JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

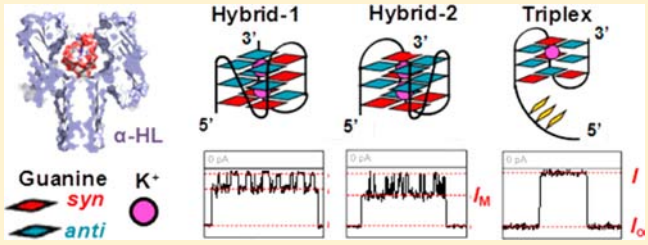
Interactions of the Human Telomere Sequence with the Nanocavity of the α-Hemolysin Ion Channel Reveal Structure-Dependent Electrical Signatures for Hybrid Folds

Na An, Aaron M. Fleming, and Cynthia J. Burrows*

Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112-0850, United States

Supporting Information

ABSTRACT: Human telomeric DNA consists of tandem repeats of the sequence 5′-TTAGGG-3′, including a 3′ terminal single-stranded overhang of 100−200 nucleotides that can fold into quadruplex structures in the presence of suitable metal ions. In the presence of an applied voltage, the α-hemolysin (α-HL) protein ion channel can produce unique current patterns that are found to be characteristic for various interactions between G-quadruplexes and the protein nanocavity. In this study, the human telomere in a complete sequence context, 5′-TAGGG(TTAGGG)$_3$TT-3′, was evaluated with respect to its multiple folding topologies. Notably, the coexistence of two interchangeable conformations of the K$^+$-induced folds, hybrid-1 and hybrid-2, were readily resolved at a single-molecule level along with triplex folding intermediates, whose characterization has been challenging in experiments that measure the bulk solution. These results enabled us to profile the thermal denaturation process of these structures to elucidate the relative distributions of hybrid-1, hybrid-2, and folding intermediates such as triplexes. For example, at 37 °C, pH 7.9, in 50 mM aqueous KCl, the ratio of hybrid-1:hybrid-2:triplex is approximately 11:5:1 in dilute solution. The results obtained lay the foundation for utilizing the α-HL ion channel as a simple tool for monitoring how small molecules and physical context shift the equilibrium between the many G-quadruplex folds of the human telomere sequence.



INTRODUCTION

The human telomere is composed of a telomere-specific protein complex and long stretches of a repetitive sequence, 5′-TTAGGG-3′, that is terminated with a 100−200 nucleotide single-stranded overhang.[1−3] This complex protects the ends of linear chromosomes from deterioration via the DNA damage response pathway.[4] Telomere length regulation is complicated by the end-replication problem that results from incomplete telomerase activity and homologous recombination.[5,6] Despite cells' incessant efforts to maintain the telomere structure, they shorten with age as a result of chronic stress, oxidation, and inflammation.[7,8] Consequently, telomere length serves as a molecular clock for cellular lifespan, and severely truncated ends lead to cell senescence and death.[9,10] Furthermore, ∼85% of human cancer cells maintain excellent telomeres by means of upregulating telomerase, providing a state of immortality.[11,12] This discovery highlights telomere maintenance in cancer cells as a ubiquitous druggable target with high potential for controlling this disease.[13,14]

Essential to telomere function, the 3′ terminal single-stranded overhang can be "hidden" by forming higher order secondary structures that interact with various proteins to retain genomic stability.[15,16] One of the proposed secondary structures is folding of the overhang to a G-quadruplex, which features Hoogsteen hydrogen-bonded assemblies of four guanine bases (G-tetrad) with adjacent G-tetrads assembled together by π−π stacking interactions and coordination of appropriate cations within the quadruplex channel.[17] Recent studies have detected G-quadruplexes in mammalian cells and have confirmed that small molecules do bind these structures in vivo with high affinity.[18] The G-quadruplexes that fold from the human telomere sequence can adopt myriad structures that are dependent on the cation, the medium, and the presence or absence of overhanging ends.[19] Cells maintain K$^+$ cations in the highest concentration (140 mM);[20] therefore, G-quadruplex folds observed in the presence of K$^+$ are thought to dominate in vivo.

The truncated human telomere sequence 5′-TAGGG-(TTAGGG)$_3$-3′ folds unimolecularly in an antiparallel fashion to yield a three-tetrad G-quadruplex with two K$^+$ ions coordinated within the interior channel. NMR analysis has characterized the dominant fold to have a lateral loop closest to the 5′-end, followed by two edgewise loops; this topology has been termed the hybrid-1 fold (Figure 1a). However, in the natural human telomere sequence 5′-TAGGG(TTAGGG)$_3$TT-3′ that includes a 3′ TT overhang, the hybrid-2 conformation is thought to dominate according to NMR studies.[21−24] The key difference for this fold lies in the loop orientation, in which the lateral loop is closest to the 3′-end (Figure 1a). Biophysical studies typically avoid using telomere sequences that contain a 3′-TT tail to simplify the model system, because this tail causes

Received: January 28, 2013
Published: May 17, 2013

   dx.doi.org/10.1021/ja400973m | J. Am. Chem. Soc. 2013, 135, 8562−8570

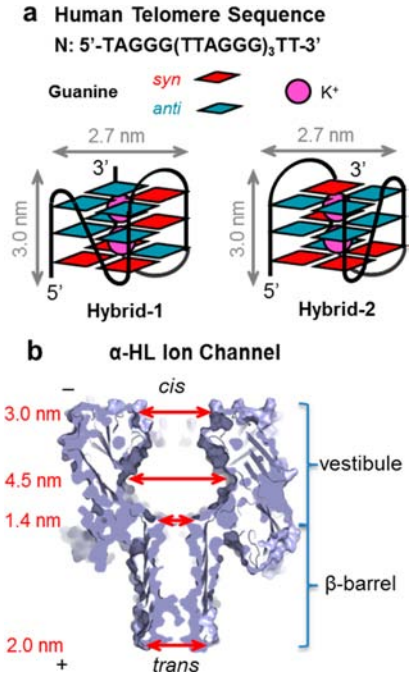


Figure 1. Structures of the G-quadruplexes formed by the human telomere sequence. (a) The sequence, structures, and dimensions of hybrid-1 and hybrid-2 folds.[43] (b) Structure and dimensions of the α-HL ion channel.[44,45]

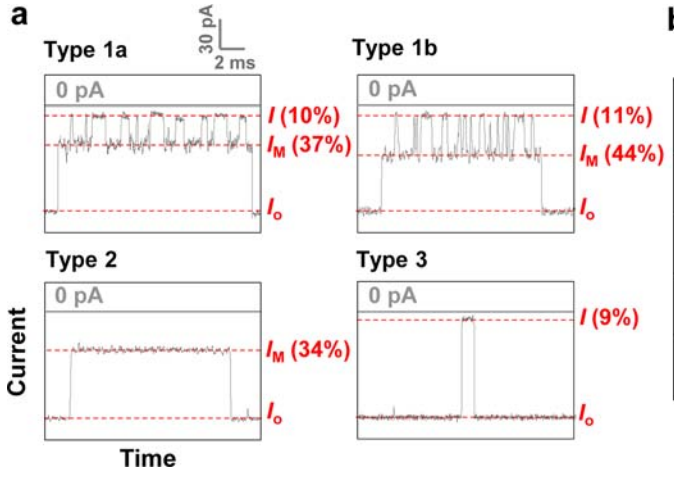
increased structural diversity.[22] Furthermore, these hybrid structures are polymorphic and interconvertible in solution due to the small energetic differences between them.[25] Moreover, a significant amount of a triplex-folded intermediate has recently been observed,[25,26] which has compounded the challenge in studying the human telomere sequence. In a last set of studies concerning the structure of the human telomere sequence by X-ray crystallography and by NMR in solutions with dehydrating agents, the propeller fold was observed.[17,27] This fold is unimolecular with parallel strands and all G nucleotides adopting the *anti* conformation that forces all three loops to be oriented laterally.

The α-hemolysin (α-HL) ion channel-forming protein has advanced as a potential next-generation DNA sequencing platform because of its ability to thread single-stranded, but not double-stranded, DNA through its narrow β-barrel.[28−33] Additionally, it has been used as a tool to provide a wealth of biophysical data concerning DNA and RNA.[34−39] By monitoring the duration and current signatures while electrically drawing the DNA molecules through the nanopore ion channel embedded in a lipid bilayer, the kinetics of the duplex unzipping processes and DNA−protein interactions can be evaluated. Recently, Shim et al. reported studies of the thrombin binding aptamer (TBA), which is a simple G-quadruplex comprising one cation and two tetrads, by encapsulating it in the α-HL nanocavity.[40,41] Furthermore, Rotem et al. utilized an α-HL equipped with the TBA to interrogate the interactions between thrombin and its aptamer as a protein detection platform that also provided thermodynamic and kinetic properties of this complex.[42]

In the present work, we describe the structure-dependent interactions of the human telomeric G-quadruplexes in KCl solutions with the α-HL nanopore ion channel. For these studies, we chose the natural human telomere sequence N = 5′-TAGGG(TTAGGG)$_3$TT-3′ that has a two-nucleotide overhang on both the 5′ and 3′ ends. By monitoring the current−time (i−t) trace patterns, we have identified the hybrid-1, hybrid-2, and triplex conformations that have been difficult to study by ensemble experiments that monitor the properties of the bulk solution. The nanopore method has allowed the interrogation of the relative ratio of these folds at variable temperatures <55 °C on a single-molecule level. This provides key evidence demonstrating that hybrid-1 dominates over hybrid-2 for the human telomere in a complete context (i.e., with both 5′ and 3′ tails) in dilute concentrations of this sequence in KCl solution.

RESULTS AND DISCUSSION

Characterization of G-Quadruplexes. In order to determine if the high ionic strength of the electrolyte (50 mM KCl and 950 mM LiCl) required for these studies influenced the structures, circular dichroism (CD) spectroscopy and thermal melting (T_m) measurements were conducted (Figures S1 and S2). In both 50 mM KCl and mixed salt electrolyte solutions, the CD characteristics were consistent with literature reports of the hybrid folds at lower total salt



b

Event Types	Percentage Distribution
1a	51 ± 5
1b	20 ± 2
2	8 ± 1
3	21 ± 2

Figure 2. The various interactions between the human telomeric sequence and α-HL. (a) Three types of current−time traces observed. (b) The percentage distributions of these event types. These results were obtained in 50 mM KCl, 950 mM LiCl in 25 mM Tris buffer (pH 7.9) at 25 °C.

dx.doi.org/10.1021/ja400973m | J. Am. Chem. Soc. 2013, 135, 8562−8570

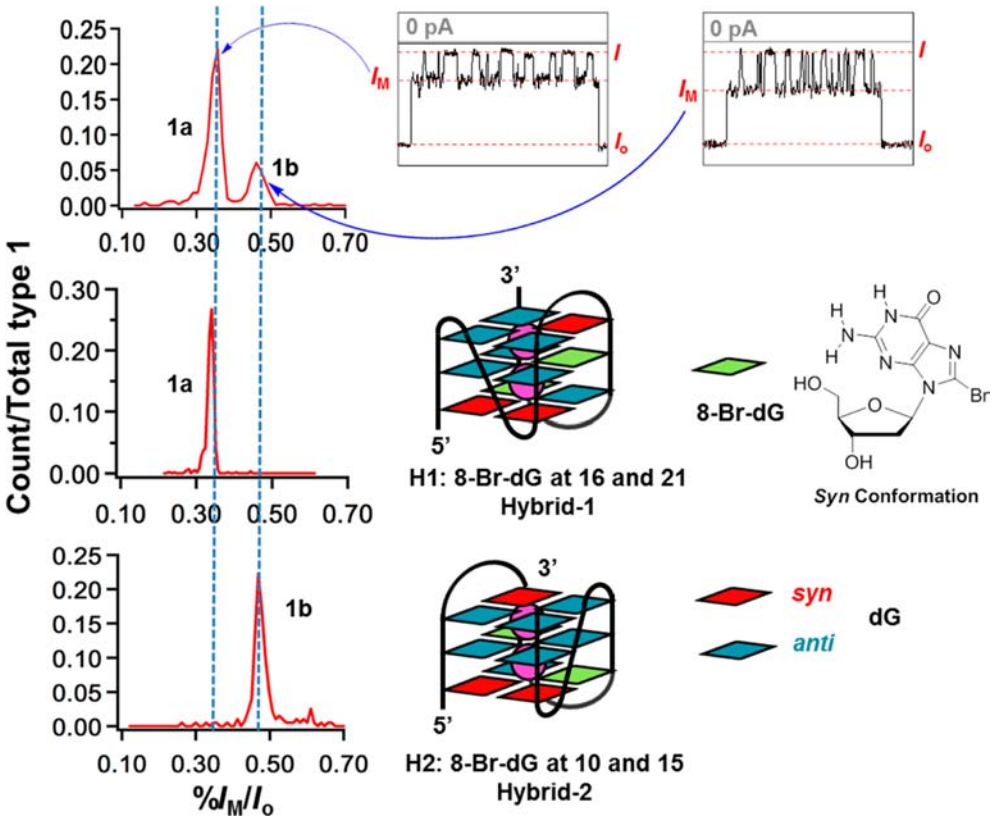


Figure 3. Resolving the coexistent hybrid-1 and hybrid-2 folds at a single-molecule level. Incorporation of two 8-Br-dGs at different positions into the natural sequence can induce the G-quadruplexes to adopt exclusively hybrid-1 or hybrid-2 folds. $\%I_M/I_o$ histograms of the type 1 events from the natural sequence without any substitutions, N = 5′-TAGGG(TTAGGG)$_3$TT-3′ (top), were compared to those of the substituted strands, H1 = 5′-TAGGGTTAGGGTTAG**X**GTTA**X**GGTT-3′ (middle) and H2 = 5′-TAGGGTTAG**X**GTTA**X**GGTTAGGGTT-3′ (bottom), where X = 8-Br-dG. All three histograms represent analysis of ∼500 molecules for each experiment.

concentrations.[46] The only detectable difference for the present samples studied resides in the increased T_m values observed for the higher salt conditions.

Interactions between the Telomeric Sequence and α-HL. We used a previously reported experimental setup to insert a single heptamer of the bacterial ion channel-forming peptide α-HL into a lipid bilayer painted across a ∼600-nm-radius orifice in a glass capillary.[47,48] An electrical potential was applied across the bilayer while DNA was added to the solution on only one side of the bilayer. Being electrophoretically driven to the proximity of the *cis* side of the α-HL ion channel, the folded oligomers were able to enter into the large vestibule of the protein ion channel, producing long current blockages (Figure 2a). There were mainly three types of current−time (i−t) traces observed, which we refer to as type 1, type 2, and type 3 events. As will be further discussed below, these event types appear to correlate with different interactions between the DNA molecule and the protein. Type 1 events featured alternations between intermediate current levels (I_M) and deep blockage current levels (I), which constituted 71 ± 6% of all the recorded events. Interestingly, I_M presented two distinct amplitude subpopulations with different percentage residual currents $\%I_M/I_o$ (Figure 3), where I_o represents the open channel current, of 36 ± 3% and 45 ± 4%, respectively, and having a frequency ratio of I_{M1a}:I_{M1b} = 2.6:1 at −120 mV (*cis* vs *trans*). Type 1 and type 2 events both ended with an intermediate current level, which is an indication that they did not translocate through the narrow constriction zone (∼1.4 nm) of the β-barrel to the *trans* side of the channel. This observation is in stark contrast to that observed for the TBA in KCl solution[40,41] and DNA hairpins[49,50] that were all capable of unraveling and traversing through the ion channel. Type 3 (Figure 2) events contributed 21 ± 2% of the total recordings, and they typically showed simple deep blockage current levels. In order to confirm that the presence of excess Li^+ ion does not affect the distribution of these structures, we conducted the same experiment in 1 M KCl at 25 °C in the absence of LiCl, and similar event distributions were observed (Figure S8). These event types are further discussed below and were shown to reflect folding intermediates that this sequence adopts in KCl solutions.

Hybrid-1 vs Hybrid-2 Folds of K^+-Induced G-Quadruplexes. Interactions between the K^+-induced hybrid folds and α-HL exhibited two subgroups of events within type 1 traces (types 1a and 1b, Figure 2a). Studies have shown that four-repeat human telomere sequences lead to at least two forms of intramolecular G-quadruplexes in K^+ solution. The hybrid-1 and hybrid-2 folds are the dominant conformations based on NMR studies,[22,23] and we suspected them to be the sources of these two different behaviors, type 1a vs type 1b. In order to further investigate the identities of the type 1a and 1b traces, we adopted a similar chemical modification approach that has successfully been used in CD and NMR studies.[22,24] Two 2′-deoxyguanosine to 8-bromo-2′-deoxyguanosine (dG→8-Br-dG) substitutions were made to lock the G-quadruplex into either the hybrid-1 or hybrid-2 fold. The presence of 8-Br-dG enforces the *syn* conformation of that particular nucleotide (Figure 3), leading to >95% hybrid-1 when 8-Br-dG is at

dx.doi.org/10.1021/ja400973m | J. Am. Chem. Soc. 2013, 135, 8562−8570

positions 16 and 21, and to >95% hybrid-2 when it occupies positions 10 and 15 of the sequence. CD and T_m studies on both sets of 8-Br-dG-substituted G-quadruplexes showed consistency with previous reports, except that the T_m values were roughly 10 °C higher in the 1 M salt solution used here (Figure S1).[24] These two modified sequences selectively eliminated one or the other type 1 subpopulation that presented in the I_M histogram of the natural human telomere sequence (Figure 3), indicating that the coexistence of these two hybrid structures can be resolved at a single-molecule level through interactions with the α-HL ion channel. Besides the distinct I_M signatures, the duration of the deep blockage (I) level serves as another parameter to help distinguish these two G-quadruplex folds, with hybrid-1 presenting a nearly 3-fold longer deep blockage level than the hybrid-2 fold (Figure S4). Individual i−t traces did not alternate between I_{M1a} and I_{M1b} within one event, indicating the G-quadruplex cannot change its conformation for the time it is trapped inside of the protein vestibule. Additionally, from these data the relative ratio of these two conformations is >2:1 for hybrid-1 vs hybrid-2, respectively. To confirm that this hybrid ratio is indeed correct, we elected to repeat this experiment with the identical 8-Br-dG substitutions conducted in previous NMR studies (8-Br-dG at only position 15 for hybrid-2),[22] and we observed the same results (Figure S6). For the sake of completeness, we also considered the possibility that the 5′ and 3′ termini of hybrid folds can interact with the α-HL ion channel differently, resulting in these two distinct I_M levels. Since such a directionality effect has been best observed and studied with single-stranded DNA translocating through the protein channel,[51,52] we designed a control sequence that has the same G content as the natural human telomere sequence but is unable to fold into quadruplexes (Figure S7). Translocation results of this sequence under identical experimental conditions showed no resolvable current modulations, suggesting the inability of α-HL ion channel to detect the 5′ and 3′ entry discrepancy of this 24-mer ODN within its most sensitive β-barrel. This argues strongly against the possibility of α-HL producing a 7% current separation for different capture orientations of a given folded quadruplex within its large vestibule, which is typically thought to be less sensitive to molecular identity. Instead, the data support the hypothesis that hybrid-1 and hybrid-2, with their slightly different shapes due to loop topology, yield the characteristic current−time signatures identified here as types 1a and 1b.

These observations provide many fascinating points: (1) The α-HL nanocavity has the sensitivity to discriminate between these two very similar hybrid folds on the basis of their unique I_M values and deep blockage current level durations (Figure 3). (2) While being encapsulated within the vestibule, the hybrid folds cannot interconvert, nor can they unfold sufficiently to translocate. (3) Assuming we could determine the relative capture rate for each hybrid, we could readily describe the ratio of the hybrid folds in dilute solutions for the natural human telomere. This, to date, has been a formidable challenge.[21,22]

G-Quadruplex Tail Interactions with α-HL. In K^+ solutions, the two-nucleotide overhang at either terminus of the human telomere sequence greatly affects the overall topology of folded G-quadruplexes, resulting in increased structural diversity.[22] To simplify the model systems, most studies avoid using sequences containing these overhangs.[25] Nevertheless, we chose to investigate the complete human telomere sequence with both 5′-TA and 3′-TT tails. Since single-stranded DNA can extend beyond the α-HL constriction and interact with the β-barrel, it added another layer of complexity to the analysis. In order to interrogate the role that these tails play in generating unique current patterns, two more sequences were tested in which one or both of the tails were truncated (Figure 4b,c). The sequence that contained only a 5′-

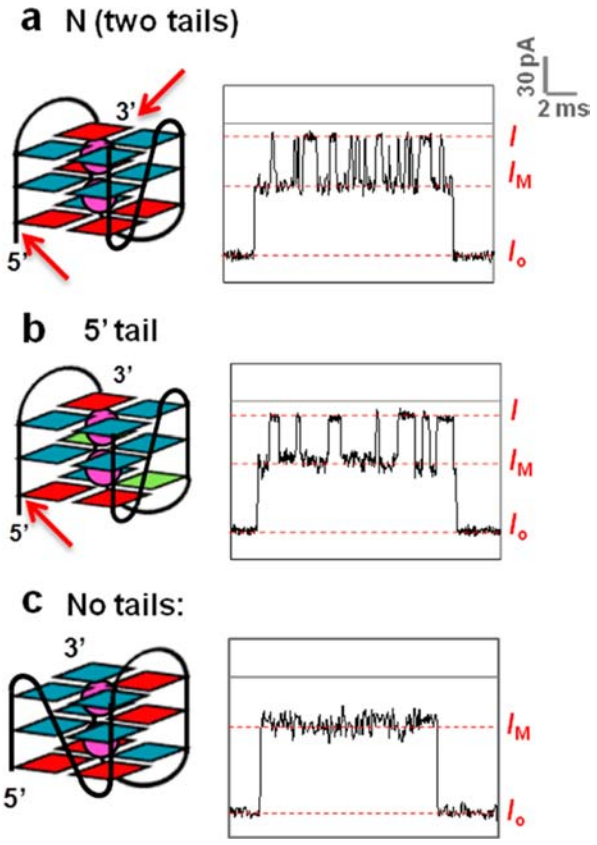


Figure 4. Studies of tail effects on the current signatures. Representative i−t traces for three sequences: (a) natural human telomere sequence N, 5′-TAGGG(TTAGGG)$_3$TT-3′; (b) 5′ tail, 5′-TAGGGTTAGXGTTAXGGTTAGGG-3′, where X = 8-Br-dG; and (c) no tails, 5′-GGG(TTAGGG)$_3$-3′.

TA overhang (Figure 4, 5′ tail) showed a current pattern similar to that of the natural sequence with both tails, while the one without overhangs (Figure 4, no tails) presented a long intermediate current level without causing deep current blockages (I). These results suggest that when the hybrid species are drawn closer to the constriction of the β-barrel, one of the two tails protrudes past the constriction zone at the base of the vestibule and into the β-barrel, producing the deep blockage current level I. Given the shape and stability of hybrid folds, the electrical force cannot unravel them under these circumstances, allowing diffusion to bring the tails back into the vestibule and generating the higher current I_M. The process of the hybrid moving up and down in the vestibule is repeated multiple times before the DNA molecule finally escapes the α-HL ion channel, giving rise to the unique alternating current patterns. It is noted that the 5′ tail sequence used here was a locked hybrid-2 conformation by 8-Br-dG incorporation in order to locate the blunt 3′ terminus at the double-chain reversal loop of this structure. Unfortunately, the I/I_M alternation current pattern was stochastic, making it difficult to determine the exact contributions of the two terminal loops.

dx.doi.org/10.1021/ja400973m | J. Am. Chem. Soc. 2013, 135, 8562−8570

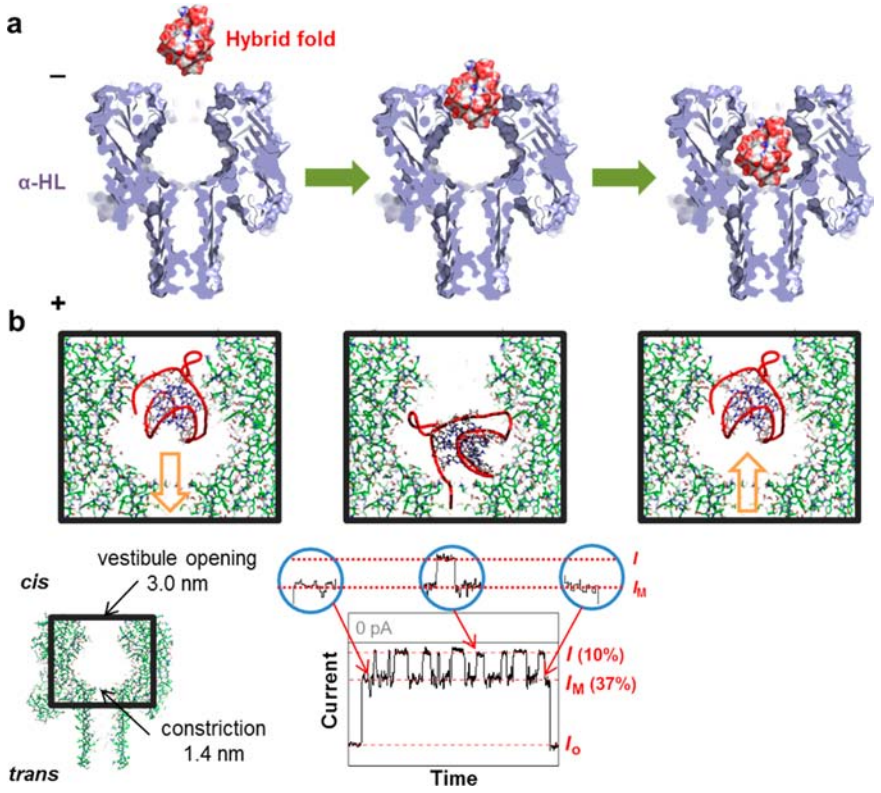


Figure 5. Illustration of the proposed mechanism for the current signatures. (a) Space-filling model of the hybrid fold entering the α-HL vestibule using PBD structures 7AHL[45] and 2JSQ.[22] (b) Stick model of the proposed interaction mechanism. Intermediate current levels (I_M) at the beginning and end of an event correlated to the interaction between the DNA and the opening of the protein vestibule. In the presence of an electrical potential (−120 mV, *cis* vs *trans*), the G-quadruplexes can be driven closer to the β-barrel, producing deeper blockages I with one tail protruding into the constriction, or diffuse back to the opening of the vestibule, leading again to I_M current levels.

Our observation of the tail effect is consistent with a similar study of the TBA interacting with the α-HL ion channel.[40] The structure of this sequence fold featured a two-tetrad, one-cation G-quadruplex without overhangs, and it produced a flat intermediate current level in the α-HL ion channel before the TBA unfolded and translocated. After the attachment of a six-nucleotide tail, a similar I/I_M alternating current pattern was observed as a result of the tail interacting with the β-barrel. These studies address the property of the G-quadruplex that gives rise to the deep blockage current levels, and the overall mechanism of the DNA/protein interaction is highlighted in the section below.

Overall Proposed Mechanism. On the basis of the observation that we can detect differences between these very similar hybrid folds, we propose the following mechanism for why this is possible at the single-molecule level. The current level differences between the hybrid folds in the α-HL vestibule may result from slight differences in their size and shape that combine to alter the molecular interactions occurring at the interface between the quadruplex and the protein vestibule's constriction sites. The vestibule of α-HL is flanked by a larger opening on the *cis* side ($d \approx 3.0$ nm) and a smaller constriction site on the *trans* side ($d \approx 1.4$ nm; Figure 1b).[44,45] Hybrid folded G-quadruplexes are roughly 2.7 × 2.7 × 3.0 nm, which narrowly allows their entry into the vestibule. The detectible differences between hybrid-1 and hybrid-2 are associated with the I_M current level, while both folds gave similar deep blockage current levels. We hypothesize the following: Similar deep current levels occur when the electrophoretic force pulls the quadruplex against the smaller constriction zone on the *trans* side of the vestibule, where the quadruplex can occlude the current flow maximally while blocking much of the ∼40 yoctoliter (yL) volume of the vestibule. The result is a ∼10% residual current due to one of the 2-nucleotide overhangs being drawn into the β-barrel (Figure 4). Because the hybrid folds are so much larger than the ∼1.4 nm constriction, they give similar deep blockage residual currents (Figure 5). However, when diffusion allows the quadruplex to drift up against the larger opening on the *cis* side of the vestibule, a difference in current level (I_M) is detected. Apparently the size match between the hybrid folds (∼2.7 nm) and the *cis* constriction site (∼3.0 nm) allows these distinct I_M levels to occur. Thus, the subtle differences in the hybrid fold structures lead to interactions with the protein surface in a unique fashion to bring about the ∼7% difference (37% vs 44%) observed in the intermediate current level, I_M. An additional I_M current level of ∼34% is observed in ∼8% of the events (type 2, Figure 2a), but these events show no deep current blockage. We propose that these represent entry of the hybrid fold edge-first rather than face-first (type 1, Figure 2a), corresponding to an orientation that cannot strongly block the entry to the β-barrel, probably because the 3′ and 5′ tails are oriented laterally or, in any case, do not approach near enough to enter the β-barrel. However, the exact bending, stretching, or compressing interactions that are contributing to the I_M currents cannot be further deciphered by this experimental technique.

Support for this overall hypothesis comes from the 8-Br-dG studies that validate the current differences obtained in the locked H1 and H2 models matching those of the captured hybrid-1 and hybrid-2 folds from solution, thereby providing a

dx.doi.org/10.1021/ja400973m | J. Am. Chem. Soc. 2013, 135, 8562−8570

readout of the folding topology. Furthermore, because the event types always initiate and terminate with a mid-level block, we interpret the mid-level current readings as resulting from quadruplex interactions at the *cis* side of the vestibule. A further test to demonstrate the inability of hybrid quadruplexes to translocate through the β-barrel was observed in voltage-dependent studies, in which increasing the applied voltage only increased the capture time, and did not affect the deep or intermediate blockage current levels or their fluctuation times (Figure S9). This observation is consistent with a study of interactions between peptides and the α-HL ion channel.[53] Additionally, these observations demonstrate that α-HL can be a powerful analytical tool for monitoring the structural topology distribution for G-quadruplexes in solution at the single molecule level, in which the data are readily comparable to known standards; such discrimination has not been easily achievable by other methods.[21,22,54]

Human telomere G-quadruplexes are polymorphic in vitro, and this fact increases the complexity of finding these structures in vivo. Excitingly, recent studies have found the G-quadruplex fold in mammalian cells utilizing a fluorescent antibody that bound many different folding topologies (e.g., hybrid and propeller folds).[55] Biophysical studies have suggested that the human telomere sequence under dehydrating conditions gives the propeller fold,[27,56] while studies utilizing BSA as a crowding agent suggest the hybrid fold persists.[57] In the studies reported herein, the hybrid folds were analyzed within the confined (i.e., crowded) context of the vestibule in α-HL. Hybrid folds, once captured, showed current levels that alternated between a deep block and a midlevel block, and these current levels did not change during a capture event (~300 ms). This observation is interpreted to mean that a hybrid fold, once in the nanocavity of the protein ion channel, does not change its conformation dramatically, although it may still be bending, stretching, or compressing. That is to say, hybrid-1 and hybrid-2 do not interconvert while captured in the α-HL cavity, nor does a hydrid convert to other known structures such as the propeller fold that is proposed under conditions of dehydration and molecular crowding. Further support for this claim again resides in the 8-Br-dG studies, where we found that forcing a critical G into the *syn* conformation does not allow the propeller fold to exist even in conditions known to favor this fold (e.g., 40% CH_3CN, Supporting Information).[27,56] Upon capture of a locked hybrid with 8-Br-dG, current signatures were identical to those observed for a hybrid fold in the native sequence. Therefore, the nanocavity, though small (~40 yL in volume) and crowded, was not capable of inducing the propeller fold, and the hybrid fold persisted.

Existence of G-Quadruplex Folding Intermediates in K^+ Solution. Even though the structures and energetics of the human telomeric G-quadruplexes have been widely investigated, the actual folding pathway is not clearly established.[58] Reports have suggested the existence of many folding intermediates, in which the triplex intermediate fold has recently been highlighted in optical tweezer and spectroscopic experiments.[25,26,54] Interestingly, we noted a third event type in KCl solution that constituted 21 ± 2% of the overall events and presented a simple deep blockage current level consistent with translocation (type 3, Figure 2a). Because the hybrid folds can be readily assigned in 8-Br-dG experiments, and translocation was never observed for the hybrid structure, we hypothesize that the third event type may correspond to the triplex intermediates. In order to closely examine this, we synthesized a model sequence with the GGG close to the 5′ terminus replaced with TTT to induce the triplex structure yielding a longer 5′-tail (Figure 6a, 5′T). This triplex sequence was

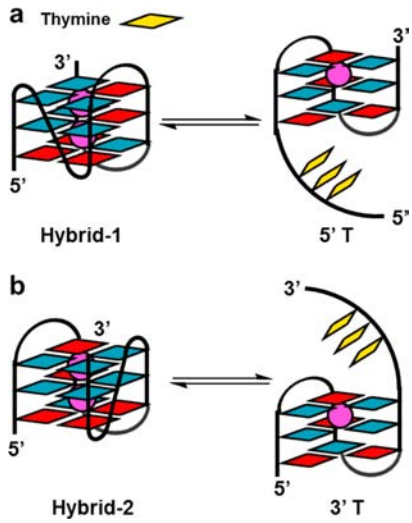


Figure 6. Proposed triplex folding intermediates of the hybrid folds. Replacement of 5′ vs 3′ GGG with TTT leads to 5′ tail (5′T) and 3′ tail (3′T) triplexes, proposed to be the folding intermediates of (a) hybrid-1 and (b) hybrid-2 folds.

selected as a likely intermediate that would form from the hybrid-1 fold dominating in solution. Characterization of this intermediate by CD provided results similar to those previously described by the laboratories of Chaires, Sugiyama, and Parrinello (Figures S2 and S3)[25,26,59] and a most recent report ($T_m \approx 35$ °C) (Figure S2).[59] Translocation studies on the 5′-T triplex sequence within α-HL provided a similar i–t trace signature and duration time compared to the type 3 events produced by the native human telomere sequence (N), supporting the hypothesis that the third event type corresponds to the presence of a triplex intermediate at 25 °C (Figure 7a). For the sake of completeness, we also studied the triplex that would be formed from the less abundant hybrid-2 fold that would present with a 3′-tail; for these experiments, we replaced the 3′-most GGG run with TTT (3′T). This induced triplex was also capable of translocating through the protein channel based on voltage-dependent studies (Figure 7c) in which a decrease in τ with increasing applied voltage was observed for all proposed triplex structures.

A summation of the observed events for the hybrid and triplex folds allows us to describe >95% of the folded states for the natural human telomere sequence in dilute solutions. Hybrid folded G-quadruplexes are notoriously polymorphic, so other uncharacterized structures likely exist in solution and may describe the small amount of unidentifiable event types.[54] Overall, the human telomere sequence folds to many topologies in KCl solutions,[54] among which we have identified the two hybrid folds and their triplex folding intermediates by utilizing the α-HL nanopore as a size-selective sieve.

Thermal Denaturation Profile of Human Telomeric G-Quadruplexes. Armed with the ability to detect hybrid folds and the triplex folding intermediates, and because these structures display a high degree of polymorphism and dynamic behavior in solution, we chose to interrogate the thermal profile for these structures in solution. Additionally, we recognize that the α-HL ion channel can discriminate molecules of various

dx.doi.org/10.1021/ja400973m | J. Am. Chem. Soc. 2013, 135, 8562–8570

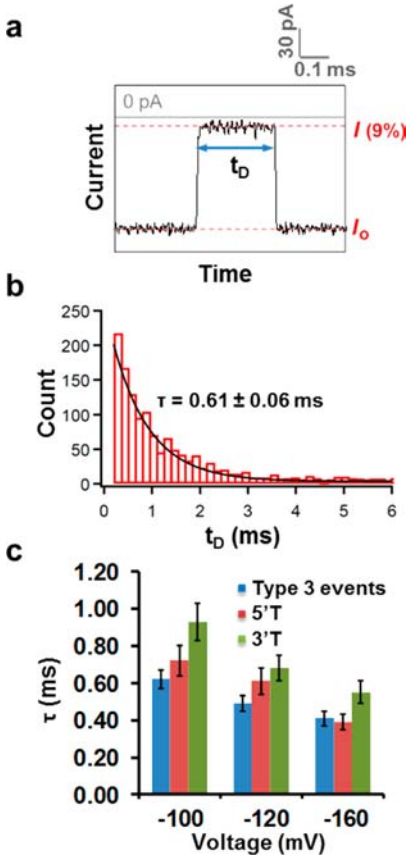


Figure 7. Nanopore measurements of the triplexes 5′T and 3′T. (a) Examples of *i*–*t* traces of the triplexes and the durations of the events (t_D) that were fit to a single exponential decay model, shown in (b) with a time constant τ. (c) Plot of the decay constants of translocation events (blue, type 3 events from the natural sequence; red, 5′-T triplex; green, 3′-T triplex) as a function of applied voltage (*cis* vs *trans*).

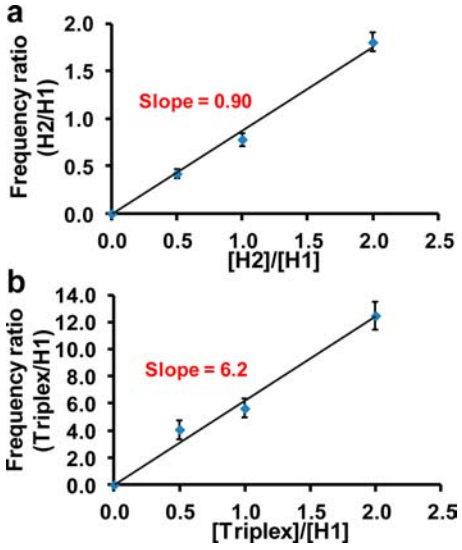


Figure 8. Entry rate corrections between the hybrid species and the triplex. Plots of event frequency ratio vs concentration ratio for (a) hybrid-2 (H2) and hybrid-1 (H1) model sequences and (b) triplex and H1. The ODNs sequences used are as follow: H1, 5′-TAGGGTTAGGGTTAG**X**GTTA**X**GGTT-3′; H2, 5′-TAGGGT-TAG**X**GTTA**X**GGTTAGGGTT-3′ (X = 8-Br-dG); and triplex, 5′-TATTT(TTAGGG)$_3$TT-3′.

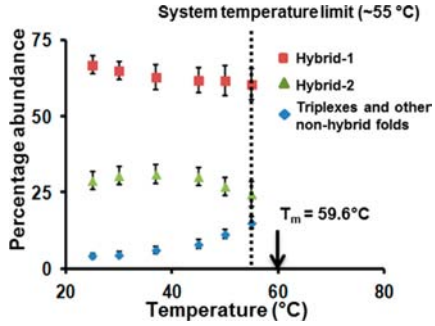


Figure 9. Thermal denaturation profile of the human telomeric sequence. Plot of percentage abundance vs temperature after entry rate corrections for hybrid-1, hybrid-2, and triplex plus other non-hybrid folds, respectively. The system temperature limit is ~55 °C.

sizes and shapes, resulting in inconsistent entry rates between these structures. In order to quantify this difference, experiments were performed to investigate the relative event frequencies of the model sequences at known concentrations (Figure 8). To compare the capture rates of hybrid-1 and hybrid-2 folds, aliquots of the 8-Br-dG-substituted sequences were combined in known proportions and events were analyzed. Similarly, the 5′T triplex-forming sequence could be compared to the H1 sequence to obtain relative entry rates. Given the close similarity in size and shape of the hybrid species, it was not surprising to observe only slightly different entry rates between H1 and H2 (H1:H2 = 1.1:1). As expected, triplexes bearing a single-stranded tail exhibited much higher (~6.2 times) capture rates compared to bulky, fully folded quadruplexes. Corrected by these entry rate factors, thermal denaturation for the natural human telomere sequence was studied up to the temperature limit of the instrumentation (55 °C, Figure 9). By sorting the events using the knowledge ascertained above, we classified the relative distribution of the hybrid-1, hybrid-2, and triplex intermediates over a range from 25 to 55 °C. The relative abundance for the hybrid-1 and hybrid-2 folds decreased slowly compared to their initial values, while the triplex intermediates became more abundant. Due to the system limitation, we were unable to collect data above the T_m value of this sequence (~59.6 °C in 50 mM KCl + 950 mM LiCl); however, the changes were clear when the temperature approached ~50 °C and were consistent with spectroscopically observed results.[25] This study demonstrates that the hybrid folded ratio can be determined across a range of temperatures, while observing their partially denatured intermediates increase with rising temperature.

The studies described herein lead to the conclusion that the hybrid-1 structure is preferred over the hybrid-2 fold by a factor of 2.3 for the sequence 5′-TAGGG(TTAGGG)$_3$-3′ in dilute (5 μM) solution in the presence of K^+. This ratio is similar to the conclusion reached by optical tweezer studies of hybrid-1 vs hybrid-2,[60,61] but opposite that determined by NMR studies of the same sequence.[21] Perhaps the defining difference is that the NMR studies were conducted at 1000-fold higher concentration (5 mM), suggesting that intermolecular interactions,[62] in addition to solvent, cation, and the presence of overhanging tails, play a significant role in shifting the equilibrium between topological structures of similar energy.

dx.doi.org/10.1021/ja400973m | *J. Am. Chem. Soc.* 2013, 135, 8562–8570

CONCLUSIONS

Herein, we harnessed the size-limiting properties of the α-HL ion channel to detect subtle differences in the folded conformations of the human telomere sequence in solution. This has allowed us to determine that, under dilute conditions at 25 °C, the sequence 5′-TAGGG(TTAGGG)$_3$TT-3′ folds predominantly to hybrid-1 and hybrid-2 conformations in which the ratio is ∼2.3:1. This analysis was achieved by monitoring the intermediate residual currents that result from the fine molecular interactions of these two folds within the vestibule of α-HL. Furthermore, we observed current types that correspond to the triplex folding intermediates that have been proposed theoretically[54] and recently observed.[25,26] These observations allowed us to conduct temperature-dependent studies to monitor the initial unfolding progress of these species. These studies present an intriguing entry into the investigation of DNA structure and dynamics in a crowded environment.[17,27,57] Because ∼85% of all cancers show increased telomerase activity, an attractive molecular target to combat cancer is the G-quadruplexes of the human telomere. The present studies lay the groundwork for utilizing α-HL to monitor folding ratios for the human telomere sequence that will be affected by small-molecule binders, permitting a more rapid screening method for future cancer therapeutic agents.

METHODS

The oligodeoxynucleotides (ODNs) were synthesized by the DNA-Peptide Core Facility at the University of Utah with commercially available phosphoramidites (Glen Research, Sterling, VA), and processed following standard protocols (see Supporting Information). A custom-built, high-impedance, low-noise amplifier and data acquisition system was used for the ion channel recordings, and it was designed and constructed by Electronic Biosciences (EBS, San Diego, CA). A 100 kHz low-pass filter was used to collect and analyze the data with 500 kHz sampling rate. The sample i−t traces presented in Figures 2, 3, and 6 were refiltered to 20 kHz for presentation. Only events with duration longer than 10 μs (filter bandwidth) were extracted and analyzed using QUB 1.5.0.31 and fitted using Igor Pro 6.1. These experiments were conducted at 25 °C, except for the thermal denaturation profiling that was performed at 25, 30, 37, 45, 50, and 55 °C, respectively, with an error of ∼1 °C.

ASSOCIATED CONTENT

Supporting Information

Materials, experimental methods, CD spectra, T_m data, sample i−t traces, and event duration histograms. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

burrows@chem.utah.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors are grateful to Prof. Henry S. White for helpful discussions and to the National Institutes of Health for financial support (GM093099).

REFERENCES

(1) Moyzis, R. K.; Buckingham, J. M.; Cram, L. S.; Dani, M.; Deaven, L. L.; Jones, M. D.; Meyne, J.; Ratliff, R. L.; Wu, J. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6622−6626.
(2) Wright, W. E.; Tesmer, V. M.; Huffman, K. E.; Levene, S. D.; Shay, J. W. *Genes Dev.* **1997**, *11*, 2801−2809.
(3) de Lange, T. *Genes Dev.* **2005**, *19*, 2100−2110.
(4) de Lange, T. *Science* **2009**, *326*, 948−952.
(5) Grider, C. W.; Blackburn, E. H. *Cell* **1985**, *43*, 405−413.
(6) Dunham, M. A.; Nuemann, A. A.; Fasching, C. L.; Reddel, R. R. *Nat. Genet.* **2000**, *26*, 447−450.
(7) Wolkowitz, O. M.; Mellon, S. H.; Epel, E. S.; Lin, J.; Dhabhar, F. S.; Su, Y.; Reus, V. I.; Rosser, R.; Burke, H. M.; Kupferman, E.; Compagnone, M.; Nelson, J. C.; Blackburn, E. H. *PLoS One* **2011**, *6*, e17837.
(8) O'Donovan, A.; Tomiyama, A. J.; Lin, J.; Puterman, E.; Adler, N. E.; Kemeny, M.; Wolkowitz, O. M.; Blackburn, E. H.; Epel, E. S. *Brain, Behav., Immun.* **2012**, *26*, 573−579.
(9) Armanios, M.; Blackburn, E. H. *Nat. Rev. Genet.* **2012**, *13*, 693−704.
(10) Savage, S. A.; Bertuch, A. A. *Genet. Med.* **2010**, *12*, 753−764.
(11) Pereira, B.; Ferreira, M. G. *Curr. Opin. Oncol.* **2013**, *25*, 93−98.
(12) Gomez, D. E.; Armando, R. G.; Farina, H. G.; Menna, P. L.; Cerrudo, C. S.; Ghiringhelli, P. D.; Alonso, D. F. *Int. J. Oncol.* **2012**, *41*, 1561−1569.
(13) Xu, Y. *Chem. Soc. Rev.* **2011**, *40*, 2719−2740.
(14) De Cian, A.; Cristofari, G.; Reichenbach, P.; De Lemos, E.; Monchaud, D.; Teulade-Fichou, M. P.; Shin-Ya, K.; Lacroix, L.; Lingner, J.; Mergny, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 17347−17352.
(15) Maizels, N. *Nat. Struct. Mol. Biol.* **2006**, *13*, 1055−1059.
(16) de Lange, T. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 323−329.
(17) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature* **2002**, *417*, 876−880.
(18) Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. *Nat. Chem.* **2013**, *5*, 182−186.
(19) Lane, A. N.; Chaires, J. B.; Gray, R. D.; Trent, J. O. *Nucleic Acids Res.* **2008**, *36*, 5482−5515.
(20) Yang, D.; Okamoto, K. *Future Med. Chem.* **2010**, *2*, 619−646.
(21) Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2006**, *34*, 2723−2735.
(22) Phan, A. T.; Kuryavyi, V.; Luu, K. N.; Patel, D. J. *Nucleic Acids Res.* **2007**, *35*, 6517−6525.
(23) Phan, A. T.; Luu, K. N.; Patel, D. J. *Nucleic Acids Res.* **2006**, *34*, 5715−5719.
(24) Xu, Y.; Noguchi, Y.; Sugiyama, H. *Bioorg. Med. Chem.* **2006**, *14*, 5584−5591.
(25) Gray, R. D.; Buscaglia, R.; Chaires, J. B. *J. Am. Chem. Soc.* **2012**, *134*, 16834−16844.
(26) Koirala, D.; Mashimo, T.; Sannohe, Y.; Yu, Z.; Mao, H.; Sugiyama, H. *Chem. Commun.* **2012**, *48*, 2006−2008.
(27) Heddi, B.; Phan, A. T. *J. Am. Chem. Soc.* **2011**, *133*, 9824−9833.
(28) Branton, D.; Deamer, D. W.; Marziali, A.; Bayley, H.; Benner, S. A.; Butler, T.; Ventra, M. D.; Garaj, S.; Hibbs, A.; Huang, X.; Jovanovich, S. B.; Krstic, P. S.; Lindsay, S.; Ling, X. S.; Mastrangelo, C. H.; Meller, A.; Oliver, J. S.; Pershin, Y. V.; Ramsey, J. M.; Riehn, R.; Soni, G. V.; Tabard-Cossa, V.; Wanunu, M.; Wiggin, M.; Schloss, J. A. *Nat. Biotechnol.* **2008**, *26*, 1146−1153.
(29) Manrao, E. A.; Derrington, I. M.; Laszlo, A. H.; Langford, K. W.; Hopper, M. K.; Gillgren, N.; Pavlenok, M.; Niederweis, M.; Gundlach, J. H. *Nat. Biotechnol.* **2012**, *30*, 349−353.
(30) Cherf, G. M.; Lieberman, K. R.; Rashid, H.; Lam, C. E.; Karplus, K.; Akeson, M. *Nat. Biotechnol.* **2012**, *30*, 343−348.
(31) Henrickson, S. E.; Misakian, M.; Robertson, B.; Kasianowicz, J. J. *Phys. Res. Lett.* **2000**, *85*, 3057−3060.
(32) Gu, L. Q.; Braha, O.; Conlan, S.; Chelye, S.; Bayley, H. *Nature* **1999**, *398*, 686−690.
(33) Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13770−13773.
(34) Mathe, J.; Visram, H.; Viasnoff, V.; Rabin, Y.; Meller, A. *Biophys. J.* **2004**, *87*, 3205−3212.
(35) Muzard, J.; Martinho, M.; Mathe, J.; Bockelmann, U.; Viasnoff, V. *Biophys. J.* **2010**, *98*, 2170−2178.

(36) Jin, Q.; Fleming, A. M.; Burrows, C. J.; White, H. S. *J. Am. Chem. Soc.* **2012**, *134*, 11006−11011.
(37) Viasnoff, V.; Chiaruttini, N.; Bockelmann, U. *Eur. Biophys. J.* **2009**, *38*, 263−269.
(38) Sutherland, T. C.; Dinsmore, M. J.; Kraatz, H. B.; Lee, J. S. *Biochem. Cell Biol.* **2004**, *82*, 407−412.
(39) Liu, A.; Zhao, Q.; Krishantha, D. M. M.; Guan, X. *J. Phys. Chem. Lett.* **2011**, *2*, 1372−1376.
(40) Shim, J. W.; Gu, L.-Q. *J. Phys. Chem. B* **2008**, *112*, 8354−8360.
(41) Shim, J. W.; Tan, Q.; Gu, L.-Q. *Nucleic Acids Res.* **2008**, *37*, 972−982.
(42) Rotem, D.; Jayasinghe, L.; Salichou, M.; Bayley, H. *J. Am. Chem. Soc.* **2012**, *134*, 2781−2787.
(43) Patel, D. J.; Phan, A. T.; Kuryavyi, V. *Nucleic Acids Res.* **2007**, *35*, 7429−7455.
(44) Jung, Y.; Cheley, S.; Braha, O.; Bayley, H. *Biochemistry* **2005**, *44*, 8919−8929.
(45) Song, L.; Hobaugh, M. R.; Shustak, C.; Cheley, S.; Bayley, H.; Gouaux, J. E. *Science* **1996**, *274*, 1859−1866.
(46) Karsisiotis, A. I.; Hessari, N. M. a.; Novellino, E.; Spada, G. P.; Randazzo, A.; Webba da Silva, M. *Angew. Chem., Int. Ed.* **2011**, *50*, 10645−10648.
(47) Zhang, B.; Galusha, J.; Shiozawa, P. G.; Wang, G.; Bergren, A. J.; Jones, R. M.; White, R. J.; Ervin, E. N.; Cauley, C. C.; White, H. S. *Anal. Chem.* **2007**, *79*, 4778−4787.
(48) White, R. J.; Ervin, E. N.; Yang, T.; Chen, X.; Daniel, S.; Cremer, P. S.; White, H. S. *J. Am. Chem. Soc.* **2007**, *129*, 11766−11775.
(49) Vercoutere, W.; Winters-Hilt, S.; Olsen, H.; Deamer, D.; Haussler, D.; Akeson, M. *Nat. Biotechnol.* **2001**, *19*, 248−252.
(50) Vercoutere, W. A.; Winters-Hilt, S.; DeGuzman, V. S.; Deamer, D.; Ridino, S. E.; Rodgers, J. T.; Olsen, H. E.; Marziali, A.; Akeson, M. *Nucleic Acids Res.* **2003**, *31*, 1311−1318.
(51) An, N.; Fleming, A. M.; White, H. S.; Burrows, C. J. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 11504−11509.
(52) Purnell, R. F.; Mehta, K. K.; Schmidt, J. J. *Nano Lett.* **2008**, *8*, 3029−3034.
(53) Christensen, C.; Baran, C.; Krasnigi, B.; Stefureac, R. I.; Nokhrin, S.; Lee, J. S. *J. Pept. Sci.* **2011**, *17*, 726−734.
(54) Mashimo, T.; Yagi, H.; Sannohe, Y.; Rajendran, A.; Sugiyama, H. *J. Am. Chem. Soc.* **2010**, *132*, 14910−14918.
(55) Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. *Nat. Chem.* **2013**, *5*, 182−186.
(56) Lannan, F. M.; Mamajanov, I.; Hud, N. V. *J. Am. Chem. Soc.* **2012**, *134*, 15324−15330.
(57) Miller, M. C.; Buscaglia, R.; Chaires, J. B.; Lane, A. N.; Trent, J. O. *J. Am. Chem. Soc.* **2010**, *132*, 17105−17107.
(58) Chaires, J. B. *FEBS J.* **2010**, *277*, 1098−1106.
(59) Limongelli, V.; De Tito, S.; Cerofolini, L.; Fragai, M.; Pagano, B.; Trotta, R.; Cosconati, S.; Marinelli, L.; Novellino, E.; Bertini, I.; Randazzo, A.; Luchinat, C.; M., P. *Angew. Chem., Int. Ed.* **2013**, *52*, 2269−2273.
(60) Yu, Z.; Mao, H. *Chem. Rec.* **2013**, *13*, 102−116.
(61) Dhakal, S.; Cui, Y.; Koirala, D.; Ghimire, C.; Kushwaha, S.; Yu, Z.; Yangyuoru, P. M.; Mao, H. *Nucleic Acids Res.* **2013**, *41*, 3915−3923.
(62) Do, N. Q.; Phan, A. T. *Chemistry* **2012**, *18*, 14752−14759.