Guanine anchoring : a strategy for specific targeting of a G-quadruplex using short PNA, LNA and DNA molecules

Tan, Derrick Jing Yang; Das, Poulomi; Winnerdy, Fernaldo Richtia; Lim, Kah Wai; Phan, Anh Tuấn

2020

Tan, D. J. Y., Das, P., Winnerdy, F. R., Lim, K. W., & Phan, A. T. (2020). Guanine anchoring : a strategy for specific targeting of a G-quadruplex using short PNA, LNA and DNA molecules. Chemical Communications, 56(44), 5897-5900. doi:10.1039/d0cc01778g

<https://hdl.handle.net/10356/142097>

<https://doi.org/10.1039/d0cc01778g>

© 2020 The Royal Society of Chemistry. All rights reserved. This paper was published in Chemical Communications and is made available with permission of The Royal Society of Chemistry.

Downloaded on 26 Apr 2023 19:00:10 SGT

COMMUNICATION

Guanine anchoring: A strategy for specific targeting of a G-quadruplex using short PNA, LNA and DNA molecules

Received 00th January 20xx,
Accepted 00th January 20xx

Derrick Jing Yang Tan,^a Poulomi Das,^a Fernaldo Richtia Winnerdy,^a Kah Wai Lim^a and Anh Tuấn Phan^{*ab}

DOI: 10.1039/x0xx00000x

Two separate structural elements of a G-quadruplex (G4), a vacant site and a flanking single-strand, could provide an opportunity for specific targeting of particular G4 structures via dual recognition. Here, we show that a short peptide nucleic acid (PNA) can specifically recognize and bind to a G4 at sub-micromolar affinity based on both G-tetrad vacant site filling and complementary duplex formation. This sequence-guided guanine-anchoring strategy can be further developed to improve specificity of targeting G4 structures.

G-quadruplex (G4) is a nucleic acid secondary structure formed by the stacking of multiple planar G-tetrads, each consisting of four Hoogsteen hydrogen-bonded guanines.^{1–5} Genomic regions with key functions, such as telomeres, gene promoters, replication start sites and 5′-untranslated regions (5′-UTRs), were found to be enriched with G4-forming sequences.^{6,7} G4 formations have been implicated in cellular processes⁸ such as telomere maintenance⁹ and gene regulation.¹⁰ Strategies to alter the stability of G4 structures are of great interest for therapeutic applications.^{10–13} Small-molecule ligands, which selectively bind to G4s and induce G4 stability, have been reported.^{14–29} Targeting approaches utilizing a combination of G4 and duplex binders have also been proposed,³⁰ wherein their conjugation was shown to result in a synergistically strong binding ligand with some sequence specificity.³¹

Previously, a guanine-deficit G4 having $4n-1$ core guanines (n denotes the number of tetrads) was reported to form together with G-tetrads a triad layer,^{32, 33,34} which readily accepts guanine-containing molecules to form a complete G-tetrad.^{34,35} Around 230,000 potential $4n-1$ G4-forming sequences have been identified in the human genome.³⁴ The

binding of guanine-containing molecules to a vacancy-bearing G4 increases the stability of the structure,^{33,34} suggesting that the anchoring of a guanine to the G4 vacant site represents an attractive targeting approach.

Peptide nucleic acids (PNAs)³⁶ which display high biological stability,³⁷ has been widely used in nucleic acid therapeutics for regulation of gene expression.³⁸ PNA was shown to bind strongly to DNA or RNA via Watson-Crick base pairing, where the stronger binding affinity is attributed to the lack of electrostatic repulsion between the neutral backbone of PNA and the negatively-charged phosphate backbone of DNA or RNA.³⁹ G-rich PNA has been shown to self-assemble into G4 PNA,⁴⁰ and interplay of PNA with duplex and quadruplex structures have been reported previously.⁴¹ For example, C-rich complementary PNA was shown to invade and disrupt G4 structure by inducing heteroduplex formation,⁴² whereas G-rich homologous PNA was shown to invade G4 to form a heteroquadruplex.^{42–44} Chimeric recognition based on both homologous and complementary binding modes has also been proposed.^{45–47}

Here, we describe the use of PNA in a guanine anchoring approach to target guanine-deficit G4 structures. We showed that with the guanine residue from several short PNAs ($p(AG)$, $p(GA)$ and $p(Lys-GA)$) readily occupying the vacant site in the G4 target, the PNA molecules bind to the target at sub-micromolar affinity and stabilize the G4 structures. We further demonstrated a sequence-guided guanine-anchoring strategy where PNA could recognize a guanine-deficit G4 target through guanine anchoring and complementary duplex formation with the overhang sequence extended from the G4.

A $4n-1$ DNA sequence *delta1* (Figure 1 and Table S1), which was previously examined to bind guanine-containing small molecules ($d(AG)$ and $cGAMP$),³⁴ was utilized as the substrate. Two PNA 2-mers, $p(AG)$ and $p(GA)$ (p stands for a PNA strand in the N-to-C directionality), were used as the ligands (Figure 1A,B), in order to test the effect of strand orientation on the binding.⁴⁸ Another PNA molecule incorporated with an additional lysine residue, $p(Lys-GA)$ (Figure 1C), was also

^a School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371. E-mail: phantuan@ntu.edu.sg

^b NTU Institute of Structural Biology, Nanyang Technological University, Singapore 636921

† Electronic supplementary information (ESI) available. See DOI: 10.1039/x0xx000000

included in the study to test the effect of positive charge for enhancing PNA binding affinity to DNA.⁴⁹ The short PNA ligands were separately titrated to 100 μM of the DNA target (*delta1*) up to 1:1 DNA:ligand ratio to reach the complete binding as monitored by NMR spectroscopy. In each case, the guanine base of the short PNA occupied the vacant site via the formation of four additional hydrogen bonds (Figure 1D), as indicated by a transition of the imino proton spectrum of *delta1* (from 10 peaks for the free *delta1* to 12 peaks for the *delta1*-PNA complexes in the 10-12 ppm region) (Figure 1E). In all three cases, an imino proton peak at 12.5-13.0 ppm was observed, indicating the formation of an A•T base pair between the PNA adenine residue and the thymine residue at the 3'-end of *delta1* (Figure S1). The linewidth of this peak varied among the three complexes, reflecting different stability or dynamics of the base-pair (Figure S1). The data also show that PNA can bind to the site in different strand directionality.

Figure 1 Chemical structure of (A) *p*(AG) (compound 1), (B) *p*(GA) (compound 2) and (C) *p*(Lys-GA) (compound 3). (D) Schematic of *delta1*-*p*(GA) complex. (E) NMR imino proton region of free *delta1*, *delta1* complex with (i) compound 1, (ii) compound 2 and compound 3 recorded at 25 °C. PNAs were added to 100 μM *delta1* to 1 eq. ratio. Spectra of all complexes shows the transition of 10 initial peaks (free form) to 12 final peaks (complex form). Peaks of free form are indicated by black circles and peaks from complex form are indicated by red asterisks. *delta1*: 5'TGGTGGGTGGGTGGGT^{3'}

Table 1. Dissociation constant (K_D) of *delta1*-ligand complexes as measured by ITC

Ligand	K_D (μM)
<i>p</i> (AG)	0.106 ± 0.002
<i>p</i> (GA)	0.196 ± 0.005
<i>d</i> (AG) [#]	$7.4 \pm 0.3^{\#}$
<i>d</i> (GA) [#]	$69.5 \pm 2.5^{\#}$

[#]Data obtained from ref. ³⁴

Binding of PNA ligands provides a strong stabilization effect on the G4 structure. NMR melting experiments in buffer containing 8 mM K^+ showed elevations of the melting temperatures (T_m) from 45.6°C for *delta1* to 64.1°C ($\Delta T_m = 18.5^\circ\text{C}$), 65.2°C ($\Delta T_m = 19.6^\circ\text{C}$) and 67.9°C ($\Delta T_m = 22.3^\circ\text{C}$) upon additions of *p*(AG), *p*(GA) and *p*(Lys-GA) to *delta1* at the 1:1 ratio, respectively (Figure S2). These ΔT_m values are 3-7°C larger than the value of 15.2°C obtained upon adding 3-fold excess of the DNA counterpart *d*(AG).³⁴ The *delta1*-*p*(Lys-GA) complex was measured to have the largest ΔT_m , inferring an additional stabilization provided by the positively-charged lysine side chain.

The dissociation constants (K_D) of the *delta1*-PNA complexes were measured by isothermal titration calorimetry (ITC). Sub-micromolar K_D values of 106 nM and 196 nM for *p*(AG) and *p*(GA) (Figure S3, Table 1), indicating 70- and 350-fold stronger affinity than that of their respective DNA counterparts.³⁴ The absence of electrostatic repulsion between PNA and *delta1* may be the reason for their greater stabilization and binding affinity in comparison to data obtained for their DNA counterparts. Note a secondary lower-affinity ($K_D > 100 \mu\text{M}$) non-specific binding of these PNA molecules toward G4 DNA, as supported by experiments with a parallel G4 without any vacant site, T95-27⁵⁰ (Table S1, Figure S4, Figure S5).

To achieve more specific targeting, we exploited the flanking sequence of a 4n-1 G4 and its complementary strand. We demonstrated the strategy of sequence-guided guanine-anchoring by using the flanking sequence of a 4n-1 G4 to guide a guanine to its vacant site. A G4-forming sequence derived from *delta1* containing 10-nt flanking residues at the 3'-end (named *delta1-ex*, Table S1) was used as the substrate. As a ligand, we used an 11-nt DNA construct (DNA-G) consisting of a 10-nt sequence, which complements the 10-nt flank of *delta1-ex*, connected to a guanine residue at the 3'-end. In addition, as locked nucleic acid (LNA) and PNA have seen widespread applications in antisense oligonucleotide technology due to their strong binding affinity with DNA and RNA,⁵¹ we designed LNA mixmer (LNA-G) and PNA (PNA-G) counterparts of DNA-G (Table S1). The three ligands were utilized to probe the binding ability to *delta1-ex*. With the addition of DNA-G, a quadruplex-duplex hybrid (QDH) structure readily forms at the *delta1-ex*:DNA-G ratio of 1:1. A total of 10 peaks were observed in the imino proton region of 12-14 ppm (Figure 2A and Figure S6), indicative of the formation of a Watson-Crick complementary duplex between the flanking sequence of *delta1-ex* and DNA-G (Figure 2B). The observation of 12 imino proton peaks in the 10-12 ppm region (Figure 2A) indicated the vacant site filling by the 3'-guanine of DNA-G, resulting in a complete G4 structure (Figure 2B). Addition of LNA-G to *delta1-ex* resulted in a

complex exhibiting a similar characteristic QDH imino proton signature as a major species, with a minor conformation being observed (Figure 2A). Addition of PNA-G to *delta1-ex* completely altered the imino proton spectrum the free DNA (Figure 2A) indicating the complex formation, as supported by native gel electrophoresis (Figure S7). The observation of imino protons in the 10–12 and 12–14 ppm regions indicate the formation of both G-quadruplex and Watson-Crick duplex structures, respectively. However, the number and relative intensity of the imino proton peaks suggest the formation of at least two conformations of comparable populations. Due to spectral overlap, we could not provide conclusive evidence for the vacant site filling, although one conformation seems to share the QDH imino proton signature. The above results show that DNA could be effectively designed to use the sequence-guided guanine-anchoring strategy, while in cases of LNA and PNA, other conformations were formed (Figure 2A). These might arise from (i) different adaptable duplex-quadruplex junction interfaces⁵² or (ii) the invasion of the ligands to the G-rich part of *delta1-ex* given that there are four guanines in the ligand sequence.⁴⁷ This situation might be specific only for the studied sequence. Further sequence variations and structural studies on QDH junctions containing LNA and PNA will be required to answer these questions.

To explore further the sequence-guided guanine-anchoring strategy, we designed a 4n-2 G4 substrate (*chl1-mod*, Table S1)

based on a previously reported G4 structure (*chl1*)⁵³ (Figure S8). A 12-nt DNA sequence with 10 complementary residues and two additional guanine nucleotides at its 5'-end (*G₂-DNA*, Table S1) was used as the ligand. At the 1:1 *chl1-mod* to *G₂-DNA* concentration ratio, 12 imino proton peaks (10–12 ppm) arising from the G4 structure and 10 imino proton peaks (12–14 ppm) arising from the Watson-Crick duplex were observed in the ¹H NMR spectrum (Figure S9). The data revealed the insertion of the two terminal guanines of *G₂-DNA* into the G4 core and duplex formation from the flanking sequences (Figure S9), thereby giving rise to a complete QDH structure.

In conclusion, we showed that a guanine in 2-mer short PNA molecules could fill the vacant site of a 4n-1 G4, displaying binding affinities in sub-micromolar range and increasing the melting temperature by about 20°C. Sequence-specific targeting of 4n-1 G4 was achieved by using a strand complementary to the flanking sequence to guide the terminal guanine to anchor at the vacant site of the target G4. We extended the sequence-guided guanine-anchoring strategy to target a 4n-2 G4, thereby forming a stable QDH structure. The sequence-guided anchoring of guanine(s) to vacant site(s) to generate complete G4 structure represents an attractive strategy for specific targeting G4 with incomplete tetrad(s).

This work was supported by the Singapore National Research Foundation Investigatorship (NRF-NRFI2017-09).

Figure 2. (A) NMR imino proton region of *delta1-ex:DNA-G*, *delta1-ex:LNA-G* and *delta1-ex:PNA-G* complex at 1:1 concentration ratio. Imino protons of the G-quadruplex and duplex are tentatively assigned and marked by asterisks and diamonds, respectively. The imino proton of the terminal base pair of the duplex was broad at room temperature (marked by an arrow) due to fast proton exchange with solvent. (B) Schematic of *delta1-ex:ligand* quadruplex-duplex hybrid complex. *delta1-ex*: 5'-TTGGTGGGTGGGTGGG-CATGGACAGC-3'; *DNA-G*: 5'-GCTGTCCATGG-3'; *LNA-G*: 5'-GcTgTcCaTgG-3'; *PNA-G*: 5'-GCTGTCCATGG-3'; complementary segments are underlined; LNA nucleotides are indicated by lowercase.

Conflicts of interest

There are no conflicts to declare.

Notes and references

- 1 M. Gellert, M. N. Lipsett and D. R. Davies, *Proc. Natl. Acad. Sci. U.S.A.*, 1962, **48**, 1013–1018.
- 2 D. Sen and W. Gilbert, *Nature*, 1988, **334**, 364–366.
- 3 S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd and S. Neidle, *Nucleic Acids Res.*, 2006, **34**, 5402–5415.
- 4 D. J. Patel, A. T. Phan and V. Kuryavyi, *Nucleic Acids Res.*, 2007, **35**, 7429–7455.
- 5 J. T. Davis, *Angew. Chem. Int. Ed. Engl.*, 2004, **43**, 668–698.

- 6 V. S. Chambers, G. Marsico, J. M. Boutell, M. Di Antonio, G. P. Smith and S. Balasubramanian, *Nat. Biotechnol.*, 2015, **33**, 877-881.
- 7 N. Maizels and L. T. Gray, *PLoS Genet*, 2013, **9**, e1003468.
- 8 D. Rhodes and H. J. Lipps, *Nucleic Acids Res.*, 2015, **43**, 8627-8637.
- 9 A. M. Zahler, J. R. Williamson, T. R. Cech and D. M. Prescott, *Nature*, 1991, **350**, 718-720.
- 10 A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 11593-11598.
- 11 G. W. Collie and G. N. Parkinson, *Chem. Soc. Rev.*, 2011, **40**, 5867-5892.
- 12 J. L. Mergny and C. Helene, *Nat. Med.*, 1998, **4**, 1366-1367.
- 13 S. Balasubramanian, L. H. Hurley and S. Neidle, *Nat. Rev. Drug Disc.*, 2011, **10**, 261-275.
- 14 D. Sun, B. Thompson, B. E. Cathers, M. Salazar, S. M. Kerwin, J. O. Trent, T. C. Jenkins, S. Neidle and L. H. Hurley, *J. Med. Chem.*, 1997, **40**, 2113-2116.
- 15 D. Monchaud and M. P. Teulade-Fichou, *Org. Biomol. Chem.*, 2008, **6**, 627-636.
- 16 I. M. Dixon, F. Lopez, A. M. Tejera, J. P. Esteve, M. A. Blasco, G. Pratviel and B. Meunier, *J. Am. Chem. Soc.*, 2007, **129**, 1502-1503.
- 17 J. Alzeer, B. R. Vummidi, P. J. Roth and N. W. Luedtke, *Angew. Chem. Int. Ed.*, 2009, **48**, 9362-9365.
- 18 J. Dai, M. Carver, L. H. Hurley and D. Yang, *J. Am. Chem. Soc.*, 2011, **133**, 17673-17680.
- 19 G. W. Collie, R. Promontorio, S. M. Hampel, M. Micco, S. Neidle and G. N. Parkinson, *J. Am. Chem. Soc.*, 2012, **134**, 2723-2731.
- 20 J. M. Nicoludis, S. T. Miller, P. D. Jeffrey, S. P. Barrett, P. R. Rablen, T. J. Lawton and L. A. Yatsunyk, *J. Am. Chem. Soc.*, 2012, **134**, 20446-20456.
- 21 C. Bazzicalupi, M. Ferraroni, A. R. Bilia, F. Scheggi and P. Gratteri, *Nucleic Acids Res.*, 2013, **41**, 632-638.
- 22 W. J. Chung, B. Heddi, M. Tera, K. Iida, K. Nagasawa and A. T. Phan, *J. Am. Chem. Soc.*, 2013, **135**, 13495-13501.
- 23 W. J. Chung, B. Heddi, F. Hamon, M. P. Teulade-Fichou and A. T. Phan, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 999-1002.
- 24 A. Kotar, B. Wang, A. Shivalingam, J. Gonzalez-Garcia, R. Vilar and J. Plavec, *Angew. Chem. Int. Ed. Engl.*, 2016, **55**, 12508-12511.
- 25 Y. Pavan Kumar, P. Saha, D. Saha, I. Bessi, H. Schwalbe, S. Chowdhury and J. Dash, *ChemBioChem*, 2016, **17**, 388-393.
- 26 J. Wirmer-Bartoschek, L. E. Bendel, H. R. A. Jonker, J. T. Grün, F. Papi, C. Bazzicalupi, L. Messori, P. Gratteri and H. Schwalbe, *Angew. Chem. Int. Ed.*, 2017, **56**, 7102-7106.
- 27 W. Liu, Y. F. Zhong, L. Y. Liu, C. T. Shen, W. Zeng, F. Wang, D. Yang and Z. W. Mao, *Nat. Commun.*, 2018, **9**, 3496.
- 28 D. R. Calabrese, X. Chen, E. C. Leon, S. M. Gaikwad, Z. Phyto, W. M. Hewitt, S. Alden, T. A. Hilimire, F. He, A. M. Michalowski, J. K. Simmons, L. B. Saunders, S. Zhang, D. Connors, K. J. Walters, B. A. Mock and J. S. Schneekloth, Jr., *Nat. Commun.*, 2018, **9**, 4229.
- 29 S. Asamitsu, S. Obata, Z. Yu, T. Bando and H. Sugiyama, *Molecules*, 2019, **24**.
- 30 T. Q. N. Nguyen, K. W. Lim and A. T. Phan, *Sci. Rep.*, 2017, **7**.
- 31 S. Asamitsu, S. Obata, A. T. Phan, K. Hashiya, T. Bando and H. Sugiyama, *Chem. Eur. J.*, 2018, **24**, 4428-4435.
- 32 B. Heddi, N. Martin-Pintado, Z. Serimbetov, T. M. Kari and A. T. Phan, *Nucleic Acids Res.*, 2016, **44**, 910-916.
- 33 X. M. Li, K. W. Zheng, J. Y. Zhang, H. H. Liu, Y. D. He, B. F. Yuan, Y. H. Hao and Z. Tan, *Proc. Natl. Acad. Sci. U.S.A.*, 2015, **112**, 14581-14586.
- 34 F. R. Winnerdy, P. Das, B. Heddi and A. T. Phan, *J. Am. Chem. Soc.*, 2019, **141**, 18038-18047.
- 35 K. B. Wang, J. Dickerhoff, G. Wu and D. Yang, *J. Am. Chem. Soc.*, 2020, **142**, 5204-5211.
- 36 M. Egholm, O. Buchardt, P. E. Nielsen and R. H. Berg, *J. Am. Chem. Soc.*, 1992, **114**, 1895-1897.
- 37 E. Uhlmann, A. Peyman, G. Breipohl and D. W. Will, *Angew. Chem. Int. Ed. Engl.*, 1998, **37**, 2796-2823.
- 38 J. C. Hanvey, N. J. Peffer, J. E. Bisi, S. A. Thomson, R. Cadilla, J. A. Josey, D. J. Ricca, C. F. Hassman, M. A. Bonham, K. G. Au and et al., *Science*, 1992, **258**, 1481-1485.
- 39 M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden and P. E. Nielsen, *Nature*, 1993, **365**, 566-568.
- 40 Y. Krishnan-Ghosh, E. Stephens and S. Balasubramanian, *J. Am. Chem. Soc.*, 2004, **126**, 5944-5945.
- 41 V. L. Marin and B. A. Armitage, *Biochemistry*, 2006, **45**, 1745-1754.
- 42 V. L. Marin and B. A. Armitage, *J. Am. Chem. Soc.*, 2005, **127**, 8032-8033.
- 43 A. Paul, P. Sengupta, Y. Krishnan and S. Ladame, *Chemistry*, 2008, **14**, 8682-8689.
- 44 B. Datta, C. Schmitt and B. A. Armitage, *J. Am. Chem. Soc.*, 2003, **125**, 4111-4118.
- 45 M. Hagihara, L. Yamauchi, A. Seo, K. Yoneda, M. Senda and K. Nakatani, *J. Am. Chem. Soc.*, 2010, **132**, 11171-11178.
- 46 D. Bhattacharyya, K. Nguyen and S. Basu, *Biochemistry*, 2014, **53**, 5461-5470.
- 47 B. A. Armitage, *Methods Mol Biol*, 2019, **2035**, 333-345.
- 48 P. E. Nielsen, M. Egholm and O. Buchardt, *Bioconjugate Chem.*, 1994, **5**, 3-7.
- 49 G. Wang and X. S. Xu, *Cell Res.*, 2004, **14**, 111-116.
- 50 N. Q. Do and A. T. Phan, *Chem. Eur. J.*, 2012, **18**, 14752-14759.
- 51 D. A. Braasch and D. R. Corey, *Chem. Biol.*, 2001, **8**, 1-7.
- 52 K. W. Lim and A. T. Phan, *Angew. Chem. Int. Ed.*, 2013, **52**, 8566-8569.
- 53 V. Kuryavyy and D. J. Patel, *Structure*, 2010, **18**, 73-82.