J|A|C|S COMMUNICATIONS

Published on Web 05/11/2010

A U-Tetrad Stabilizes Human Telomeric RNA G-Quadruplex Structure

Yan Xu,* Takumi Ishizuka, Takashi Kimura, and Makoto Komiyama*

Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan

Received November 20, 2009; E-mail: xuyan@mkomi.rcast.u-tokyo.ac.jp; komiyama@mkomi.rcast.u-tokyo.ac.jp

Telomeres have long been considered to be transcriptionally silent. A recent finding demonstrated that telomere DNA is transcribed into telomeric repeat-containing RNA in mammalian cells.[1] The telomeric RNA molecules were detected in different human and rodent cell lines, containing mainly UUAGGG repeats of heterogeneous length.[1] These findings raise the crucial question of how telomeric RNA is specifically associated with chromosome ends. The existence of telomeric RNA may reveal a new level of regulation and protection of chromosome ends. Knowledge of the structure and function of telomere RNA will be essential for understanding telomere biology and telomere-related diseases.

To define the structural features of human telomere RNA sequences, we examined the conformation of human telomere RNA sequence r(UAGGGU) ORN-1 by CD spectroscopy. The CD spectrum of ORN-1 showed a positive band at 265 nm and a negative band at 240 nm in the presence of K^+ or Na^+ at 25 °C (Supporting Information, Figure S1) which are the characteristic CD signature of a parallel G-quadruplex structure.[2] We further investigated the structure of telomere RNA ORN-1 by NMR.[3] In the imino proton region of the 600 MHz ^{1}H NMR spectrum of ORN-1 in the presence of Na^+ or K^+, we surprisingly observed four sharp peaks that are assigned to the imino protons of the G-quadruplex structure at 10.5−12.0 ppm (Figure 1). Since the strands of the UAGGGU parallel G-quadruplex are equivalent, three G quartet rings respectively formed by twelve G residues should give three imino protons. We hypothesized that a way the fourth imino proton can appear is if the U residues at one end of the molecule hydrogen bond with each other, (U)N3−H···O4(U). To test this hypothesis, we prepared the RNA sequence r(UUAGGG) ORN-2 that changes the 3′-end uridine to the 5′-end compared to ORN-1 (Figure S2). ORN-2 gives three imino protons, and the fourth imino proton from the hydrogen bonds of the U-tetrad did not appear, suggesting that the single base mutation in the 3′ and 5′ ends influences the U-tetrad formation and only the 3′-end uridine can form the U-tetrad. Another experiment was performed to identify the U-tetrad formation by using zebularine. The zebularine moiety lacking the oxygen at position 4 of uridine is unable to form hydrogen bonds. As expected from ^{1}H NMR experiment, the fourth imino proton from the hydrogen bonds of the U-tetrad disappeared (Figure S3). These results suggested that the U residues form hydrogen bonds to participate in the formation of the ORN-1 G-quadruplex. To further understand the molecular basis of the telomere RNA G-quadruplex structure, we assigned the imino proton signals on the basis of Nuclear Overhauser Effect (NOE) connectivities (Figure 2). The sequential NOE interactions G3H1/G4H1 and G4H1/G5H1 correspond to imino protons of G on adjacent G-quartets for G3-G4-G5 steps on each of the four strands in the presence of Na^+ or K^+ (cross peaks G3/G4 and G4/G5, Figure 2a, Figure S4). Importantly, we detected NOEs between adjacent imino protons of the G5H1 and U6H3 (cross peaks G5/U6, Figure 2a, Figure S4). The chain from U1 to U6 by monitoring the H8/H6−H1' sequential connectivity of ORN-1 could be traced in NOESY spectra (Figure 2b, Figure S5). We detect the NOE connections of the imino protons to the aromatic protons in the G3-G4-G5-U6 segment of the G-quadruplex to further assign the G-quadruplex (Figure S6). The imino proton of U6 that gives NOEs to H5 and H6 resonances provides further evidence for the U−U base pair (Figure 2c). The strong NOE from the imino proton to H5 must be an intermolecular signal of U moieties, because the NOE to H5 is too strong to be an intramolecular signal of U (the fixed intramolecular distance of imino proton to H5 in U (≈ 4.22 Å) is unable to give such a strong NOE). Furthermore, the intramolecular distance from the imino proton to H6 is out of NOE range (the fixed intramolecular distance of imino proton to H6 (>5 Å) is unable to give an NOE cross peak), indicates that the NOE from the imino proton to H6 must be an intermolecular signal of U. Thus these NOEs are indicative of the presence of N3−O4 hydrogen bonds of U−U base pairs and provides unequivocal evidence for the U-tetrad plane in the G-quadruplex structure (Figure 2f, g). We additionally performed ^{1}H−^{15}N HSQC experiments to assign the U imino proton and identify the hydrogen bonds of U−U base pairs in the U-tetrad.[4] The imino proton-^{15}N cross peaks of the 3′-end uridine were observed (Figure 2e),[5] providing further evidence for the U−U base pair. The chemical shifts for C4 provide insight into the identity of the hydrogen bonding of the U−U base pair. We performed a H(C)C experiment to compare the chemical shift of H5−C4 of U6 (U-tetrad) and H5−C4 of U1 (no U-tetrad).[6] The C4 signal of U6 shifted downfield as compared with the C4 of U1 (Figure 2d), consistent with the result that chemical shifts for hydrogen-bonded carbonyls are shifted downfield relative to those of non-hydrogen-bonded carbonyls.[6] The H5/H3−C5 correlations are obtained to assign the U imino proton and identify the hydrogen bonds of U−U base pairs in the U-tetrad (Figure S7).[7] We noted that uridines from the 5′-end do not participate in the U-tetrad core formation, consistent with their imino protons not being observed (Figures 1a and 2a) and with the fact that no H5/6-H3 NOE cross peaks were found (Figure 2c). Presumably, the A at the 5′-end of the UA sequence may cause a steric hindrance effect on the formation of the U-tetrad at the 5′-end; however, the 3′-end U is not interrupted in forming the U-tetrad

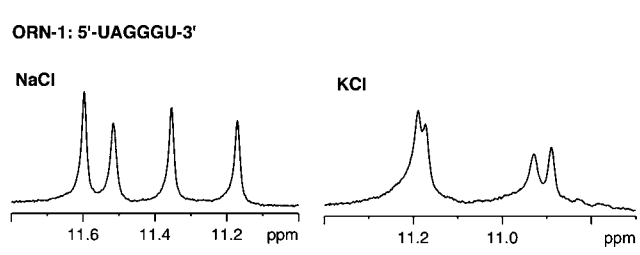

Figure 1. One-dimensional 600 MHz imino proton spectra of telomeric RNA r(UAGGGU) ORN-1 in solution containing 200 mM NaCl (10 mM Na-phosphate, 25 °C) or 200 mM KCl (10 mM K-phosphate, 7 °C).

 © 2010 American Chemical Society

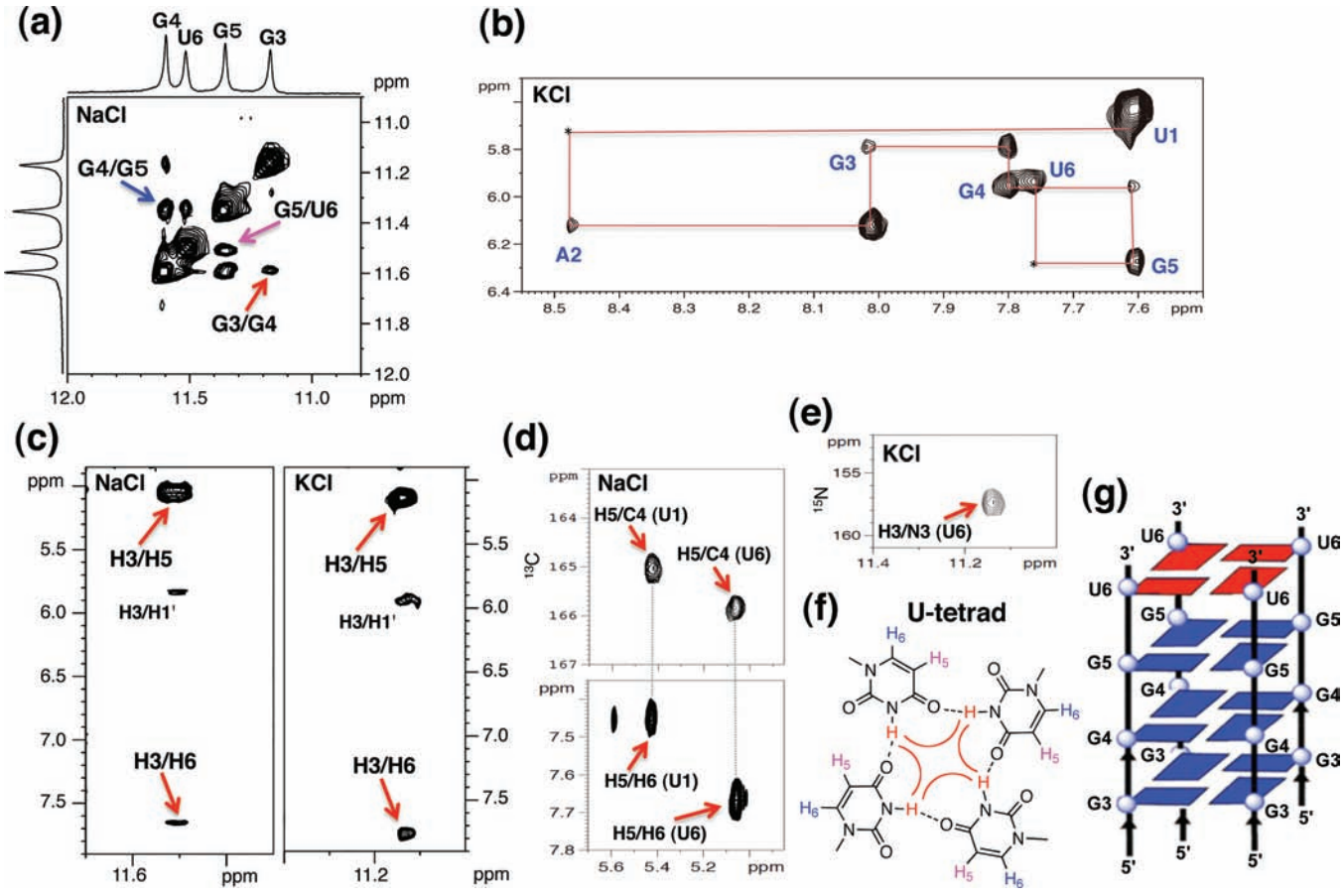

Figure 2. (a) Imino region of 2D-NOESY spectrum of ORN-1 in the presence of 200 mM NaCl and 10 mM Na-phosphate, 25 °C. Cross peaks for G3/G4, G4/G5, and G5/U6 were represented respectively. (b) The H8/6-H1′ proton region of NOESY spectrum (mixing time, 200 ms) of the ORN-1 G-quadruplex in K^+ solution. The sequential assignment pathway is shown (solid line). (c) The H3−H5/H6 (U6) proton region of NOESY spectra of ORN-1 in the presence of 200 mM NaCl and 10 mM Na-phosphate, or 200 mM KCl and 10 mM K-phosphate, 25 °C. Cross peaks for H3−H5, H3−H6, and H3−H1′ were shown. (d) Correlation of C4 to H5 using 2D HCC experiments (upper panel) in Na^+ solution. The H5/H6 (U) TOCSY confirms the 2D HCC H5 assignments (bottom panel). (e) $^1H-^{15}N$ HSQC spectrum showing the H3−N3 (U6) correlation in K^+ solution. (f) A U-tetrad is formed by hydrogen bonds between adjacent uridines at the 3′-end. The hydrogens of H3, H5, and H6 were represented by different colors. (g) Schematic parallel G-quadruplex structure of human telomere RNA (ORN-1) 5′-$U_1A_2G_3G_4G_5U_6$-3′, where red boxes represent uridine bases and blue boxes represent guanine bases.

due to direct binding with G bases. These results suggested that uridines at the 3′-end of the parallel human telomeric RNA G-quadruplex form the U-tetrad with one hydrogen bond between adjacent uridines (Figure 2f, g). We noted the chemical-shift changes induced by different ions, as U6 occurs at an ∼0.6 ppm shift when sodium and potassium ions are changed in solution. The chemical-shift change induced by different ion binding are caused by either a deshielding effect of the ion or structural transition of RNA.[8] For example, the metal binding sites may be different in a sodium or potassium ion-induced U-tetrad due to the different ionic radii for the two metals. This is similar to the observation that the same human telomere sequence adopts a completely different G-quadruplex architecture in a Na^+ or K^+ ion solution.[9]

Next, we investigated the effects of the U-tetrad on the thermodynamic stability of the RNA G-quadruplex by a CD melting curve. Figure 3 shows the melting curves of ORN-1 and ORN-2. Surprisingly, the U-tetrad significantly stabilizes the RNA G-quadruplex structure in the presence of Na^+, leading to an increase in melting temperature (T_m) of 29 °C ($\Delta T_m = 29$ °C). Curve fitting yielded the free energy, enthalpy, and entropy of formation as shown in Table 1. These parameters revealed a very high free energy of formation for the RNA G-quadruplex of ORN-1 with a U-tetrad ($\Delta G = -26.2$ kJ/mol), compared with the RNA G-quadruplex of ORN-2 without the U-tetrad ($\Delta G = -4.2$ kJ/mol) (Table 1). Although the ORN-1 has an unfavorable entropic gain ($\Delta S = -638$ kJ/mol K) compared with the ORN-2 ($\Delta S = -420$ kJ/mol K), the enthalpy value for the ORN-1 G-quadruplex results in being 89 kJ/mol higher ($\Delta H = -224$ kJ/mol of ORN-1 vs $\Delta H = -135$ kJ/mol of ORN-2). To further investigate the effect of the U-tetrad

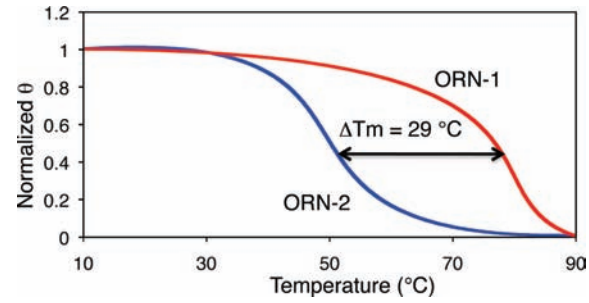

Figure 3. CD melting curves of ORN-1 and ORN-2 monitored at 264 nm in the presence of 150 mM NaCl ($T_m = 79.0$ °C for ORN-1, $T_m = 49.1$ °C for ORN-2).

Table 1. Thermodynamic Parameters for RNA G-Quadruplexes

RNA	$-\Delta H$ (kJ/mol)	$-\Delta S$ (eu)	$-\Delta G_{(37\ °C)}$ (kJ/mol K)	T_m (°C)
ORN-1	224 ± 8	638 ± 21	26.2 ± 1.0	79.0 ± 0.7
ORN-2	135 ± 2	420 ± 5	4.2 ± 0.2	49.1 ± 0.3

on the G-quadruplex stability, we prepared a 4-thiouridine (s^4U)-containing telomeric RNA sequence (Figure S8). Substitution of oxygen with sulfur in position 4 of U impairs molecular hydrogen bonding to each other.[10] We found that deletion of the U-tetrad induced a significant decrease in T_m (Figure S8 and Table S1). These results suggest that the U-tetrad gives an effective enthalpic contribution to the RNA G-quadruplex stabilization. We noted that four uridine residues located at the end of the G-quadruple are favorable for U-tetrad formation than the previously reported RNA dimer G-quadruplex,[11] in which only two uridine residues are positioned at the ends. The CD melting experiments showed that the tetramer structure is significantly more stable (with T_m being 20 °C higher) than the dimer G-quadruplex in K^+ solution (Figure S9).

To further characterize the RNA G-quadruplex structure, MALDI-TOFMS was used to directly observe the G-quadruplex formation.[12] We observed the peaks near m/z [7927.2 + nK^+] or [7658.9 + nNa^+] in correspondence to the molecular weight of ORN-1 (MW = 1915.2, m/z = 4MW + nK^+ or 4MW + nNa^+) (Figure S10), suggesting that the RNA G-quadruplex remains stable even in the gas phase. K^+ and Na^+ ion adducts are clearly observed for the G-quadruplex.[12b] These MALDI-TOFMS data are in excellent agreement with the results obtained from CD and NMR studies, indicating a stable telomere RNA G-quadruplex structure. In several previous studies it has been reported that associated cations can locate at different positions of the G-quadruplex.[9] Different monovalent cations can therefore alter the electronic states of the G-quadruplex, resulting in the differences in the 1D NMR spectrum in Na^+ versus K^+. Furthermore, this remarkable ΔG implies that the four extra hydrogen bonds of the U-tetrad are involved, suggesting that the U-tetrad may be binding an extra K^+ or Na^+ ion in the central channel of the G-quadruplex. A further indication of its stability is the observation that imino protons can be detected in a sample that has been exchanged into D_2O several times and incubated for 24 h in D_2O at 40 °C (Figure S11). These imino peaks are as strong as the nonexchangeable aromatic proton peaks at ~8 ppm. We know of no Watson−Crick duplex as stable as this. In a native PAGE experiment, a single major band was observed for ORN-1 with increasing K^+ and Na^+ concentration (0−200 mM) (Figure S12), indicating a compact G-quadruplex formation by ORN-1 even at low salt concentration.

The finding of telomere RNA molecules opens new doors to better understanding the essential biological role of telomere. There is a clear need to revisit structural and functional mechanisms of telomeres accompanying telomere RNA participation. We and two other groups have focused effort toward the identification of folding topologies for telomere RNA architectures.[11,13] In this study, a key discovery is our demonstration of a novel U-tetrad that forms the base of an RNA G-quadruplex and dramatically stabilizes a human telomeric RNA G-quadruplex structure. The U-tetrad-stabilized telomeric RNA G-quadruplex structure adds considerably to our understanding of the diversity of RNA G-quadruplex architectures. It shows that the structure of base "quartets" is important in RNA assembly.[14] The unique structural feature may provide new targets for efforts directed toward the design and generation of potent and selective RNA G-quadruplex-interacting molecules.

Acknowledgment. This work was partially supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports, Culture, and Technology of Japan. Support by the Global COE Program for Chemistry Innovation is also acknowledged.

Supporting Information Available: General method, CD (Figures S1, S8, S9, Table S1), NMR (Figures S2−S7, S11), MALDI-TOFMS (Figure S10), PAGE (Figure S12). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Azzalin, C. M.; Reichenbach, P.; Khoriauli, L.; Lingner, J. *Science* **2007**, *318*, 798–801. (b) Schoefter, S.; Blasco, M. A. *Nat. Cell Biol.* **2008**, *10*, 228–236.
(2) (a) Petraccone, L.; Trent, J. O.; Chaires, J. B. *J. Am. Chem. Soc.* **2008**, *130*, 16530–16532. (b) Dash, J.; Shirude, P. S.; Hsu, S. T.; Balasubramanian, S. *J. Am. Chem. Soc.* **2008**, *130*, 15950–15956. (c) Rosenzweig, B. A.; Ross, N. T.; Tagore, D. M.; Jayawickramarajah, J.; Saraogi, I.; Hamilton, A. D. *J. Am. Chem. Soc.* **2009**, *131*, 5020–5021. (d) Palumbo, S. L.; Ebbinghaus, S. W.; Hurley, L. H. *J. Am. Chem. Soc.* **2009**, *131*, 10878–10891. (e) Collie, G.; Reszka, A. P.; Haider, S. M.; Gabelica, V.; Parkinson, G. N.; Neidle, S. *Chem. Commun.* **2009**, 7482–7484.
(3) (a) Rankin, S.; Reszka, A. P.; Huppert, J.; Zloh, M.; Parkinson, G. N.; Todd, A. K.; Ladame, S.; Balasubramanian, S.; Neidle, S. *J. Am. Chem. Soc.* **2005**, *127*, 10584–10589. (b) Hu, L.; Lim, K. W.; Bouaziz, S.; Phan, A. T. *J. Am. Chem. Soc.* **2009**, *131*, 16824–16831. (c) Hsu, S. T.; Varnai, P.; Anthony., B.; Reszka, A. P.; Neidle, S.; Balasubramanian, S. *J. Am. Chem. Soc.* **2009**, *131*, 13399–13409. (d) Cosconati, S.; Marinelli, L.; Trotta, R.; Virno, A.; Mayol, L.; Novellino, E.; Olson, A. J.; Randazzo, A. *J. Am. Chem. Soc.* **2009**, *131*, 16336–16337. (e) Hnsel, R.; Foldynov-Trantrkov, S.; Lhr, F.; Buck, J.; Bongartz, E.; Bamberg, E.; Schwalbe, H.; Dtsch, V.; Trantrek., L. *J. Am. Chem. Soc.* **2009**, *131*, 15761–15768. (f) Ambrus, A.; Chen, D.; Dai, J. X.; Bialis, T.; Jones, R. A.; Yang, D. Z. *Nucleic Acids Res.* **2006**, *34*, 2723–2735. (g) Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. J. *J. Am. Chem. Soc.* **2006**, *128*, 9963–9970. (h) Ida, R.; Wu, G. *J. Am. Chem. Soc.* **2008**, *130*, 3590–3602.
(4) Zhan, Y.; Rule, G. S. *J. Am. Chem. Soc.* **2005**, *127*, 15714–15715.
(5) The synthesis of the ^{15}N-labeled RNA was shown in the Supporting Information.
(6) (a) Fiala, R.; Munzarova, M. L.; Sklenar, V. *J. Biomol. NMR* **2004**, *29*, 477–490. (b) Sashital, D. G.; Venditti, V.; Angers, C. G.; Cornilescu, G.; Butcher, S. E. *RNA* **2007**, *13*, 328–338.
(7) (a) Phan, A. T.; Guéron, M.; Leroy, J. L. *Methods Enzymol.* **2001**, *338*, 341–371. (b) Phan, A. T. *J. Biomol. NMR* **2000**, *16*, 175–178. (c) Majumdar, A.; Patel, D. J. *Acc. Chem. Res.* **2002**, *35*, 1–11. (d) Theimer, C. A.; Finger, L. D.; Trantirek, L.; Feigon, J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 449–454.
(8) Furtig, B.; Richter, C.; Wohnert, J.; Schwalbe, H. *ChemBioChem* **2003**, *4*, 936–962.
(9) (a) Patel, D. J. *Nature* **2002**, *417*, 807–808. (b) Parkinson, G. N.; Lee, M. P.; Neidle, S. *Nature* **2002**, *417*, 876–880. (c) Xu, Y.; Noguchi, Y.; Sugiyama, H. *Bioorg. Med. Chem.* **2006**, *14*, 5584–5591. (d) Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. *J. Am. Chem. Soc.* **2006**, *128*, 9963–9970. (e) Ambrus, A.; Chen, D.; Dai, J. X.; Bialis, T.; Jones, R. A.; Yang, D. Z. *Nucleic Acids Res.* **2006**, *34*, 2723–2735.
(10) Kumar, R.K.; Davis, D. R. *Nucleic Acids Res.* **1997**, *25*, 1272–1280.
(11) (a) Xu, Y.; Kaminaga, K.; Komiyama, M. *J. Am. Chem. Soc.* **2008**, *130*, 11179–11184. (b) Martadinata, H.; Phan, A. T. *J. Am. Chem. Soc.* **2009**, *131*, 2570–2578.
(12) (a) Datta, B.; Bier, M. E.; Roy, S.; Armitage, B. A. *J. Am. Chem. Soc.* **2005**, *127*, 4199–4207. (b) Ma, L.; Iezzi, M.; Kaucher, M. S.; Lam, Y. F.; Davis, J. T. *J. Am. Chem. Soc.* **2006**, *128*, 15269–15277.
(13) Randall, A.; Griffith, J. D. *J. Biol. Chem.* **2009**, *284*, 13980–13986.
(14) (a) Pan, B.; Xiong, Y.; Shi, K.; Deng, J.; Sundaralingam, M. *Structure* **2003**, *11*, 815–823. (b) Pan, B.; Xiong, Y.; Shi, K.; Sundaralingam, M. *Structure* **2003**, *11*, 825–831. (c) Deng, J.; Xiong, Y.; Sundaralingam, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 13665–13670. (d) Cheong, C.; Moore, P. B. *Biochemistry* **1992**, *31*, 8406–8414.

JA909708A