deliver compounds which does not penetrate BBB to the brain such as NA. Additionally, combining these brain delivery studies with tissue clearing techniques can provide new insights for CNS targeted drug research. For example, information on the method of administration and the brain coordinates of the administered compound can lead the optimal brain delivery method and compound structure-brain delivery.

Results & Discussion

In the experiments with the administration to both nostrils, clearing protocol was being improved, thus the procedure was different from the clearing protocol in Material & Method, but it did not significantly affect the results.

・Bilateral Nasal Administration

Intranasal administration of NA-Alexa594 was used to verify its delivery to the brain. NA-Alexa594 was administered in 2 μL doses using a micropipette through alternately both bilateral nostrils of mice until the total volume reached 24 μL. The administered mice were woken spontaneously from anesthesia and remained alive until the next day. One day after administration, the mice were sacrificed and fixed by perfusion to reveal the drug distribution in three dimensions according to the tissue clearing-LSFM protocol. (Fig. 1) In the image, the fluorescently labeled compounds were present in the olfactory bulb of the brain and especially on the surface of the olfactory bulb. Delivery to the brain via intranasal administration was possible, but it was not delivered to the more internal parts of the brain, such as the striatum.

(a) (b)

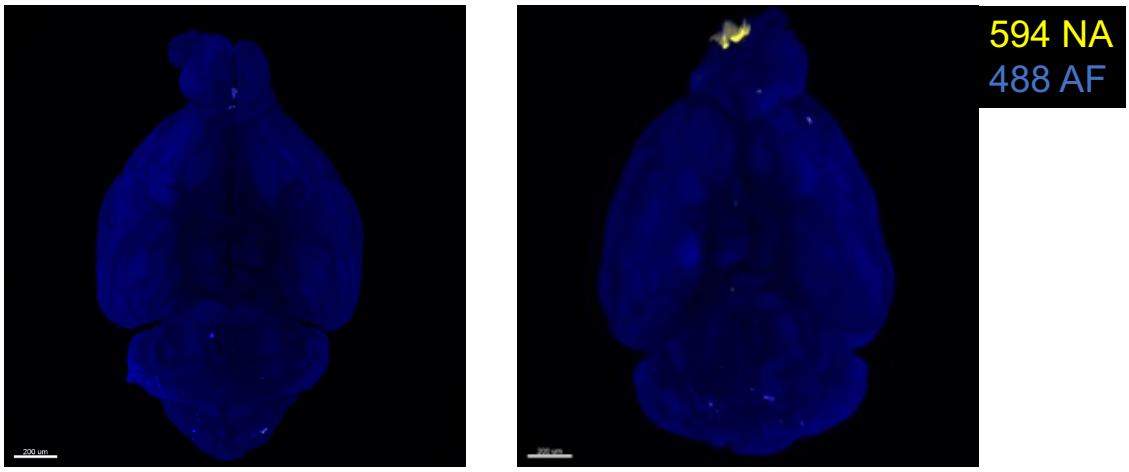


Fig. 1 LSFM image of two mouse brains (intranasal administration of only PBS (a) and NA-Alexa594 solution (b).

• Bilateral Nasal Administration – Long Term and Multiple Times

To further investigate the possible transfer to the brain by intranasal administration, multiple doses were administered over 7 days. From the first day of administration until the perfusion fixation one week later, 24 μL was administered once, three times, or five times. Tissue clearing protocol was applied and 3D imaging was obtained by LSFM. (Fig. 2) The fluorescence was observed near the olfactory bulb, while no delivery to the striatum or other region of the brain was observed.

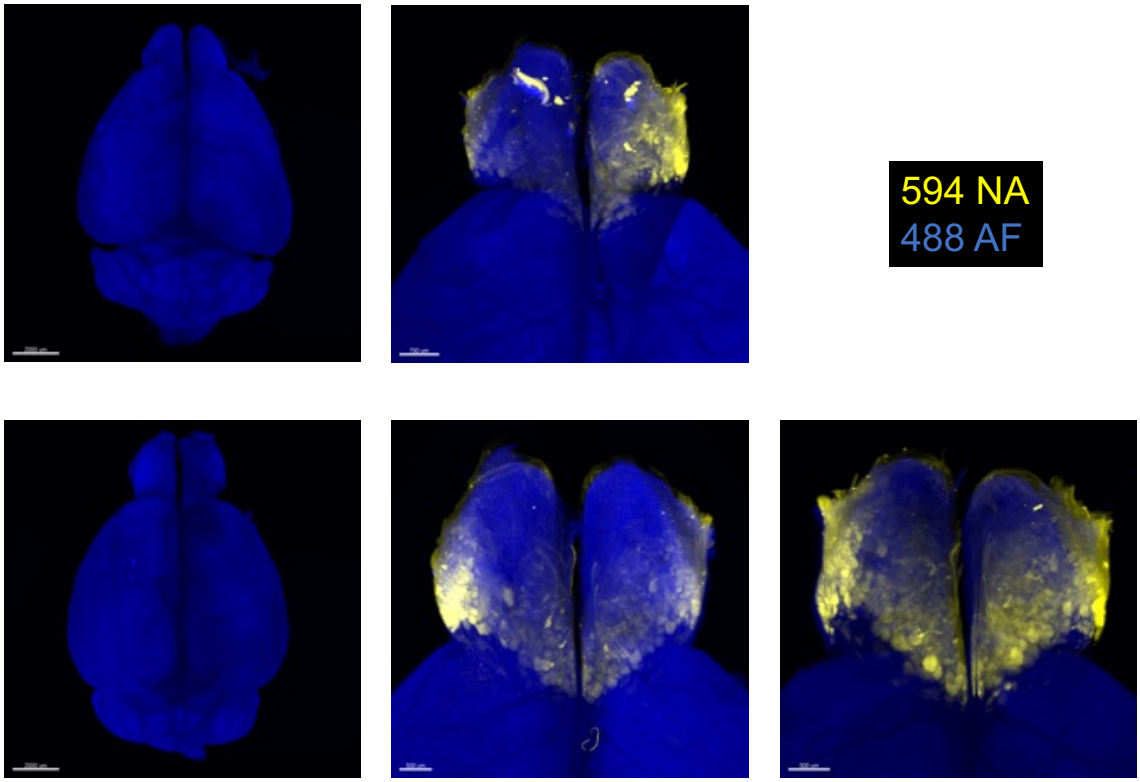


Fig. 2 LSFM image of five mouse brains (intranasal administration of only PBS(1) and NA-Alexa594 solution(2) after one day breeding and one(3), three(4) and five(5) injections NA-Alexa594 solution over 7 days.)

· Bilateral Nasal Administration – Delivery Carriers

In intranasal delivery to the brain, there are reportedly developed techniques for drug delivery. I expected to improve the efficiency of delivery to the brain by co-administering NA-Alexa594 with the available peptides *L*-penetratin and tetradecyl-β-D-maltoside (TDM). [11,12] (Fig. 3) The samples co-administered with *L*-penetratin only showed fluorescence in the olfactory bulb. (Fig. 4) The results showed a different trend from the results of the administration of NA-Alexa594 alone. The reason for changing the localization of NA-Alexa594 in the brain may be that *L*-penetratin was a peptide that increases cell permeability and co-injected NA-Alexa594 has increased cell permeability.

L-penetratin

Peptide: RQIKIWFQNRRMKWKK

Tetradecyl-β-D-maltoside (TDM)

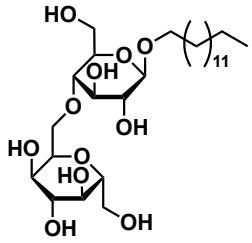


Fig. 3 Structure of Nose-to-Brain carriers

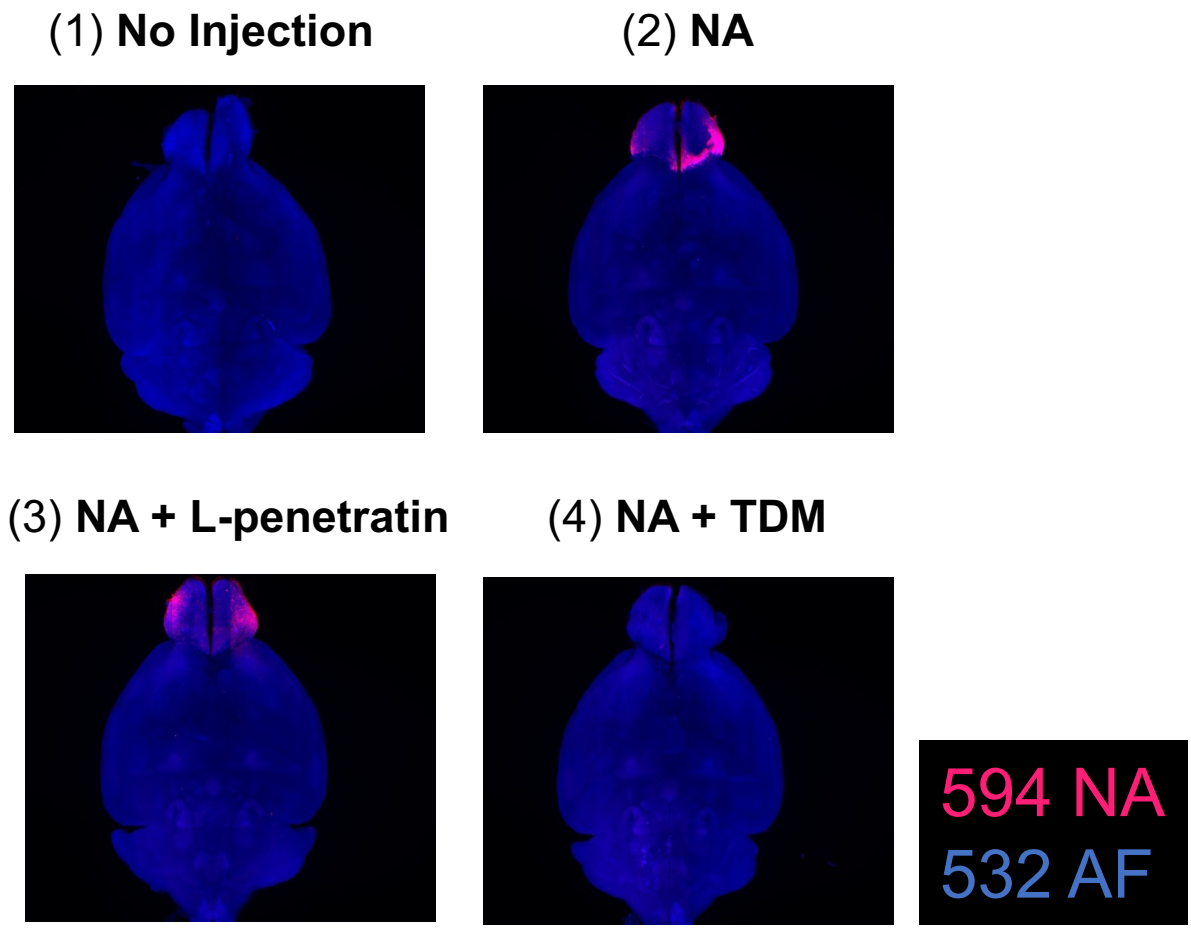


Fig. 4 LSFM image of four mouse brains. Key: no-injection (1) and intranasal administration of NA-Alexa594 solution (2) and with *L*-penetratin (3), TDM (4).

• Intracerebroventricular Administration 0 h, 1 day

Using the established protocol, the localization of NA-Alexa594 in the brain using a delivery method that allowed for delivery to the brain was examined. First, the administration into the left lateral ventricle was performed. (Fig. 5) All ventricles were considered to be connected, and thus, the administrated compound would diffuse in the ventricles even if administered to one side of the brain. 3D imaging was acquired according to the established protocol for both the immediate fixed samples and the perfusion-fixed samples. In both 0 h and 1 day data, the injection point of the needle was clearly visible and the administration into the ventricle was confirmed. In the intraventricular administration-0 h, the left injections diffused inside the ventricles and the fluorescence was observed in and around the walls of the bilateral lateral ventricles, as well as in the cerebellum and blood vessels. When the injection was initiated in front of the ventricles, no fluorescence of the compound was observed in the ventricles, and it diffused into the brain cortex

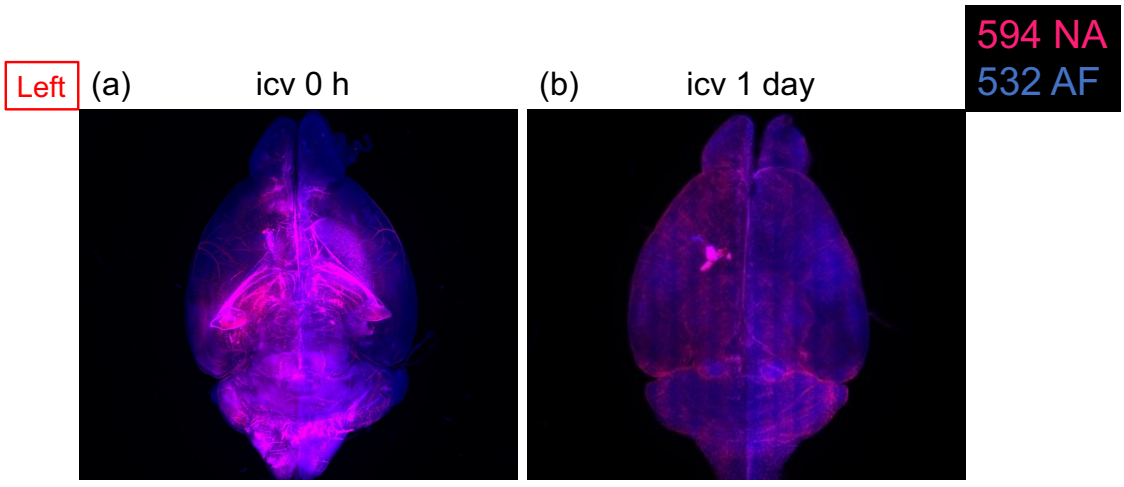


Fig. 5 LSFM image of three mouse brains. Key (a) icv 0 h: Intracerebroventricular administration of NA-Alexa594 solution, perfusion fixation immediately after administration and (b) icv 1 day: 1 day after administration.

around the injection. (data not shown) These experimental facts supported the idea that injection into the ventricles led to distribution around the ventricles and other areas. In the fixed samples after one day of intracerebroventricular injection, the strongest signal was observed at the injection point, and weak signals were present in the entire mouse brain. These results of intracerebroventricular administration suggested that the compound immediately spread to the brain, mainly to the ventricles, after administration, and that the clearance mechanism of the brain worked within a day.

・Intranasal Administration 0 h, 1 day

The drug was administered into the left nasal cavity according to the protocol. 3D images of the brains of the perfusion-fixed and one-day-old animals were acquired immediately after administration as well as the ventricular administration. In the brain samples that were perfusion-

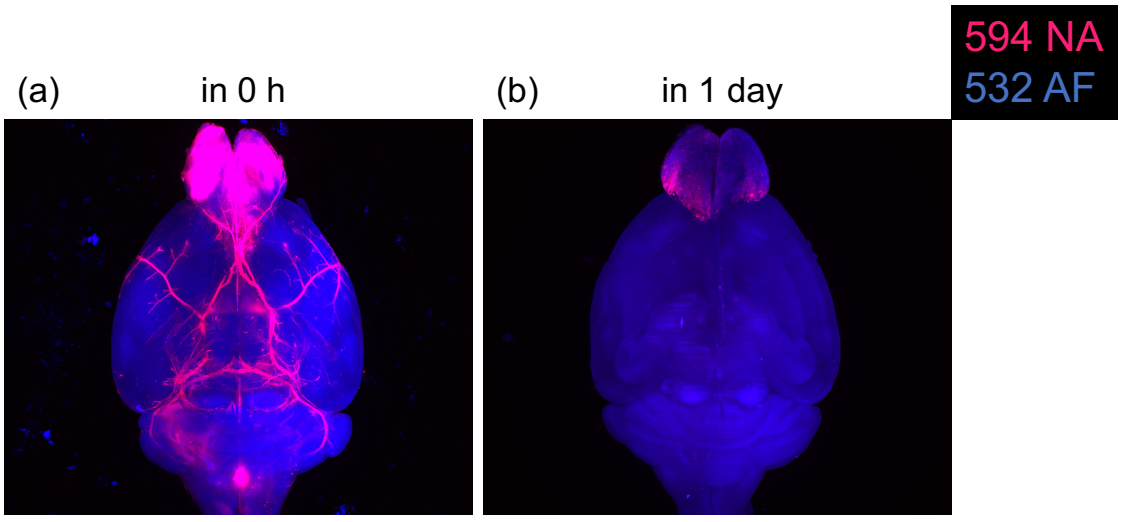


Fig. 6 LSFM image of three mouse brains. Key (a) in 0 h: intranasal administration of NA-Alexa594 solution, perfusion fixation immediately after administration and (b) in 1 day: 1 day after administration.

fixed immediately after the administration, the detection of NA-Alexa594-derived fluorescence and the intranasal delivery of the compound to the brain could be simultaneously confirmed. (Fig. 6) In intranasal administration 0 h and 1 day, the fluorescence was visible not only in the olfactory bulb but also in the blood vessels in the 0 h sample. In the 1 day sample, fluorescence was detected in the left olfactory bulb. It was consistent with the fact that intranasal delivery to the brain is highly blood-transferable in the blood.

The results of compound distribution by the two administration methods showed that it was possible to detect intracerebroventricularly and intranasally administered fluorescent small molecules with the established protocol.

・Intranasal Administration 0, 0.5, 1.0, 3.0, 6.0 h

In the previous section, the localization of NA-Alexa594 was different in the 0 h and 1 day data. It may represent the complex deliveries to the brain by intranasal administration and needed to be investigated in detail.

The changes over time were followed to understand the pharmacokinetics of the administered compounds in the brain by intranasal administration with multiple routes of delivery to the brain. To determine the penetration rate of the compound in the brain and the delivery route to the brain, mice were immobilized at 0, 0.5, 1.0, 3.0, and 6.0 h after administration of NA-Alexa594, and the brains were removed. NA-Alexa594 immediately transferred into cerebral veins immediately after intranasal administration, and NA-Alexa594 adhered to blood vessels, consistent with the previous fact showing that NA cannot penetrate BBB. (Fig. 7) The 0 h sample showed the fluorescence in extensive range. In 0.5 and 1.0 h, the fluorescence had almost completely

disappeared or decreased. 0.5 and 1.0 h experiments were tested five times and the tendency of compound's distribution was consistently decreasing from 0 h samples in both cases. From 3.0 to 6.0 h, the fluorescence was again detected in the olfactory bulb. This result suggested that there was more than two routes of clearance and delivery to the brain in the mouse brain. It was consistent with previous studies showing that there were multiple routes of intranasal administration.[10] The drug delivery routes at 0 and 0.5 h after administration were delivered earlier in the day and tended to be more vascular. It was thought that the drug entered the blood vessels from the olfactory epithelium and was delivered to the brain. In addition, the fluorescence was also observed in the outer reticular layer and mitral cell layer of the olfactory bulb and was distributed in both the left and right olfactory bulbs, indicating that the drug was released from the olfactory epithelium through the intercellular space into the cerebrospinal fluid, and some of it entered the intercellular space of the olfactory bulb and was detected in the brain. The 3.0 and 6.0 h samples had a slower drug transfer rate than that observed at 0 and 0.5 h for delivery to the brain. (Fig. 8) The fluorescence accumulated in the olfactory bulb, especially in the glomerulus from the 0 h sample and was particularly detectable in the left olfactory bulb when the drug was administered through the left nostril. Unlike the delivery observed at 0 and 0.5 h, this result was evidence that the drug passed from the left olfactory epithelium, through the olfactory nerve cells, through the sieve plate, and reached the glomerulus of the left olfactory bulb.

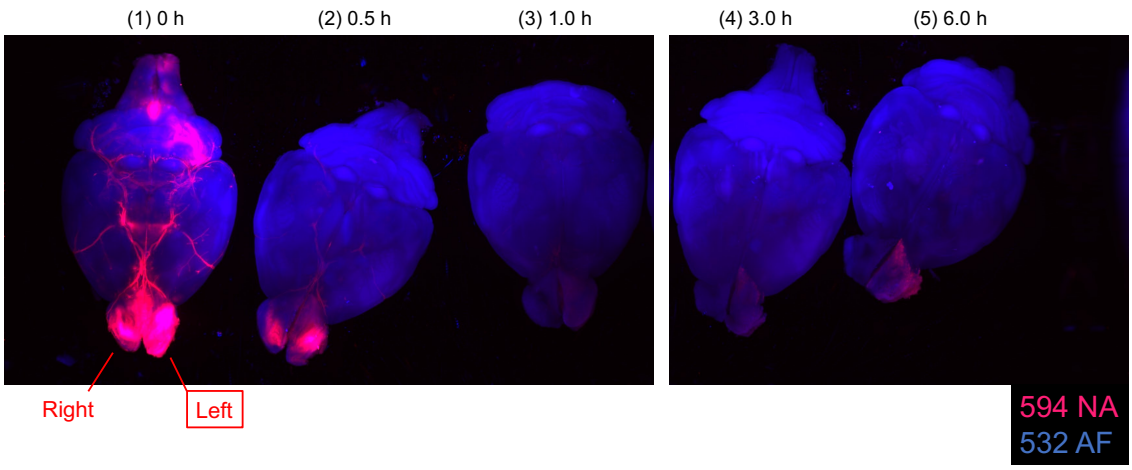


Fig. 7 LSFM image of five mouse brains. Intranasal administration of NA-Alexa594 solution, breeding time after the injection, Key (1) 0 h, (2) 0.5 h (3) 1.0 h (4) 3.0 h. (5) 6.0 h.

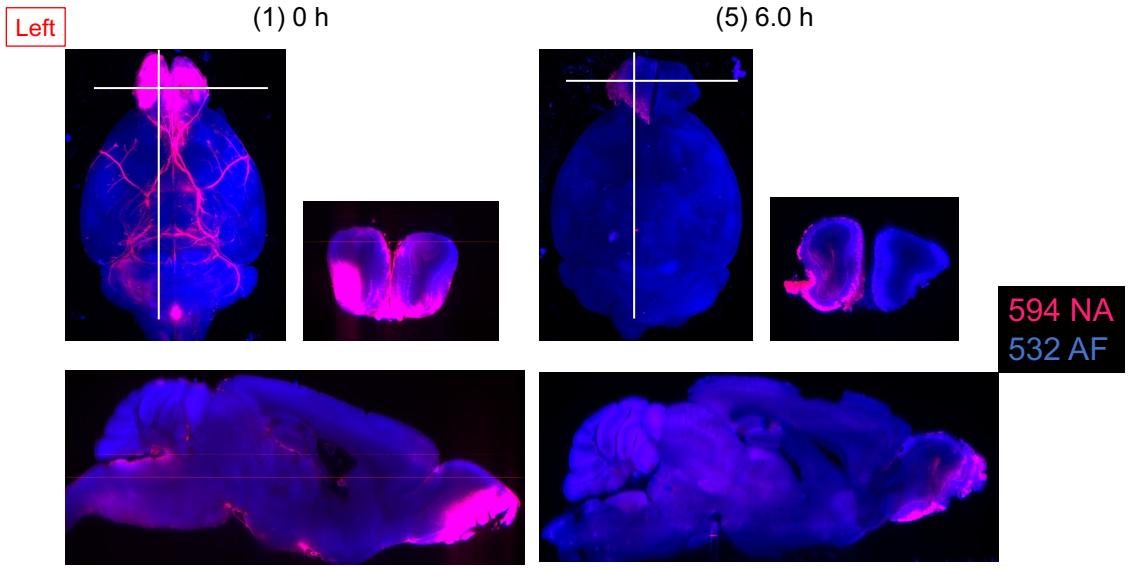


Fig. 8 LSFM image of (1) 0 h and (5) 6.0 h.

・Intranasal Administration 1 day, 1 week

Comparing these temporal changes from 0 to 6 h with the 1 day sample, the 1 day sample localized to the glomerulus as did the 3 and 6 h samples. (Fig. 9) Surprisingly, the 1 week sample was not cleared *in vivo* and localized in the olfactory bulb as well as the 1 day sample. This experimental fact suggested that NA-Alexa594 administered one week before was not cleared and remained in the glomerulus during that time. Alternatively, NA-Alexa594 was being cleared, but was passing through the olfactory nerve at a slower rate, suggesting that it was still being supplied to the glomerulus.

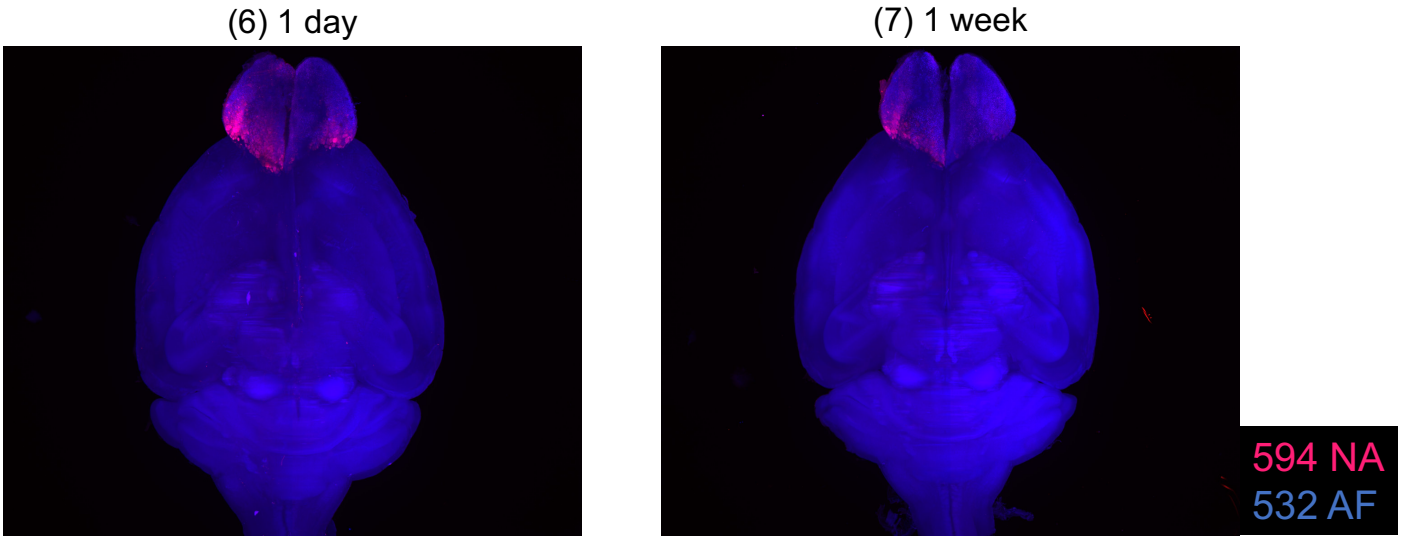


Fig. 9 LSFM image of two brains. Intranasal administration of NA-Alexa594 solution, breeding time after the injection, Key (6) 1 day, (7) 1 week.

• DEA-Alexa594

In order to understand the relationship between the structure of the drug and its delivery or distribution in the brain, DEA-Alexa594, in which NA was converted to diethylamine, was designed and synthesized. (Fig. 10) The compound was administered intranasally and perfusion-fixed immediately after and 6 h after administration. (Fig. 11) On the other hand, it showed almost the same localization in the brain as the result of NA-Alexa594. This result of brain distribution suggested that the NA site did not affect the delivery to the brain by intranasal administration and that the effect of the fluorescent site was significant.

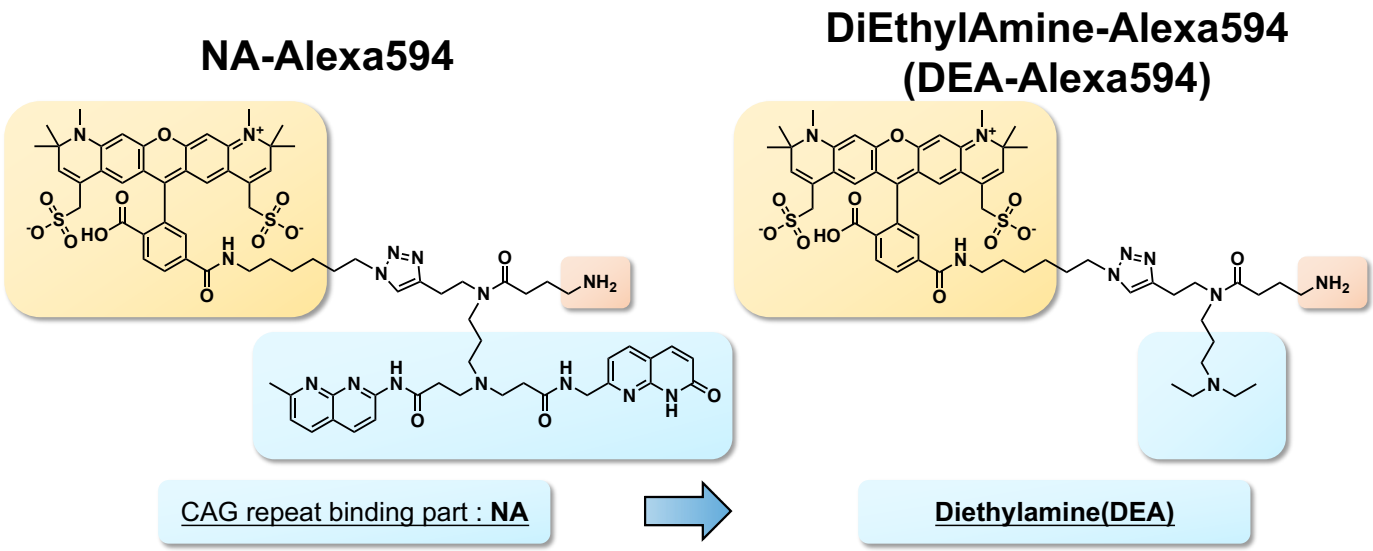


Fig. 10 Chemical structure of DEA-Alexa594

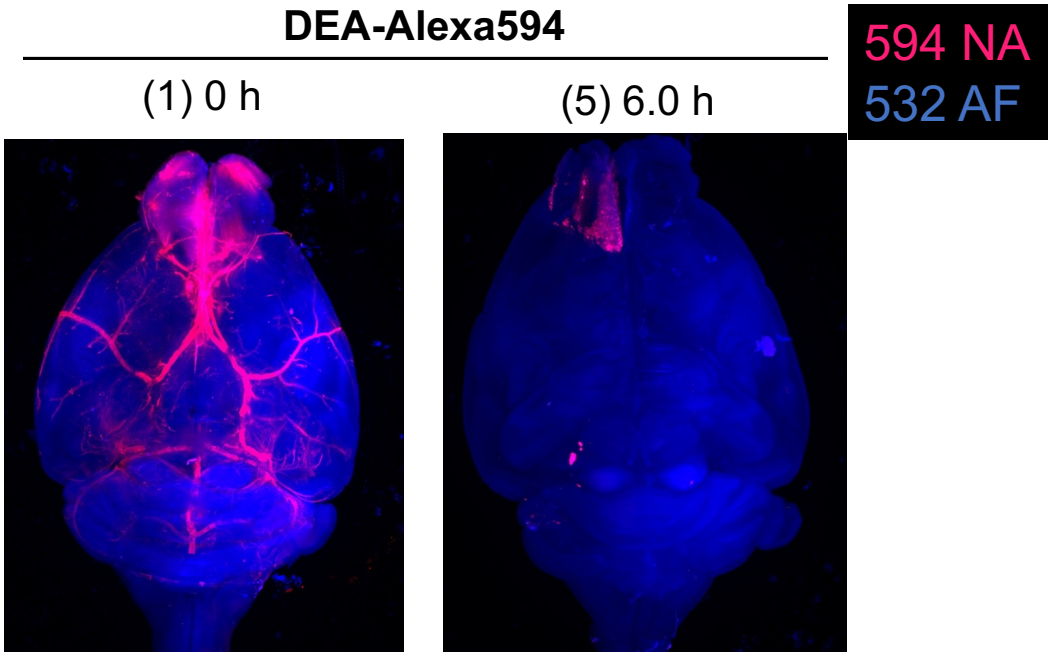


Fig. 11 LSFM image of two brains. Intranasal administration of DEA-Alexa594 solution, breeding time after the injection, Key (1) 0 h and (5) 6.0 h.

Conclusion

Using the NA-Alexa594 detection protocol, the possibility of delivery to the brain by ventricular or intranasal administration was conformed. In the case of ventricular administration, immediately after administration, NA-Alexa594 spread throughout the brain, but after one day of administration, it was found that there was almost no presence of NA-Alexa594 except at the injection site due to clearing. In the case of nasal administration, the signal was present in the blood vessels and olfactory bulb immediately after administration and was still detected near the olfactory bulb one day later, indicating that intranasal administration is capable of delivering the drug to the brain. NA-Alexa594 was also observed the temporal changes from immediately after intranasal administration to 6 h and confirmed the existence of two or more routes of delivery. On the other hand, the delivery to the striatum, which is the affected part of HD, was not observed.

Materials and Methods

Animal

Mouse handling and experimental procedures were performed following the guidelines in Osaka University for the animal welfare and were approved by the institutional review board. C57BL/6NJcl male mice (7-8 weeks) from CLEA Japan, Inc. were housed in groups (maximum of 5 animals/group) with ad libitum feeding, free access to water and food.

Stereotaxic administration

The mice were anesthetized and stereotaxically injected with 3.0 μL of 500 μM NA-Alexa594 dissolved in PBS (left side). Stereotaxic injections were delivered to one site within the cerebral ventricle with the following coordinates: (medial–lateral = 1.1 mm, anterior–posterior = 0.5 mm from the bregma, dorsal–ventral = 2.5 mm)) using a 10-μL Hamilton microsyringe within 30 sec.

Intranasal administration

After delivered mice were bred for one week, anesthetized by isoflurane gas, using a temporarily openable inhalation mask (SN-487-70-09, Shinano Seisakusho, J. Vis. Exp. (141), e58485 (2018), doi10.379158485). Mice were injected by intranasal administration with 2 μL micropipette and adjust it to 2 μL. Positioning the mouse on the mask device, put a 2 μL drop near one nostril to breathe in the solution. After waiting for one minute, as well as, put a next 2 μL drop. Repeat these steps until total volume reached 24 μL.

Perfusion fixation

The mouse was injected with PBS buffer 30 mL and 4% PFA from its left ventricle (flow rate 5 mL/min). The brain tissue was isolated from the perfused mouse. Isolated the mouse brain was immersed in 4% PFA at 4 ºC for 24 h.

Clearing

Wash

The fixed brains were washed by PBS buffer at three times for 1 h each.

Dehydration

The washed brain samples were immersed in 60% MeOH with shaking at room temperature for 4 h, followed by in 80% MeOH with shaking at room temperature for 4 h. Finally, the samples were immersed in 100% MeOH with shaking at ambient temperature overnight.

Delipidation, decolorization and RI matching

The solution exchange new 100% MeOH for 8 h. The samples were immersed in a 100:200:9 mixture of benzyl alcohol, benzyl benzoate and *N*-butyldiethanolamine (BABB-b) with shaking at ambient temperature overnight. The samples were preserved in new BABB-b solution.

References

1. Pardridge, W. M. The blood-brain barrier: Bottleneck in brain drug development. *Neurotherapeutics* **2**, 3–14 (2005).
2. Nakamori, M. *et al.* A slipped-CAG DNA-binding small molecule induces trinucleotide-repeat contractions in vivo. *Nat. Genet.* **52**, 146–159 (2020).
3. Zeniya, S. *et al.* Angubindin-1 opens the blood–brain barrier in vivo for delivery of antisense oligonucleotide to the central nervous system. *J. Control. Release* **283**, 126–134 (2018).
4. Rousselle, C. *et al.* Improved brain uptake and pharmacological activity of dalargin using a peptide-vector-mediated strategy. *J. Pharmacol. Exp. Ther.* **306**, 371–376 (2003).
5. Watanabe, S. *et al.* In vivo rendezvous of small nucleic acid drugs with charge-matched block catiomers to target cancers. *Nat. Commun.* **10**, (2019).
6. Veronesi, M. C. *et al.* Imaging of intranasal drug delivery to the brain. *Am. J. Nucl. Med. Mol. Imaging* **10**, 1–31 (2020).
7. Samaridou, E. & Alonso, M. J. Nose-to-brain peptide delivery – The potential of nanotechnology. *Bioorganic Med. Chem.* **26**, 2888–2905 (2018).
8. Erdő, F., Bors, L. A., Farkas, D., Bajza, Á. & Gizurarson, S. Evaluation of intranasal delivery route of drug administration for brain targeting. *Brain Res. Bull.* **143**, 155–170 (2018).

9. Bourganis, V., Kammona, O., Alexopoulos, A. & Kiparissides, C. Recent advances in carrier mediated nose-to-brain delivery of pharmaceutics. *Eur. J. Pharm. Biopharm.* **128**, 337–362 (2018).

10. Lochhead, J. J. & Thorne, R. G. Intranasal delivery of biologics to the central nervous system. *Adv. Drug Deliv. Rev.* **64**, 614–628 (2012).

11. Kamei, N. *et al.* Visualization and Quantitative Assessment of the Brain Distribution of Insulin through Nose-to-Brain Delivery Based on the Cell-Penetrating Peptide Noncovalent Strategy. *Mol. Pharm.* **13**, 1004–1011 (2016).

12. Arnold, J. J., Ahsan, F., Meezan, E. & Pillion, D. J. Correlation of tetradecylmaltoside induced increases in nasal peptide drug delivery with morphological changes in nasal epithelial cells. *J. Pharm. Sci.* **93**, 2205–2213 (2004).

Chapter 2-3

Detection of NA and NA-like Ligand in Brain using MS Imaging and PET

Introduction

In chapter 2-1 and 2-2, the delivery of NA-Alexa594, a CAG repeats binding moiety, NA attached to a fluorescent moiety, Alexa594, and primary amine to the brain was observed by tissue clearing and light-sheet fluorescence microscopy (LSFM). However, the results suggested that the presence or absence of NA had little effect on the delivery to the brain by intranasal administration and that the Alexa594 moiety has a significant effect on the delivery characteristics to the brain. It is necessary to know whether intranasal delivery to the brain is possible with unmodified NA or NA-like ligands without fluorescent moieties. Clarification of the possibility of intranasal administration of NA would allow for phenotypic experiments to be conducted using this injection method in mouse models for HD.

・MS Imaging

MS imaging is a technique that uses mass spectrometry to directly detect biomolecules, etc., and display the distribution of the target compound on biological tissues as an image. The obtained signal is plotted on the image as intensity in coordinates. In 1997, MS imaging was first reported for tissue imaging using MALDI, but it was not practical due to its low resolution.[1] Since then, improvements in detection lasers and other techniques have enabled us to obtain MS

imaging with high resolution. The distribution of proteins,[2] lipids,[3] neurotransmitter,[4] and drugs[5] has been clarified using MS imaging.

Since MALDI is a soft ionization technique, it can be used for relatively unstable molecules and high molecular weight samples, tens of thousands of Da. In the case of such biological samples, all the molecules to be measured require a matrix for ionization. However, when a large droplet of matrix solution is placed on the sample, the target molecules dissolved from the sample diffuse into the droplet, causing the target molecules to change their positions in the sample. The matrix was manually sprayed with the solution using an airbrush or TLC sprayer so as not to compromise spatial resolution. However, since that method had the possibility of errors due to individual and environment differences, a method of sublimating the matrix and depositing it on the sample has also been developed.[6]

・Design of Scheme

After administration of NA to the mice, dislocation of the cervical vertebra was carried out without perfusion fixation, and the brains were immediately isolated. The isolated brains were sectioned using frozen section method. The paraffin-embedded method was not used because the sections had to be immersed in paraffin once and then heated when blocking and may affect the sensitivity of MALDI. In MALDI, the targeted organic compounds are crystallized and ionized by laser irradiation. For crystallization, matrix was applied. In MALDI, organic compounds were crystallized and ionized by laser irradiation. In this study, the vapor deposition method was used, which is a method of coating the sample on the glass slide by sublimating matrix in vacuum.

Results & Discussion

• Administration Parenchymal

First, MS imaging was used to confirm the detectability of unmodified NA in the tissue environment. 3 μL of 10 mM NA was administered in the same manner as in chapter 2-2 to the brain parenchyma. Shortly thereafter, dislocation of the cervical vertebra was performed and the unimmobilized brain was isolated. The isolated brain was sectioned using the frozen section method taking into account the brain depth of administration. The sectioned brain samples were measured by MS imaging. (Fig. 1) Although the peak of the molecular ion $[M+H]^+$ was the highest intensity in the standard solution, the K^+ adduct was most abundantly produced due to the influence of salt components when dropped on the brain sections. The K^+ adduct (*m/z* 252.05) of the product ion was measured to be generated as precursor to the K^+ adduct (*m/z* 498.16). The

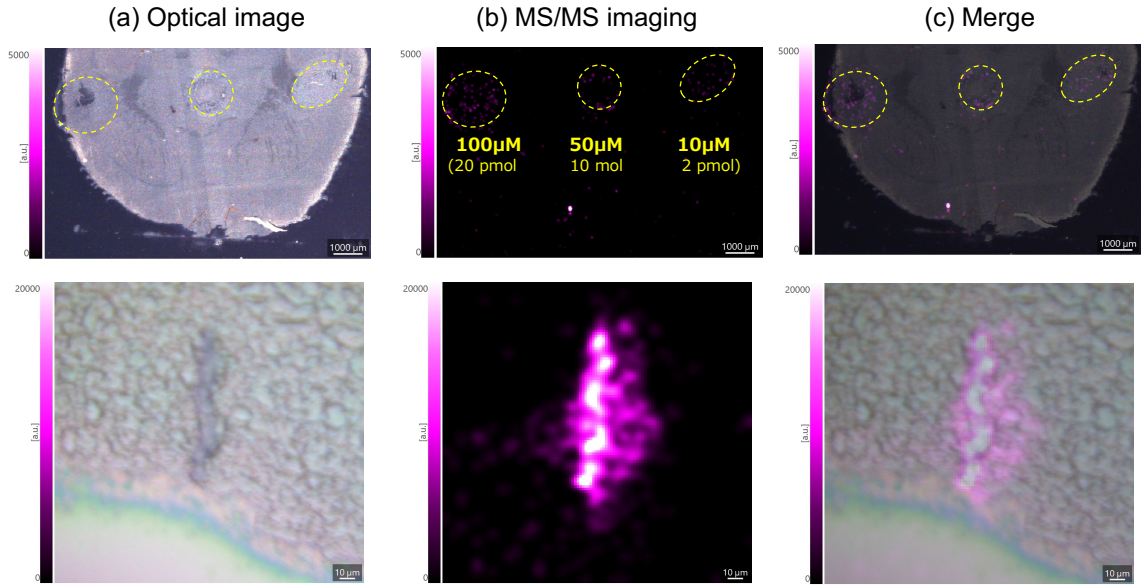


Fig. 1 (a) Optical image, (b) MALDI-MS ion images of m/z 252.05 detected in mouse brain tissue (magenta) and (c) an overlay of optical image and the figures below showed enlarged images.

signal of 252.053 measured by MS/MS was detected near the hole generated by the injection to the brain tissue. The results indicated that NA administered directly into the brain was detectable by MS imaging. The yellow dashed line in the figure showed where the NA solution was dropped on the tissue sections as the control measurement. The purpose of this process was to observe whether it was detectable and the concentration of the detected compound.

・Intranasal Administration

Next, a single dose of 10 mM-NA was administered to confirm the possibility of NA delivery into the brain by intranasal administration. Immediately after administration, the brain was isolated and sectioned as well as the previous section. MS imaging showed no detectable MS or MS/MS peaks derived from NA. (Fig. 2)

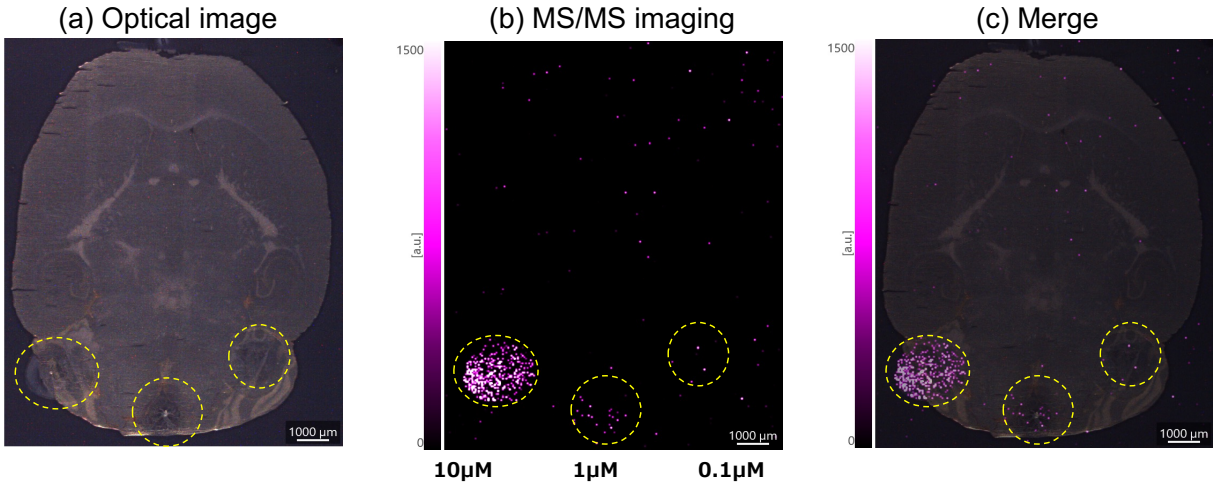


Fig. 2 (a) Optical image, (b) MALDI-MS ion images of m/z 252.05 detected in mouse brain tissue (magenta), and (c) an overlay of optical image

The following is a list of concerns.

- NA is not delivered to the brain by intranasal administration
- NA is delivered to the brain at a slow rate

・Intranasal Administration, Multiple Times

In order to address the concerns described above, it was necessary to consider the delivery rate to the brain and to sufficiently deliver the detected dose. Therefore, intranasal administration of 24 μL×5 doses per week was used to provide sufficient amount and delivery period. As before, MS imaging was performed from brain removal, but no NA-derived peaks were detected. (Fig. 3) These results indicated that NA was not expected to be delivered to the brain by intranasal administration.

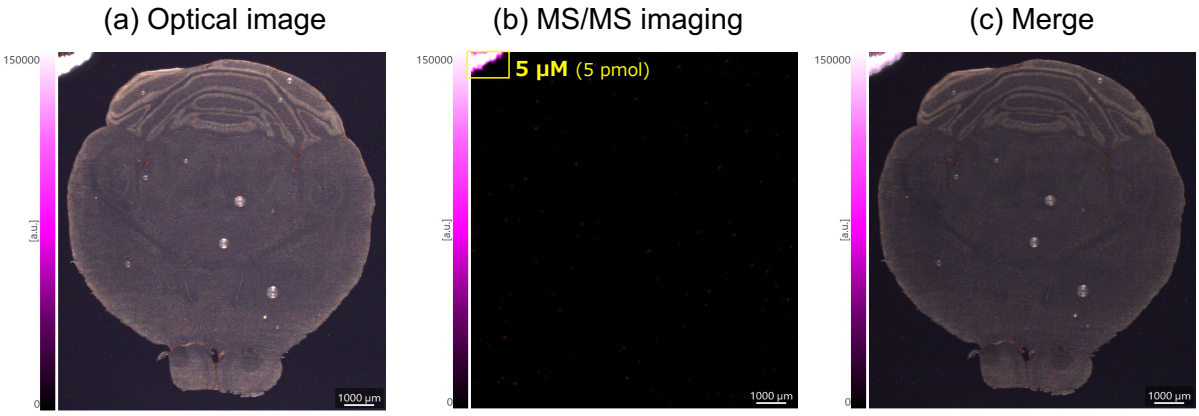


Fig. 3 (a) Optical image, (b) MALDI-MS ion images of m/z 252.05 detected in mouse brain tissue (magenta) and (c) an overlay of optical image

・Discussion in Conjunction with NA-Alexa594 Results

The LSFM results of intranasal administration of NA-Alexa594 and DEA-Alexa594 showed that there was little difference in the distribution in the brain. Although the sensitivities of fluorescence detection by LSFM and MS imaging were different, when combined with the result that unmodified NA could not be delivered into the brain by intranasal administration in MS imaging, it led that Alexa594 may have actively delivered NA into the brain.

Introduction

・PET

In this study, I attempted to detect NA in the brain using positron emission tomography (PET), which has been reported as the first choice for clarifying the distribution of exogenous drugs *in vivo*. PET is a molecular imaging technique for non-invasively visualizing the spatiotemporal distribution of substances in the body.

Results & Discussion

・Design and Synthesis of NA-Me

^{11}C was selected as positron nuclide, which has the smallest molecular weight and is the least affected by NA and there were the reports of the reactions that added ^{11}C to molecules.[7–9] The reaction for ^{11}C-methyl at position 8 of naphthyridine in NA was rejected because methyl at position 8 was formed early in the reaction of NA. The secondary amine in the linker site connecting the naphthyridine to the azaquinolone was selected as the target for ^{11}C methylation. Furthermore, [^{11}C]-MeI could be used for one-step synthesis from NA without changing the recognition moieties in NA. (Fig. 4) The effect of the change from the secondary amine of the linker to the tertiary amine with methyl attached on the physical properties of the compound is unknown. As a possible compound of concern, the nitrogen of the amide of azaquinolone may be methlyated, but in the previous study, compounds with an azaquinolone skeleton were made into

a strong base by NaH treatment and reacted with MeI.[10] Although there is also a possibility that dimethyl may be generated, this is not a problem for the labeling compound.

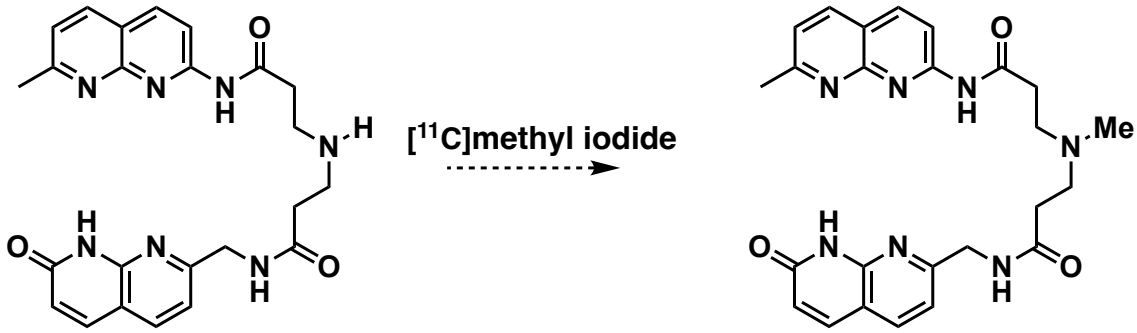


Fig. 4 Designed NA derivative for PET detection

Conclusion

I attempted to detect unmodified NA in the brain. Detection of unmodified NA was performed using MS imaging, which revealed that NA delivered to the brain parenchyma was present and detectable at the injection point. However, delivery of NA by intranasal administration was found not to be existent. It was found that the fluorescent moiety affected the delivery in the intranasal administration.

Furthermore, detection of NA-like ligands by PET was planned. The isotope-labeled NA, [^{11}C]-Me-NA, was designed.

Materials and Methods

Animal

Mouse handling and experimental procedures were performed following the guidelines in Osaka University for the animal welfare and were approved by the institutional review board. C57BL/6NJcl male mice (7-8 weeks) from CLEA Japan, Inc. were housed in groups (maximum of 5 animals/group) with ad libitum feeding, free access to water and food.

Stereotaxic administration

The mice were anesthetized and stereotaxically injected with 3.0 μL of 500 μM NA-Alexa594 dissolved in PBS (left side). Stereotaxic injections were delivered to one site within the brain parenchyma with the following coordinates: (medial–lateral = 1.1, anterior–posterior = 0.5 mm from the intersection of the inferior cerebral vein under the frontal bone with the midline, dorsal–ventral = 2.5 mm) using a 10-μL Hamilton microsyringe within 30 sec. After the injection, the cervical vertebra was dislocated.

Intranasal administration

After delivered mice were bred for one week, anesthetized by isoflurane gas, using a temporarily openable inhalation mask (SN-487-70-09, Shinano Seisakusho, J. Vis. Exp. (141), e58485 (2018), doi10.379158485). Mice were injected by intranasal administration with 2 μL micropipette and adjust it to 2 μL. Positioning the mouse on the mask device, put a 2 μL drop near one nostril to breathe in the solution. After waiting for one minute, as well as, put a next 2 μL drop. Repeat

these steps until total volume reached 24 μL. After the injection, the cervical vertebra was dislocated.

MS Imaging

MALDI MS imaging was performed using an imaging mass microscope, iMScope QT (Shimazu, Japan). The laser was set to the "Ultra" footprint setting at an ~100 μm diameter. Tissue sections were analyzed in positive reflectron ion mode with 80 laser shots fired at 1000 Hz and imaged with a 50 μm laser step size.

References

1. Caprioli, R. M., Farmer, T. B. & Gile, J. Molecular Imaging of Biological Samples: Localization of Peptides and Proteins Using MALDI-TOF MS. *Anal. Chem.* **69**, 4751–4760 (1997).

2. Goodwin, R. J. A., Pennington, S. R. & Pitt, A. R. Protein and peptides in pictures: Imaging with MALDI mass spectrometry. *Proteomics* **8**, 3785–3800 (2008).

3. Griffiths, R. L. *et al.* Formal lithium fixation improves direct analysis of lipids in tissue by mass spectrometry. *Anal. Chem.* **85**, 7146–7153 (2013).

4. Shariatgorji, R. *et al.* Spatial visualization of comprehensive brain neurotransmitter systems and neuroactive substances by selective in situ chemical derivatization mass spectrometry imaging. *Nat. Protoc.* **16**, 3298–3321 (2021).

5. Chen, Y., Tang, W., Gordon, A. & Li, B. Development of an Integrated Tissue Pretreatment Protocol for Enhanced MALDI MS Imaging of Drug Distribution in the Brain. *J. Am. Soc. Mass Spectrom.* **31**, 1066–1073 (2020).

6. Hayasaka, T., Goto-Inoue, N., Masaki, N., Ikegami, K. & Setou, M. Application of 2,5-dihydroxyacetophenone with sublimation provides efficient ionization of lipid species by atmospheric pressure matrix-Assisted laser desorption/ionization imaging mass spectrometry. *Surf. Interface Anal.* **46**, 1219–1222 (2014).

7. Helbert, H. *et al.* Cross-coupling of [11C]methyllithium for 11C-labelled PET tracer synthesis. *Chem. Commun.* **57**, 203–206 (2021).

8. Suzuki, M. *et al.* Pd0-mediated rapid cross-coupling reactions, the rapid C-[11C]methylations, revolutionarily advancing the syntheses of short-lived PET molecular probes. *Chem. Rec.* **14**, 516–541 (2014).

9. Doi, H. Pd-mediated rapid cross-couplings using [11C]methyl iodide: Groundbreaking labeling methods in 11C radiochemistry. *J. Label. Compd. Radiopharm.* **58**, 73–85 (2015).

10. Wawzonek, S. Synthesis of Pyrido[2,3-d]pyrimidine-2,4-diones. **41**, 3149–3151 (1976).

List of Publications

Chapter 1-1

"1,3-di(quinolin-2-yl)guanidine binds to GGCCCC hexanucleotide repeat DNA in C9orf72"

Tomonori Shibata, Eitaro Murakami, Kazuhiko Nakatani

Bioorg. Med. Chem. Let., **2018**, *28*, 2364-2368.

Chapter 1

"Binding mode study of DNA GGCCCC repeat with DQG derivatives"

Eitaro Murakami, Tomonori Shibata, Kazuhiko Nakatani

To be submitted

Chapter 2

"Whole-brain Molecular Distribution for Delivery to Brain by Using Tissue Clearing/3D Imaging"

Eitaro Murakami, Tomonori Shibata, Masayuki Nakamori, Kazuki Tainaka, Kazuhiko Nakatani

To be submitted

List of Presentations

1. Eitaro Murakami, Tomonori Shibata, Kazuhiko Nakatani, “Synthesis and evaluation of ligands binding to C9orf72 GGCCCC repeat DNA related to ALS/FTD”, The Second A3 Young Scientist Meeting, Tohoku University, Jul. 20, 2018.
2. Eitaro Murakami, Tomonori Shibata, Kazuhiko Nakatani, “Synthesis and evaluation of ligands binding to C9orf72 GGCCCC repeat DNA related to ALS/FTD”, The Third A3 Roundtable Meeting on Chemical Probe Research Hub, Jeju, Korea, Oct. 31-Nov. 3, 2018.
3. Eitaro Murakami, Tomonori Shibata, Kazuhiko Nakatani, “1,3-Di(quinolin-2-yl)guanidine binding to C9orf72 GGCCCC repeat DNA in ALS/FTD”, The 45th International Symposium on Nucleic Acids Chemistry 2018/The 2nd Annual Meeting of Japan Society of Nucleic Acids Chemistry, P056, Kyoto, Nov. 7-9, 2018.
4. Eitaro Murakami, Tomonori Shibata, Kazuhiko Nakatani, “N,N’-diheteroaryl guanidine derivatives targeting GGCCCC repeat DNA”, The 48th International Symposium on Nucleic Acids Chemistry 2018/The 5th Annual Meeting of Japan Society of Nucleic Acids Chemistry, 1P-44, Niigata, Nov. 10-12, 2021.
5. 村上英太郎、柴田知範、中谷和彦「ADDA 水素結合面を有する含窒素芳香族グアニジンの合成及び核酸との結合評価」日本化学会第 98 春季年会, 2D5-03, 日本大学理工学部 船橋キャンパス, 2018 年 3 月 20-23 日.
6. 村上英太郎、柴田　知範、中谷和彦「N,N’-ジヘテロアリールグアニジン誘導体の合成と GGCCCC リピート DNA への結合評価」日本化学会第 99 春季年会, 1G4-40, 甲南大学 岡本キャンパス, 2019 年 3 月 16-19 日.

7. 村上英太郎、柴田知範、中谷和彦「GGCCCC リピート結合分子 N,N’-ジヘテロアリールグアニジン誘導体の合成及び結合評価」 日本ケミカルバイオロジー学会第 14 回年, P-122, 2019.
8. 村上英太郎、柴田知範、中森雅之、田井中一貴、中谷和彦「全脳イメージングを用いた CAG リピート結合分子の薬物動態解析」日本ケミカルバイオロジー学会第 15 回年会, P-11, 2021.

Award

1. Travel Award: The Third A3 Roundtable Meeting on Chemical Probe Research Hub 2018
2. ISNAC Outstanding Poster Award in the 48th International Symposium on Nucleic Acids Chemistry 2021

Acknowledgement

I would like to express my sincere gratitude to my advisor, Professor Kazuhiko Nakatani for constructive suggestions and encouragement.

I would also like to express my sincere thanks to Assistant Professor Tomonori Shibata for his helpful discussion, suggestion, and help.

It is a pleasure to thank Associate Professor Chikara Dohno, Associate Professor Asako Murata, Assistant Professor Takeshi Yamada, Specially Appointed Assistant Professor Bimolendu Das for their kind help and discussions in my research.

I would like to thank Professor Gota Kawai for his collaborative efforts in my project of the DNA-ligand complex. I also would like to thank Professor Kazuki Tainaka for his collaborative efforts in my project of the tissue clearing and LSFM and Associate Professor Masayuki Nakamori for his collaborative efforts in my project of the administration to mice. I am also particularly grateful to Professor Shun'ichi Kuroda and Assistant Professor Masaharu Somiya for mice breeding.

I would also like to convey my gratitude to Mitsuhide Murakami and Izumi Murakami for their kindness and encouragement for my research life.