**Triptycene-Based Small Molecules Modulate (CAG)-(CTG) Repeat Junctions**

<table>
<thead>
<tr>
<th><strong>Journal</strong></th>
<th><em>Chemical Science</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manuscript ID</strong></td>
<td>SC-EDG-05-2015-001595.R1</td>
</tr>
<tr>
<td><strong>Article Type</strong></td>
<td>Edge Article</td>
</tr>
<tr>
<td><strong>Date Submitted by the Author</strong></td>
<td>26-May-2015</td>
</tr>
<tr>
<td><strong>Complete List of Authors</strong></td>
<td>Barros, Stephanie; University of Pennsylvania, Chemistry Chenoweth, David; University of Pennsylvania, Department of Chemistry</td>
</tr>
</tbody>
</table>
Triptycene-Based Small Molecules Modulate (CAG)-(CTG) Repeat Junctions

Stephanie A. Barros and David M. Chenoweth

Nucleic acid three-way junctions (3WJs) play key roles in biological processes such as nucleic acid replication in addition to being implicated as dynamic transient intermediates in trinucleotide repeat sequences. Structural modulation of specific nucleic acid junctions could allow for control of biological processes and disease states at the nucleic acid level. Trinucleotide repeat expansions are associated with several neurodegenerative diseases where dynamic slippage is thought to occur during replication, forming transient 3WJ intermediates with the complementary strand. Here, we report triptycene-based molecules that bind to a d(CAG)-(CTG) repeat using a gel shift assay, fluorescence-quenching and circular dichroism.

Introduction

Nucleic acid junctions play important roles in biological processes and serve as key structural motifs in nanotechnology and aptamer-based sensing applications. In biology, three-way junctions (3WJs) are found as transient intermediates during replication, recombination, and DNA damage repair. Junctions are also present in several viral genomes, such as HIV-1, HCV, and adeno-associated virus in addition to playing key roles in viral assembly. Nucleic acid junctions are also prevalent in the emerging field of DNA and RNA nanotechnology where the bacteriophage phi29 pRNA containing RNA three-way junctions provide a particularly impressive example. Furthermore, they occur in trinucleotide repeat expansions found in unstable genomic DNA associated with neurodegenerative diseases. The development of structure and sequence-specific nucleic acid binding molecules remains an important challenge in chemical biology. The ability to target specific motifs using small molecules would allow for the precise control of biological processes and the possibility of modulating disease states.

DNA trinucleotide repeats are present throughout the genome. Expansions of these repeats, however, are associated with a number of neurodegenerative diseases, including Huntington’s disease, spinobulbar muscular atrophy, and myotonic dystrophy. Current models of triplet repeat expansion disease suggest slippage during DNA synthesis by the formation of dynamic DNA hairpin structures. As the length of the repeat increases, the growing hairpin structure gains thermodynamic stability, with repeat length providing an important positively correlated diagnostic for disease severity. Slipped-out (CAG)$_n$-(CTG)$_n$ repeats have been implicated in the pathogenesis of triplet repeat expansion diseases such as Huntington’s disease and several others. These “slipped-out” regions are dynamic and occur along the duplex, forming three-way junctions. Current models suggest that one arm of the junction contains the excess repeats while the other arms maximize complementary pairing between adjacent strands. A dynamic ensemble of conformations are possible at the slipped junction interface, where base pairing interactions differ with each state. The slipped-out (CAG)$_n$ repeat in Figure 1 has been shown to contain one unpaired base at the center of the junction. Previous NMR studies have demonstrated that this sequence can adopt a multitude of conformations, where dynamic single nucleotide bulges at the junction interface interconvert between structures. Small molecule probes could provide important tools for gaining molecular level insight into the dynamics and repair processes associated with trinucleotide repeat junctions; however, probes of this kind are currently unknown.

Fig. 1 Slipped DNA junctions formed by (CAG)-(CTG) repeats.

Recently, we reported a new class of triptycene-based three-way junction (3WJ) binders. Here, we apply the triptycene scaffold as a first step toward developing new tools to recognize trinucleotide repeat junctions. We assessed the ability of our new triptycene based molecules to modulate the structure of d(CAG)$_n$. 
repeats using gel shift, fluorescence quenching, and circular dichroism (CD).

Results and Discussion

Utilizing a well-studied (CAG)-(CTG) repeat sequence known to form slipped-DNA three-way junctions, we developed a competitive inhibitor-based gel shift assay. This assay was inspired by previous work from our laboratory and seminal studies from the Plaxco and Ricci laboratories.\(^1,7\) We utilized the junction forming (CAG)-(CTG) repeat sequence (TNR) and an inhibitor strand (I10) shown in Figure 2a. The optimal inhibitor strand was 10 base pairs long and complementary to the 5’-end of the junction. Titration of inhibitor strand I10 into the folded junction resulted in a concentration dependent supershift of the band corresponding to hybridization of the TNR junction with I10, forming a larger molecular weight complex (TNR-I10) (Figure 2b). Addition of small molecules capable of binding the TNR junction leads to I10 strand displacement from the TNR-I10 complex in a concentration dependent manner (Figure 3a).

Inspired by nucleic acid binding proteins, we synthesized a new class of triptycene molecules containing amino acids commonly found at the protein-nucleic acid interface. Analysis of these interfaces reveals an abundance of positively charged amino acid residues such as arginine, lysine, and histidine. We synthesized triptycene derivatives substituted with arginine, lysine, and histidine (Trip\(^2\)). Compounds 2-4 were assessed using a competitive displacement gel shift assay and compared to Trip 1, which was previously shown to significantly stabilize 3WJs (Figure 3c). Folded TNR 3WJ was incubated with I10, followed by addition of triptycenes 1-4 at the same concentration.

Fig. 2 (a) Schematic of gel shift assay. The folded TNR 3WJ was incubated with different concentrations of an inhibitor strand (I10) complementary to the 5’-end, resulting in formation of TNR-I10. Non-denaturing polyacrylamide gel ran in 1X TBE buffer at 4 °C.

Fig. 3 (a) Schematic of gel shift assay. The folded TNR 3WJ was incubated with an inhibitor strand complementary to the 5’-end, opening the junction structure (TNR-I10). Addition of triptycene results in reformation of the junction (TNR-Trip). (b) Structures of triptycene derivatives (Trip\(^1\)-4). (c) Gel shift assay where TNR-I10 was incubated with triptycene derivatives at a constant concentration. (d) A plot of the difference in band intensities of complex relative to 3WJ. (e) Gel shift assay in the presence and absence of Trip 3 and Trip 4. Samples contained 0.5 μM TNR alone (where minus sign is indicated) or 0.5 μM TNR and 1.5 μM I10 (where a plus sign is indicated). Increasing concentrations of Trip 3 (lane 3, 0 μM; lane 4, 0.01 μM; lane 5, 0.10 μM; lane 6, 0.50 μM; lane 7, 1.0 μM; lane 8, 5.0 μM) and Trip 4 (lane 11, 0 μM; lane 12, 0.01 μM; lane 13, 0.10 μM; lane 14, 0.50 μM; lane 15, 1.0 μM; lane 16, 5.0 μM; lane 17, 10.0 μM). Lanes 1 and 9 are loaded with a 25 base pair DNA ladder in which the band present corresponds to 25 bases. Free TNR junction and TNR-I10 complex are indicated. Non-denaturing polyacrylamide gel ran in 1X TBE buffer at 4 °C.

Next, we tested the ability of Trip 3 and 4 to induce fluorescence-quenching upon junction formation using a double labelled (CAG)-(CTG) repeat oligonucleotide. The TNR oligonucleotide was labelled with a FAM fluorophore on the 5’-end.
Fig. 4 Fluorescence-quenching assay and circular dichroism (CD). (a) TNR 3WJ labeled with a fluorophore and quencher. When folded, low fluorescence is observed. Addition of inhibitor I10 opens the junction, resulting in an increase in fluorescence (TNR*-I10). Addition of triptycene reforms the junction, resulting in quenching of fluorescence. (b) Titration of I10 to the folded junction results in an increase in fluorescence. Fluorescence assay was conducted in 50 mM sodium phosphate buffer at pH 7.2. (c) Titration of Trip 3 and Trip 4 to TNR*-I10 results in a decrease in fluorescence. (d) Temperature-dependent circular dichroism of the TNR junction. Temperature-dependent CD in the presence of Trip 3 (e) and Trip 4 (f). Circular dichroism measurements were conducted in 50 mM sodium phosphate buffer at pH 7.2.

Conclusions

In summary, we have described triptycene-based molecules that bind to d(CAG)-(CTG) trinucleotide repeats. Trip 3 and Trip 4 bind to a model (CAG)-(CTG) repeat as determined by gel shift and fluorescence-quenching experiments. The CD spectra are also consistent with enhanced helicity of the slipped out junction upon addition of Trip 3 and 4. This new class of nucleic acid binding small molecule may serve as inspiration for creating valuable probes for diseases associated with trinucleotide repeat expansions. Trinucleotide repeat nucleic acid sequences are associated with a large number (>30) of inherited human muscular and neurological diseases. The trinucleotide repeat tract length is dynamic and often correlates with disease severity, where short stable tracts are commonplace in the non-affected population. Longer unstable triplet repeat tracts are more prone to expansion as opposed to contraction, in addition to being predisposed to generational transmission. Stabilization of a particular structure could lead to altered repair outcomes as well as recruitment of proteins. Small molecule probes will provide valuable tools to study these processes. Small molecule Binding and stabilization or modulation of these dynamic structures could lead to the development of therapeutic agents for their associated diseases.
Acknowledgements

This work was supported by funding from the University of Pennsylvania. Instruments supported by the National Science Foundation and the National Institutes of Health including HRMS (NIH RR-023444) and MALDI-MS (NSF MRI-0820996). S.A.B. wishes to thank the NIH for funding through the Chemistry-Biology Interface Training Program (T32 GM07133).

References


