Introducing a New Paradigm of "Socket-plug" Complementarity for Specific DNA-DNA Recognition and Binding

by

Yutong Huang

B.A., Simon Fraser University, 2019

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Molecular Biology and Biochemistry
Faculty of Science

SIMON FRASER UNIVERSITY
Spring 2022

Declaration of Committee

Name:	**Yutong Huang**
Degree:	**Master of Science**
Title:	**Introducing a New Paradigm of "Socket-plug" Complementarity for Specific DNA-DNA Recognition and Binding**
Committee:	**Chair: Lorena Braid** Assistant Professor, Molecular Biology and Biochemistry
	Dipankar Sen Supervisor Professor, Molecular Biology and Biochemistry
	Lisa Craig Committee Member Professor, Molecular Biology and Biochemistry
	Peter Unrau Committee Member Professor, Molecular Biology and Biochemistry
	Edgar Young Examiner Associate Professor, Molecular Biology and Biochemistry

Abstract

Nucleic acids have proved to be powerful building blocks in nanotechnology owing to their highly predictable and controllable self-assembling property. The classical Watson-Crick complementarity has been utilized to build complex 2D and 3D architectures and nanodevices. In addition to the iconic double helix generated from Watson-Crick strand pairing, DNAs are also known to be able to form alternative pairing schemes, giving rise to form DNA triplexes and G-quadruplexes, which in turn introduce novel possibilities for the engineering of DNA architectures. Herein, we describe a class of “sticky-ended” DNA triplex-quadruplex (TQ) composites that recognize and bind to each other via a wholly new paradigm of so-called “socket-plug” (as opposed to Watson-Crick) complementarity. The formation of these “socket-plug” hybrid composites is dependent on the presence of specific counterions under modestly acidic pH; the structures can in turn be disassembled by changes to either one of these requirements. Using gel electrophoresis and Förster Resonance Energy Transfer experiments, I have demonstrated the formation of different “socket-plug” hybrids in the presence of K^+ versus Na^+ counter-cations. Such a sensitivity to specific counterions is predicted to be a highly useful hybridization property, that will find wide application in nanotechnology, bioanalytical chemistry, as well as in other fields.

Keywords: DNA triplexes; G-quadruplexes; complementarity; ion sensitivity; DNA nanotechnology

Dedication

To my lovely feline angel, Captain Chan.

Acknowledgements

I would first like to express my deepest gratitude to my supervisor, Dr. Dipankar Sen, for the great opportunity to conduct my research experiments in his exceptional lab that I was fortunate to be a part of. When I first joined the Sen lab as an undergraduate student, your guidance and willingness to engage with your students made it an easy decision to continue my studies as a graduate student.

Next, I would like to thank all the past and present members of the Sen lab, especially Kun and Prince. When I first joined the lab, I had absolutely no experience in working in a research lab. It was Kun who taught me everything, starting from the basic lab techniques (such as making gels) to critical thinking in experimental designs (always think of more controls). His patience, kindness, and humour (and funny podcasts and singings) will forever be remembered by my heart, as they made the tough days easier and the good days even better. The work presented in this thesis would not have been possible without Prince, as he is the person who came up with this brilliant project idea and did all the preliminary hard work to lay a solid foundation for me to build upon. He was the first person I always turned to whenever I encountered a problem, and he would always help me troubleshoot. Throughout this project, he also contributed a lot of great ideas and thoughts about the experiments which helped improve the overall quality of the work. It is my pleasure to work with such a friendly, supportive, and helpful group of people, and I can confidently say that this work would not have been possible without all of you.

I would also like to thank Dr. Lisa Craig and Dr. Peter Unrau, for being a part of my supervisory committee and providing valuable feedback and support throughout my degree. I would especially like to thank Elaine Wang, Kennedy Chen, and Kristen Kong, for sharing the ups and downs of graduate school.

These acknowledgments wouldn’t be completed without a final thank you to my furry roommate, Captain the cat. You have always been the delight of my days ever since you came into my life. Your gentle kneading with your soft paws and the smooth purring sound you make while napping over my legs are the best comforting things in the world that can instantaneously take my stress away.

Table of Contents

List of Tables

List of Figures

List of Acronyms

AFM	Atomic Force Microscopy
CD	Circular Dichroism
DMS	Dimethyl Sulphate
DNA	Deoxyribonucleic Acid
GQ/GQs	G-Quadruplex / G-Quadruplexes
NMR	Nuclear Magnetic Resonance
RNA	Ribonucleic Acid
T_m	Melting Temperature
TQ	Triplex-Quadruplex
W-C	Watson-Crick

Chapter 1. General Introduction

1.1. DNA

The story of deoxyribonucleic acid (DNA) began in 1869, when a Swiss physiological chemist, Friedrich Miescher, obtained the first crude purification of DNA (1). While uncovering the molecular basis of cellular life had been one of the most fundamental problems in the scientific community at that time, the significance of Miescher's discovery of nucleic acids was generally underappreciated. For many years, it was believed by most scientists that proteins were the macromolecules responsible for the genetic properties of living organisms. DNA, which simply seemed less structurally complex than proteins, was thought to be inadequate for specifying the necessary information needed to make up a genome.

In the decades following Miescher's initial discovery of DNA, a handful of scientists – notably, Oswald T. Avery and Erwin Chargaff – continued to carry out research efforts that revealed more and more details about this obscure substance, and they eventually succeeded in demonstrating that DNA was the carrier of genetic information (2,3). Since then, a widespread interest in DNA began. The year 1953 marked one of the most important discoveries in the history of modern science – the groundbreaking construction of the three-dimensional, double-helical model for the structure of DNA by James D. Watson and Francis Crick (4), based on the crucial X-ray fiber diffraction work done by Rosalind Franklin (5). This astonishing work by Watson and Crick, together with accumulating evidence on the biological significance of DNA, led to a transformation in the landscape of modern molecular biology.

1.2. Canonical Secondary Structure of DNA: the Watson-Crick Duplex

The Watson-Crick double-helical model for the secondary structure of DNA is of such importance because, in addition to providing the structure of what is arguably the central molecule of life, it suggests the molecular mechanism of heredity. In the Watson-Crick model, DNA consists of two complementary polydeoxynucleotide strands that wind around each other in antiparallel fashion to form a right-handed double helix. Each

nucleotide monomer, the fundamental repeating unit in DNA, is composed of a phosphate group, a deoxyribose sugar, and a nitrogenous base (Figure 1.1). In a given strand of DNA, successive nucleotides are covalently linked through phosphodiester bonds, in which the 5'-phosphate group of one nucleotide unit is joined to the 3'-hydroxyl group of the next nucleotide. The alternating phosphate groups and deoxyriboses together form the hydrophilic backbones, or the external surfaces, of the double helix, while the hydrophobic bases of both strands occupy the interior core of the helix.

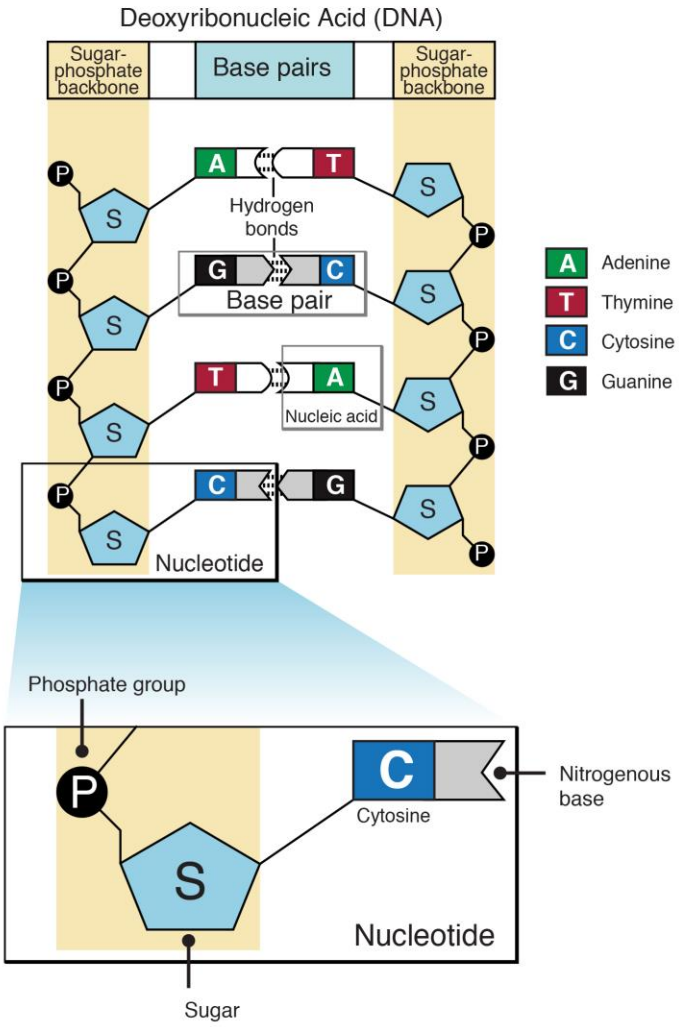


Figure 1.1. Nucleotides as the building blocks of deoxyribonucleic acid (DNA). Nucleotides are the monomeric units of nucleic acid polymers. A DNA nucleotide consists of a deoxyribose sugar molecule attached to a phosphate group and a nitrogen-containing aromatic base. Courtesy: National Human Genome Research Institute

The naturally occurring nitrogenous bases for DNA are of four types: adenine (A), guanine (G), cytosine (C), and thymine (T). These nitrogenous bases are categorized into two classes: the purines (A and G), each with two fused heterocyclic rings, and the pyrimidines (C and T), each with a single heterocyclic ring (Figure 1.2).

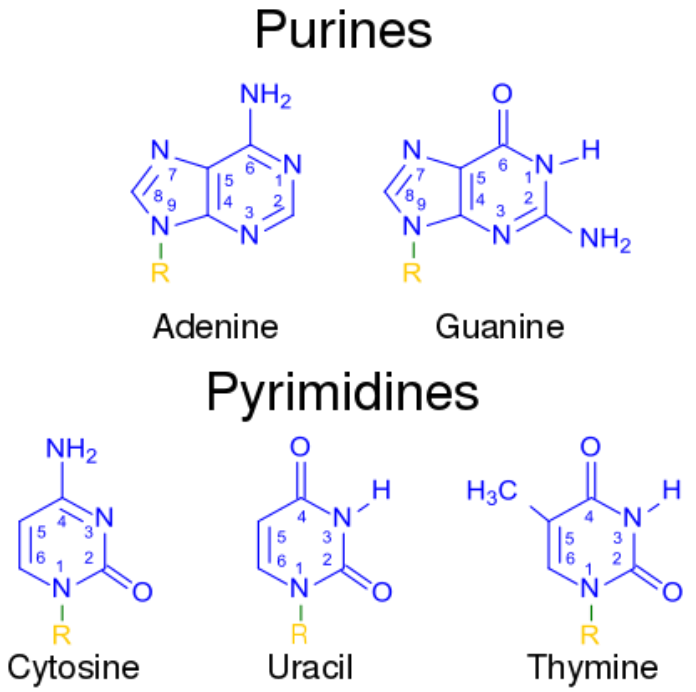


Figure 1.2. Canonical purine and pyrimidine bases of nucleic acids.
The four naturally occurring nucleobases in DNA are adenine (A), guanine (G), cytosine (C), and thymine (T). In RNA, the base uracil (U) takes the place of thymine.

The specific pairing of nucleotide bases is one of the most remarkable features of the Watson-Crick model of DNA, which posits that a Watson-Crick base pair consists of a purine hydrogen-bonded with a pyrimidine. Only a “right” pair of bases (such as G with C, and A with T) with appropriate geometrical correspondence, is able to form stable hydrogen bonds between the H-bond donor and the acceptor functionalities on the bases. More specifically, in this pairing scheme, an adenine always pairs with a thymine by forming two hydrogen bonds, whereas a guanine pairs with a cytosine via three hydrogen bonds (Figure 1.3). Other combinations of bases, for instance, purine-purine pairings are energetically unfavourable because the molecules are forced to be too close to each other, resulting in overlap repulsion; pyrimidine-pyrimidine pairings are also unfavourable because the molecules are too far apart within the standard double-helix geometry to establish hydrogen bonds. The only other purine-pyrimidine pairing option

would be AC and GT, which are mismatches because the patterns of hydrogen donors and acceptors in the respective partners do not correspond. All of these non-Watson-Crick base pairs would require considerable re-orientation of the sugar-phosphate chain, thereby significantly distorting and destabilizing the double helix structure. The unique and specific hydrogen-bonding patterns between the two Watson-Crick base pairs not only contribute to the overall stability of the DNA structure but more importantly, make DNA eminently suited for genetic information storage and in active processes such as DNA replication.

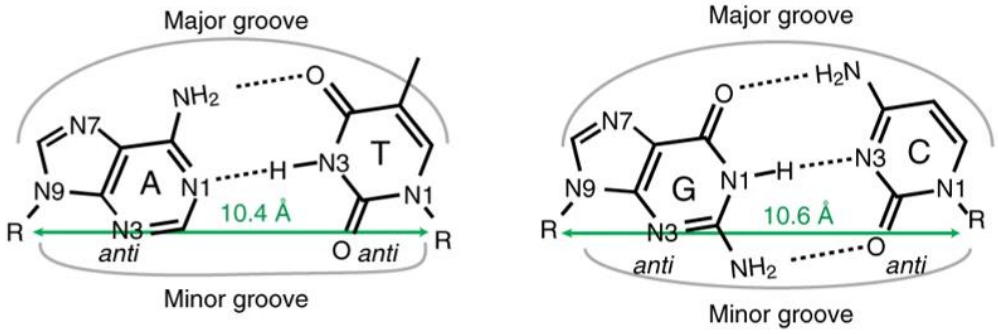


Figure 1.3. The canonical Watson-Crick AT and GC base pairs.
Two hydrogen bonds are formed between A and T in the AT base pair while three are formed in between G and C in the GC base pair. The near identical overall dimensions of base pairs of both types allow the formation of uniform helical conformations of the two polynucleotide strands. R indicates the position of the deoxyribose. Adapted from (6)

The Watson-Crick model of DNA, most commonly found in a helical architecture, known as B-DNA, represents the form of DNA crystallized in the presence of sodium counterion under highly humid conditions. B-DNA is also by far the most prevalent form of DNA in the cell. Double helical DNAs can assume other distinct conformations *in vitro* under certain conditions (Figure 1.4). At lower humidity conditions, typically below 75%, B-DNA undergoes a reversible conformational change to the so-called A-DNA, which is a wider and flatter right-handed helix. In biology, A-form DNA is typically formed when DNA is under dehydrating conditions, such as extreme desiccation of bacteria(7). Crystallographic analyses of A-DNA structure and protein-DNA complexes suggest that A-DNA may also form upon binding to certain proteins, such as the human immunodeficiency virus type 1 reverse transcriptase (HIV-1RT), a DNA polymerase that cuts the DNA at the (O3'-P) phosphodiester linkage(8). The polymerase-induced B-to-A

conformational transition has been identified in crystallographic studies of HIV reverse transcriptase bound to DNA(9). It is believed that such conformational change at the polymerase active site may improve the base pair fit in the nascent template-primer duplex due to a lower sequence-dependent structural variability in A-DNA compared with B-DNA. At high salt concentrations (such as 1.8 M $NaClO_4$, 2.5 M NaCl, or 0.7 M $MgCl_2$)(10), DNAs with alternating purine and pyrimidine base sequences can adopt a radically different double-helical conformation, called Z-DNA, which is a thinner and more elongated left-handed double helix. The biological importance of Z-DNA was underappreciated for a long time after its accidental discovery. It has now become clear that Z-DNA can be bound with high affinity by the Zα domain from the dsRNA editing protein ADAR, whose loss-of-function mutations can lead to a number of human diseases(11). Since the B form of double-helix is the most stable structure for random-sequence DNA under physiological conditions, it has been considered as the standard point of reference and exploited intensively in all studies of the properties of DNA.

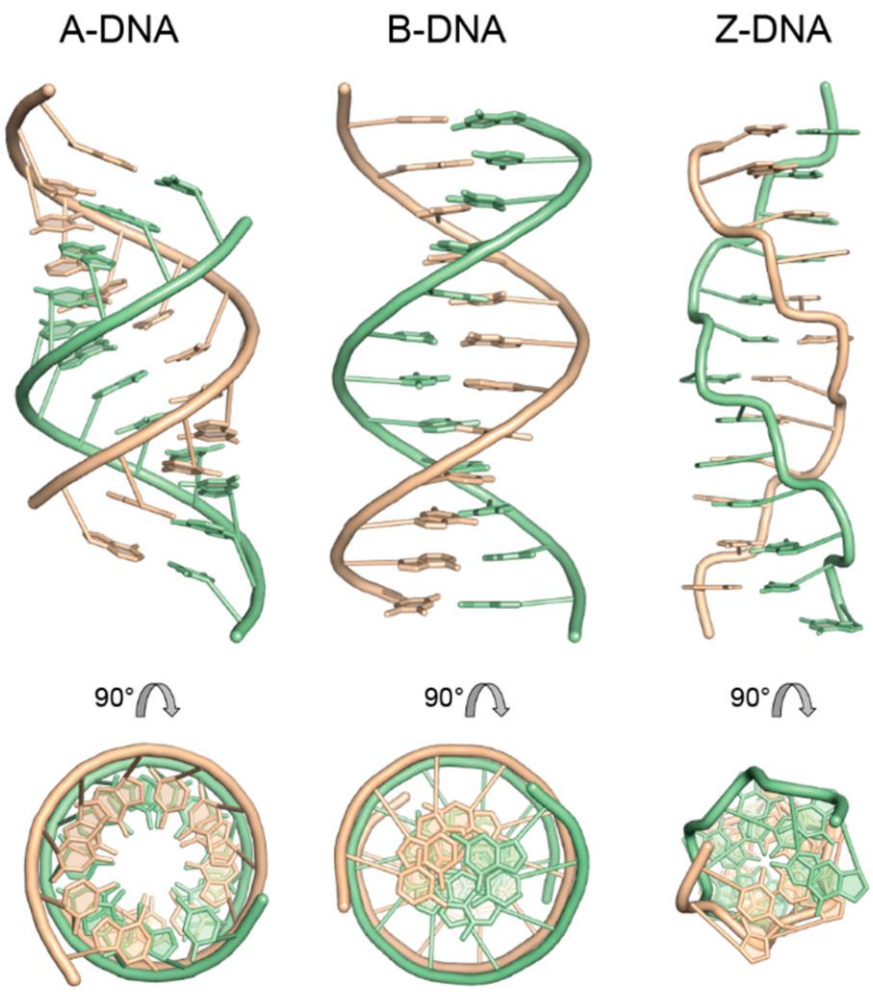


Figure 1.4. A, B, and Z forms of double-helical DNA based on X-ray fiber diffraction analysis.

The top row presents the side view of different forms of DNA where the helix axis is vertical. The bottom row presents the view down the helix axis after a 90° rotation from the top. The A-DNA and B-DNA are 12-bp DNA fragments with arbitrary sequence d(ATCGATCGATCG), whereas the Z-DNA contains an alternating sequence of d(CGCGCGCGCGCG). The models were generated based on the X-ray fiber diffraction determined by S. Arnott, Oxford University. In A-DNA, the base pairs are inclined to the helix axis by 15 to 20° and the axis has a hollow core. In B-DNA, the helix axis passes through the base pairs so that the inclination angle is approximately 0° and the helix has a solid core. In Z-DNA, the helix is left-handed, and the sugar-phosphate chains follow a zigzag course. Adapted from (12).

1.3. Alternative Base Pairing Scheme: Hoogsteen Base Pairing

In the decades following Watson and Crick's initial proposal of DNA's double-helix model, with the advancement in X-ray crystallography, scientists discovered that apart from the classical Watson-Crick base pairing interactions, DNA nucleobases can

also adopt alternative modes for pairing by having different hydrogen bonding patterns. One such alternative base pairing scheme, termed Hoogsteen base pairing, was first reported by Karst Hoogsteen in the 1960s, when he used single crystal X-ray analysis to determine the co-crystal structure of 9-methyladenine and 1-methylthymine (13). Hoogsteen observed that the adenine base was flipped upside down by a 180-degree rotation of the base around the glycosidic bond, changing the base from *anti* to a *syn* conformation (Figure 1.5). A few years later, the Hoogsteen base pairing scheme was also observed in G-C^+ base pairs in poly(dG)-poly(dC) at acidic conditions (pH 3 – 4) (14). As in the A-T base pairs, formation of the G-C^+ base pair also entails the flipping of the guanine from an *anti* to a *syn* conformation (Figure 1.5). In Watson-Crick base pairs, the C1-N6 face of the purines is involved in hydrogen bonding with the N3-C4 face of the pyrimidines. In contrast, the Hoogsteen base pairing scheme utilizes the C6-N7 face of purine bases to form hydrogen bonds with the N3-C4 face of the pyrimidines. In the G-C^+ base pairs, the formation of the hydrogen bond with N7 of the flipped guanine requires the N3 of cytosine to be protonated. For this reason, the formation of G-C^+ Hoogsteen base pairs requires low pH conditions.

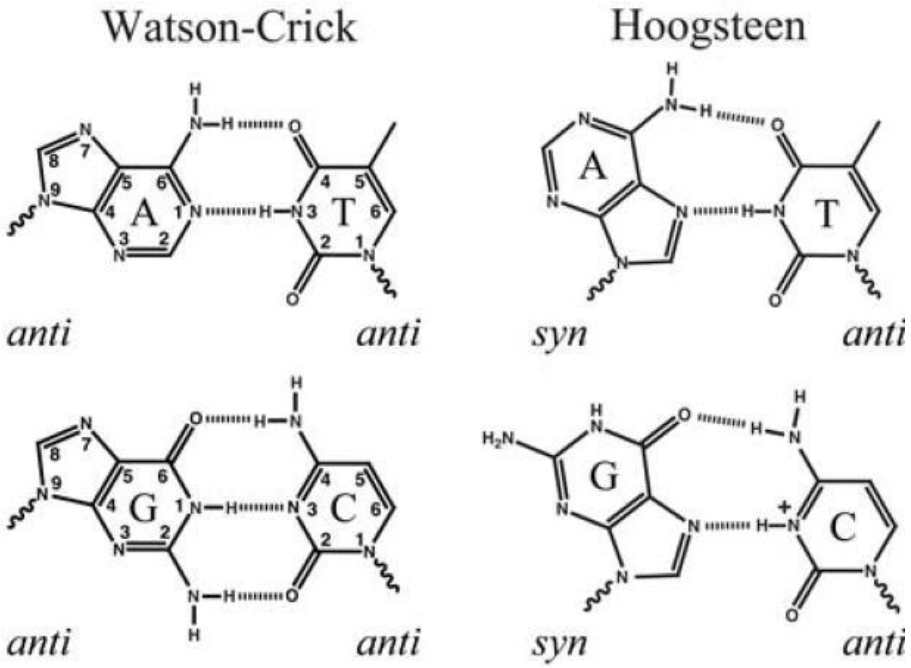


Figure 1.5. A-T and G-C base pairs in Watson-Crick and Hoogsteen pairing schemes.

An adenine base forms two hydrogen bonds with a thymine base in both Watson-Crick and Hoogsteen base pairs. By contrast, a guanine base forms three hydrogen bonds with a cytosine in the Watson-Crick scheme, the G-C Hoogsteen base pair only involves two hydrogen bonds, one of which requires the N3 position of cytosine to be protonated. The Hoogsteen base pairs are

formed by the flipping of purine bases by 180-degree around the glycosidic bond, changing the base from a *syn* to an *anti* glycosidic conformation. Adapted from (15).

For a long time since the discovery of this novel base-pairing geometry, its biological relevance remained elusive. However, recent studies have provided compelling experimental evidence suggesting that Hoogsteen base pairs play important roles in a variety of biological processes encompassing DNA replication, transcription, gene expression, recombination, DNA repair, and telomere length maintenance (16). Hoogsteen base pairing significantly expands the structural and functional versatility of DNA beyond that which can be achieved based solely on Watson-Crick base pairing, and more importantly, gives rise to several non-canonical secondary structures of DNA.

1.4. Non-Canonical Secondary Structures of DNA

From the earliest fiber X-ray diffraction studies to the more recently developed NMR studies, DNA has manifested its polymorphic nature and the flexibility to adopt different conformations under certain environmental conditions and sequence contexts. In addition to the iconic Watson-Crick double helix, DNA can also exist in other helical forms involving canonical Watson-Crick base pairs (such as hairpins and cruciform), as well as non-canonical base-pairing (including triplexes, and G-quadruplexes). Although these structures of DNA have been known for decades, for a long time their biological significance remained obscure and therefore they were generally considered fascinating phenomena but with little biological relevance or practical use. With decades of research efforts by different scientists and mounting experimental evidence, there is now little doubt that these structures almost certainly form under physiological conditions within cells and likely participate in important biological processes, including genome recombination, regulation of gene expression, and cancer cell growth. There is growing interest also in utilizing these non-canonical DNA structures for bio/nanotechnology.

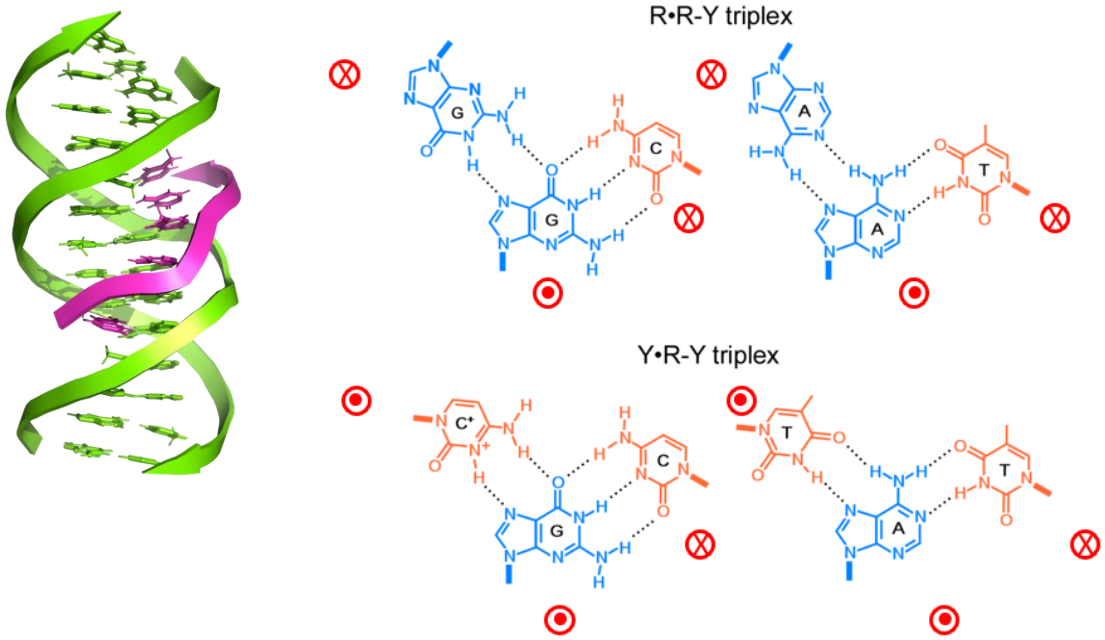
1.4.1. DNA Triplexes

One of several notable non-canonical secondary structures are DNA triplexes, first discovered by Felsenfeld and Rich in 1957 (17). A DNA triple-helix is a three-stranded complex generated by the binding of a third strand within the major groove of a Watson-Crick duplex in which one strand is wholly purine (“R”) and the other wholly

pyrimidine (“Y”) (Figure 1.6.a). Depending on the content of the third strand, canonical DNA triple helices fall into two broad classes, namely, YR*Y and YR*R (where the asterisk indicates the Hoogsteen/reverse Hoogsteen interaction) (Figure 1.6.b).

In the YR*Y triplex, a pyrimidine-containing third strand binds via Hoogsteen hydrogen-bonding to the purine strand of the W-C duplex in a parallel orientation, generating TA*T and CG*C^+ base triples (Figure 1.6.c). An important feature of the YR*Y triplexes is the formation of the CG*C^+ triple requires the protonation of the N3 of cytosine in the third strand to form Hoogsteen base-pairing with guanine residues in dsDNA. Therefore, pH is an important factor in determining the stability of YR*Y triplexes containing CG*C^+. Although this may seem to be a limitation at the first glance, the pH dependence of the parallel motif can actually be a useful property as it allows control over the assembly and/or disassembly of the triplex by simply adjusting the solution pH.

(a) **(c)**



(b)

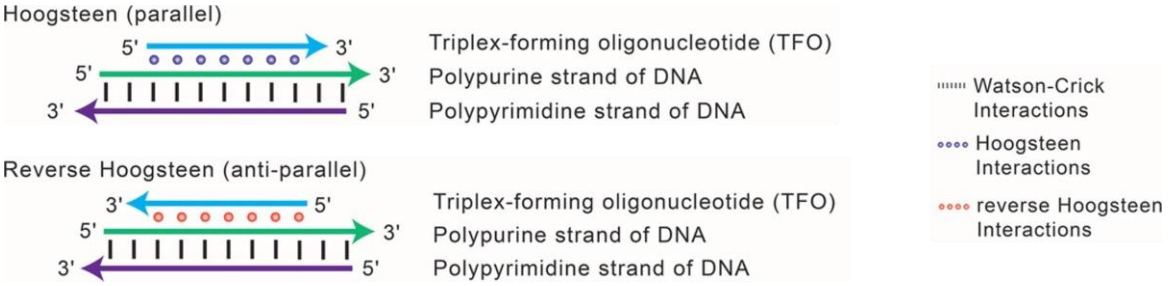


Figure 1.6. DNA triplexes and canonical triples/triplets.
(a) NMR structure of a DNA triplex (PDB: 1BWG). The green ribbons represent the Watson-Crick duplex DNA while the purple ribbon represents the triplex-forming oligonucleotide bound to the duplex via the major groove. Adapted from (18). The orientations of the strands are represented by either an "X" or a dot, meaning 5' or 3', respectively. (**b)** Parallel and antiparallel triplex configurations, adapted with permission from (19). Copyright John Wiley and Sons **(c)** In the YR*R triplex, the canonical triplets are CG*G and TA*A. In the YR*Y triplex, the canonical triplets are CG*C^+ and TA*T. The C^+ indicates a protonated cytosine. Adapted from (20).

By contrast, to their YR*Y counterparts, the YR*R triplexes do not require any protonation of bases. In these triplexes, the purine-containing third strand is found in antiparallel orientation to the purine-rich strand in the duplex. Another feature of YR*R triplexes is requirement for reverse Hoogsteen (as opposed to Hoogsteen) base pair formation within the canonical CG*G and TA*A base triples, to enable reasonable stacking interactions (Figure 1.6.c). While YR*R triplexes are generally considered more versatile than YR*Y triplexes since their stability is independent of pH and they can tolerate more diverse pairing schemes (such as TA*T), having long and uninterrupted G-tracts in the third strand sequence in YR*R is undesirable, as such tracts have the tendency to form highly stable G-quadruplexes in preference to triplexes.

Despite these differences, the two classes of triplex nevertheless share some fundamental features: (a) the Watson-Crick duplex component involved in triplex-formation must have a homopurine sequence in one strand. (b) The orientation of the two chemically homologous (i.e. either polypurine or polypyrimidine) strands within a given triplex has them positioned antiparallel to each other (21). Furthermore, (c) both parallel and antiparallel triplexes are stabilized by the presence of divalent cations such as Mg^{2+}, Ca^{2+}, and Zn^{2+}, or indeed, more highly charged cations such as formed by biological polyamines such as spermidine and spermine. Such counterions presumably reduce the electrostatic repulsion between the negatively charged phosphate sugar

backbones of the three component strands, allowing the triplexes to form more readily (17).

The discovery of the DNA triplex has proved to be of great importance to modern molecular biology, as it presents a novel strategy for targeting specific sequences of genomic DNA through oligonucleotide-based triple-helix formation. Such a property has spawned exciting studies by which triplex DNA has been characterized and investigated in detail for its potential applications in gene therapy and in nanotechnology.

1.4.2. G-quadruplexes

Another class of non-canonical DNA secondary structure that relies on the Hoogsteen base-pairing scheme is G-quadruplexes (GQs). It has been known since the early research in the 1960s by Gellert and Davies that guanine residues have the inherent propensity to self-aggregate, which gives rise to another peculiar non-canonical DNA secondary structure named G-quadruplex (22). Subsequent biochemical experiments in the late 1980s by Sen and Gilbert(23,24) and others(25,26) demonstrated that oligonucleotides found in eukaryotic telomeres containing runs of adjacent guanines can self-associate and form a four-stranded G-quadruplex structure *in vitro*.

The building blocks of G-quadruplexes are G base-quartets, which are square planar hydrogen-bonded structures formed by the association of four guanines via Hoogsteen hydrogen bonding by guanine-rich DNA or RNA sequences (Figure 1.7.a). Within one G-quartet, the N1 of the first guanine acts as a hydrogen donor to pair with the O6 on the second guanine, and simultaneously, the N2 of the first guanine acts as a hydrogen acceptor to pair with the N7 of the second guanine. The net result is a total of eight hydrogen bonds per quartet. The stacking of multiple planar G-quartets on top of one another eventually gives rise to four-stranded helical structures (Figure 1.7.b).

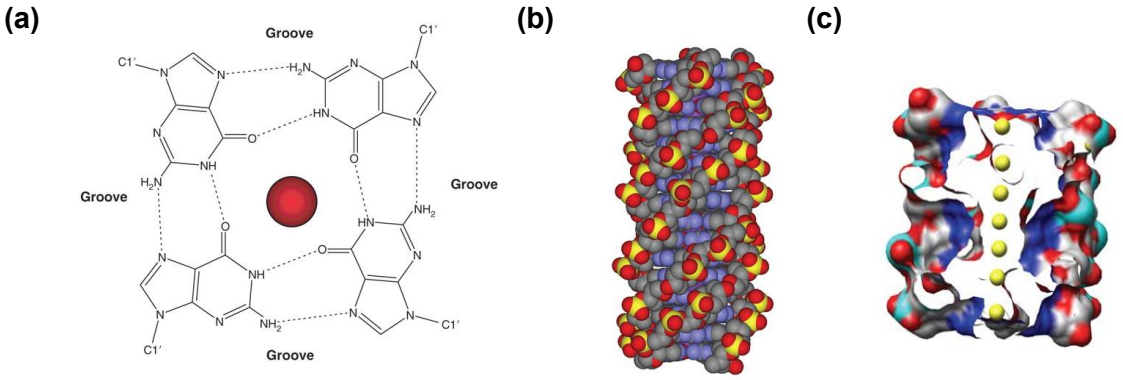


Figure 1.7. G-quadruplexe structures
(a) The arrangement of four guanine bases in a G-quartet. The hydrogen bonds are shown as dotted lines. The red circle represents a metal ion placed in the centre of the quartet. **(b)** Representation of a four-stranded intermolecular G-quadruplex formed by poly(dG). **(c)** Surface view representation of a quadruplex structures comprising eight G-quartets, with the central cavity exposed to show an array of metal ions depicted as yellow spheres. Adapted with permission from (27). Copy right 2006 Oxford University Press.

The formation of G-quadruplexes is strongly driven by the presence of certain monovalent cations such as potassium (K^+) and sodium (Na^+). In addition to the negatively charged phosphate-sugar backbone of the nucleic acids, metal cations can also, and uniquely, interact with G-quadruplexes by occupying the central cavities of the G-quartets, which makes a significant contribution to the stability of G-quadruplexes by neutralizing the electrostatic repulsion between the O6 atoms of the four guanines in a quartet in the case of Na^+, or the O6 atoms of eight guanines by the larger K^+ cation, which typically sits between successive G-quartets (Figure 1.7.c). Since K^+ and Na^+ are two physiologically relevant monovalent ions, most of the cation-GQ studies are focused on their effects on the formation as well as stability of G-quadruplex structures (28). Nonetheless, there are a handful of other monovalent ions that have been shown to play a specific role in stabilizing G-quadruplexes. Early studies in the 1990s suggested that the order of cations that stabilize G-quadruplexes is $K^+ > Rb^+ > Na^+ >> Li^+$. (29,30). The effectiveness of the stabilization by cations is dependent on (a) the ionic radii of the ions, which have a direct influence on the coordination interactions between cations and the O6 carbonyl groups of the guanine residue; and, (b) perhaps more importantly, the desolvation energy of a given cation (the hydration spheres have to be stripped from a Na^+ or K^+ cation consonant with the coordination of the cation by either one or two G-quartet(s))(31). The potassium ion has an ionic radius of 1.33 Å (Table 1.1), which is too large to be coordinated in the plane of a G-quartet. On the other hand, the size of sodium ion is small enough (ionic radius 0.95Å) so that it is able to achieve in-plane

coordination. X-ray crystallographic studies done on the telomeric sequence d($G_4T_4G_4$) from *Oxytricha nova*, which can form antiparallel bimolecular G-quadruplexes, clearly demonstrate the differences in the location of sodium and potassium ions (Figure 1.8). The crystal structures show that five K^+ cations are evenly spaced out from each other, with an average distance of 3.4 Å. Three of the five potassium ions are sandwiched between the quartet layers while the other two are coordinated between the external quartets and the loops. By contrast, only four Na^+ ions are coordinated with the G-quadruplex. While the two internal sodium ions are coordinated perfectly in the planes of the central quartets, the two external ones are slightly off the planes, towards the loops. The difference in the stability of G-quadruplex observed between K^+ and Na^+ can also be explained by the difference in the free energy of cation hydration(32,33). In solution, cations are surrounded by a hydration sphere formed by the solvent (e.g. water), but the quartet-bound cations have to be dehydrated, i.e. to lose their whole hydration sphere, to yield tight M^+-O coordination bonds. While Na^+ gives favourable energy of coordination, its binding is penalized by its strong hydration compared to K^+. Therefore, overfall, K^+ presents the best compromise, and stabilizes G4s more than Na^+ and other monovalent cations. This example highlights the fact that cation binding has a very strong influence on the stability of G-quadruplexes and the resulting structural organization.

Table 1.1. List of the effective ionic radii of monovalent cations interactive with G-quadruplexes.

Monovalent Cations	Effective Ionic Radius (pm)
Li^+	60
Na^+	95
$(NH_4)^+$	148
K^+	133
Tl^+	140
Rb^+	148
Cs^+	169

Adapted with permission from (34). Copyright John Wiley and Sons.

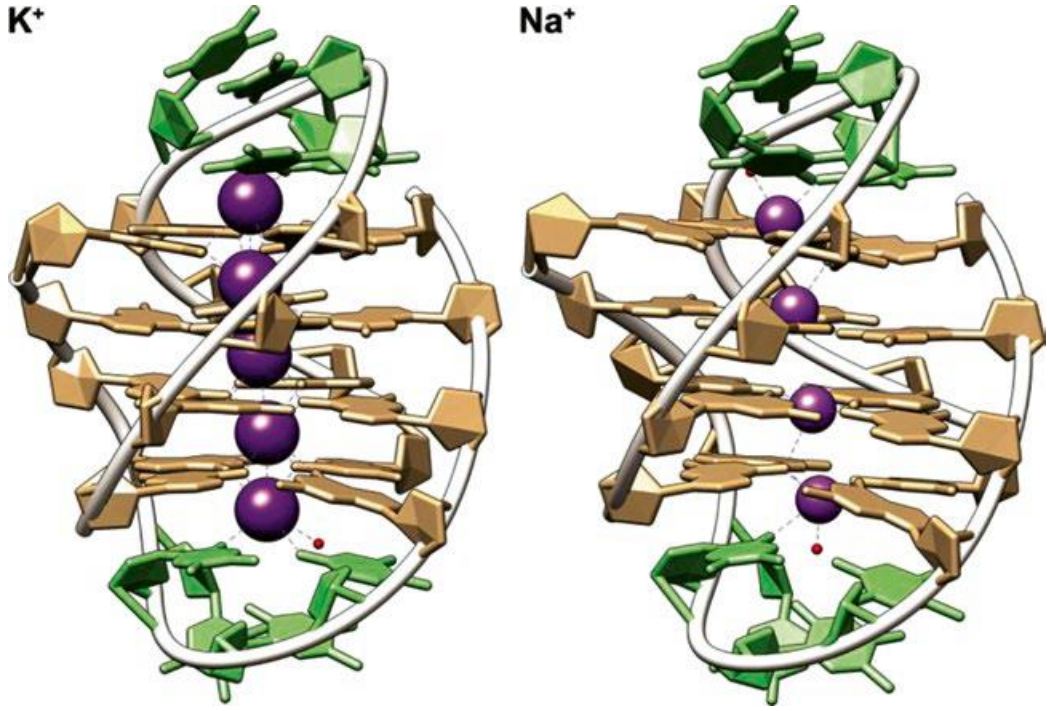


Figure 1.8. Crystal structures of $[d(T_4G_4T_4)]_2$ coordinating K^+ (PDB: 1JPQ) and Na^+ (PDB: 1JB7).

Guanine residues are shown in brown, thymines in green. The cations are portrayed as purple spheres. Five potassium ions are sandwiched between G-quartets whereas four sodium ions are located in plane with the quartets. Adapted with permission from (31). Copyright 2016, Springer International Publishing Switzerland.

One of the key features of G-quadruplexes is its intriguing structural and topological diversity, which is also one of the reasons why GQs have attracted exceptional attention from the nucleic acid research field for decades. G-quadruplex structures can be categorized into various classes according to different parameters, such as the number of strands involved, the number of G-quartets, and the orientation of the strands. Depending on the number of guanine-rich strands involved in forming the structures, which can be one, two, or four, G-quadruplexes can be classified as unimolecular (i.e. intramolecular), bimolecular, or tetramolecular quadruplexes (Figure 1.9). Depending on the orientation of strands relative to each other, a G-quadruplex can also be described as parallel, anti-parallel, or mixed orientation. A parallel G-quadruplex is one in which all the strands are oriented in the same direction. On the other hand, an antiparallel G-quadruplex means that at least one strand is running in the opposite direction relative to the rest. Antiparallel quadruplexes can be further sub-grouped into two types, based on the number of strands that are oppositely oriented. One type can be denoted as ‘aabb’, whereas another can be ‘abab’. The mixed orientation, denoted as

'aaab', is known as (3+1) hybrid, in which case only one strand is running oppositely to others (35). It has been shown by numerous studies that G-quadruplexes with different topologies can exhibit dramatically different biochemical properties. In short, the level of complexity of G-quadruplex structure and topology is beyond comparison to any other DNA secondary structures, which renders G-quadruplex an outstanding material to be exploited in nanotechnologies.

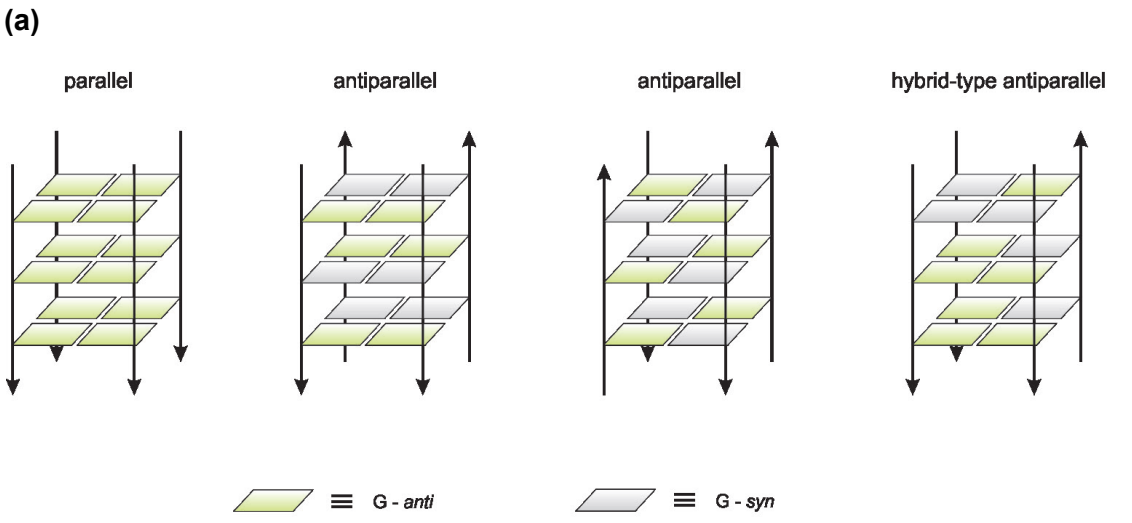


Figure 1.9. Structural and topological diversities of G-quadruplexes.
(a) Strand orientations in parallel, antiparallel and (3+1) hybrid G-quadruplexes. Strand polarities are indicated as the arrows pointing from 5' to 3' end. Adapted from(35). **(b)** Some possible

topologies for tetramolecular (left), biomolecular (middle), and unimolecular (right) G-quadruplexes. Adapted with permission from (27). Copyright 2006, Oxford University Press.

Initially, owing to the lack of proof for its existence *in vivo*, it was speculated that such higher-order DNA structure was biologically irrelevant and remained merely laboratory curiosity observed *in vitro* for a significant period of time. However, as stated by Aaron Klug, a British Nobel prize winner in Chemistry, "if G-quadruplexes form so readily *in vitro*, Nature will have found a way of using them *in vivo*" (36). Thanks to the technological advances in structural and biochemical studies, evidence has started to emerge in the recent literature consistent with occurrence and functioning of G-quadruplexes *in vivo*. Recent years have seen a renaissance in the field of bioinformatics and the development of computational algorithms for the prediction of G4 formation *in vivo*, such as Quadparser (37), AllQuads (38), ImGQfinder (39), and G4Hunter (40). Typically, the consensus motif $G_3N_{1-7}G_3N_{1-7}G_3N_{1-7}G_3$ is used to identify potential G-quadruplexes from primary sequence (41). Numerous genome-wide analyses have revealed that putative G-quadruplex-forming sequences are thoroughly distributed in the genomes of various prokaryotic and eukaryotic organisms, and are preferentially enriched in functional genomic regions, such as telomeric ends and nuclease-hypersensitive promoter regions (42).

With a growing body of computational and experimental evidence pointing to the presence of G-quadruplexes in genomic DNA and RNA, it has now become evident that G-quadruplexes are almost certainly forming under physiological conditions and likely manifest themselves in various critical cellular processes, including but not limited to genome recombination, regulation of gene expression, and cancer cell growth (36). One of the most common functional genomic regions where potential G-quadruplex forming sequences are found is the telomeric regions in chromosomes. Telomeres are highly repetitive and non-coding sequences at the ends of chromosomes that provide protection against gene erosion at cell divisions, chromosomal non-homologous end-joining and nuclease attacks (43). Telomeric DNA in vertebrates consists of tandem repeats of the sequence d(TTAGGG). Human telomeric DNA is typically 5-8 kb long with a 3' single-stranded overhang of approximately 100 to 200 nts (44). Each DNA replication results in a 50 – 200 base loss of the telomere. After reaching a critical shortening of the telomeric DNA, the cell undergoes apoptosis or programmed cell death. The formation of the DNA G-quadruplex in the human telomeric sequence has

been shown to inhibit the activity of telomerase, a cancer-specific reverse transcriptase that is activated in 80 – 90% of tumors (45). Consequently, G-quadruplex has now become an active area of research in cancer biology with wide therapeutical applications.

Considering the wide range of biological processes associated with G-quadruplexes, over the years, much progress has been made towards the direct detection and visualization of G-quadruplex structures in living cells by using small molecules, antibodies or fluorescent probes. In 2020, the Sen lab (46) has showed the use of biotin-tyramide can lead to efficient self-biotinylation of G-quadruplex, which can be utilized for labeling DNA and RNA within live, freshly dissected Drosophila larval salivary glands. More recently, Summers et al. (47) has successfully identified G4s within nuclei of live and fixed mammalian cells by using a fluorescent probe (DAOTA-M2) in conjunction with fluorescence lifetime imaging microscopy (FLIM).

Moreover, it has also received considerable attention in nanotechnology, a newly emerged interdisciplinary area of research at the crossroads of biochemistry, engineering, and medicine, as it offers novel avenues for anti-cancer drug development, such as the design of G-quadruplex-based nanoparticles for targeting cancer cells and delivering drugs (48).

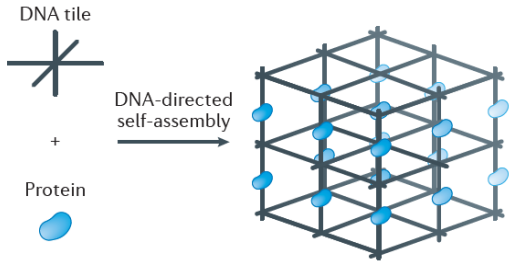
1.5. DNA Nanotechnology

Apart from its central role in biology, DNA has also proved to be a particularly powerful tool in the field of nanotechnology over the past three decades (49–51). DNA possesses a number of properties that render it an excellent candidate for nanostructure building blocks: a) the assembly of DNA is highly precise and controllable owing to its remarkable binding/base-pairing specificity; b) the cost of DNA synthesis is reasonably inexpensive; c) the continuous pi-stacking of the base pairs of DNA renders feasible charge conductivity through the DNA; d) DNA generally shows robust chemical and thermodynamic stability in aqueous solutions and in ionic liquids.

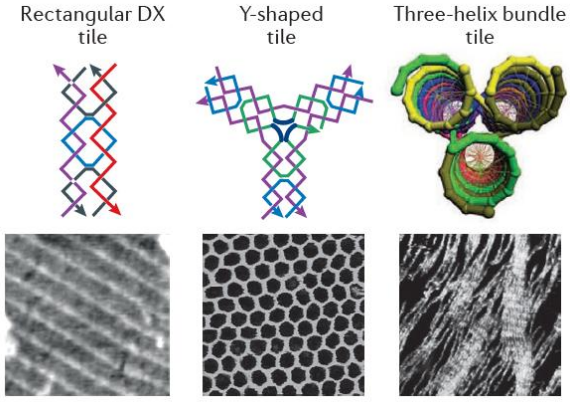
The origins of DNA nanotechnology can be traced back to 1982, when a famous nanotechnologist in the US, Nadrian Seeman, proposed the idea of using immobile DNA junctions to build repeating 3D scaffolds that could be used to organize proteins and

other biomacromolecules in regular three-dimensional arrays (52) (Figure 1.10.a). In 1993, Seeman reported the first construction of double-crossover molecules, comprising of two DNA double helices linked together by the exchange of two strands, which offer geometric rigidity as well as structural stability necessary for building DNA nanostructures with controlled geometry, connectivity, and topology (53,54). It is the early work done by Seeman and other scientists in the late 1990s that laid out the foundation for the growth of a large field, and the construction of many DNA nanostructures. Utilizing the fundamental unit of double-crossover DNA complexes with four sticky ends, which are often called DNA tiles, numerous DNA 2D nanostructures have been constructed and visualized using atomic force microscopy (AFM) in the following decades (Figure 1.10.b). In 2006, another significant breakthrough in the field arose from the construction of 'DNA origami' by Paul Rothermund, which is a method for DNA construction on a finite size scale, whereby long single-stranded DNA is folded into a desired pattern by adding short staple strands that bring together selected parts of the scaffold long DNA strand (55) (Figure 1.10.c).

(a)



(b)



(c)

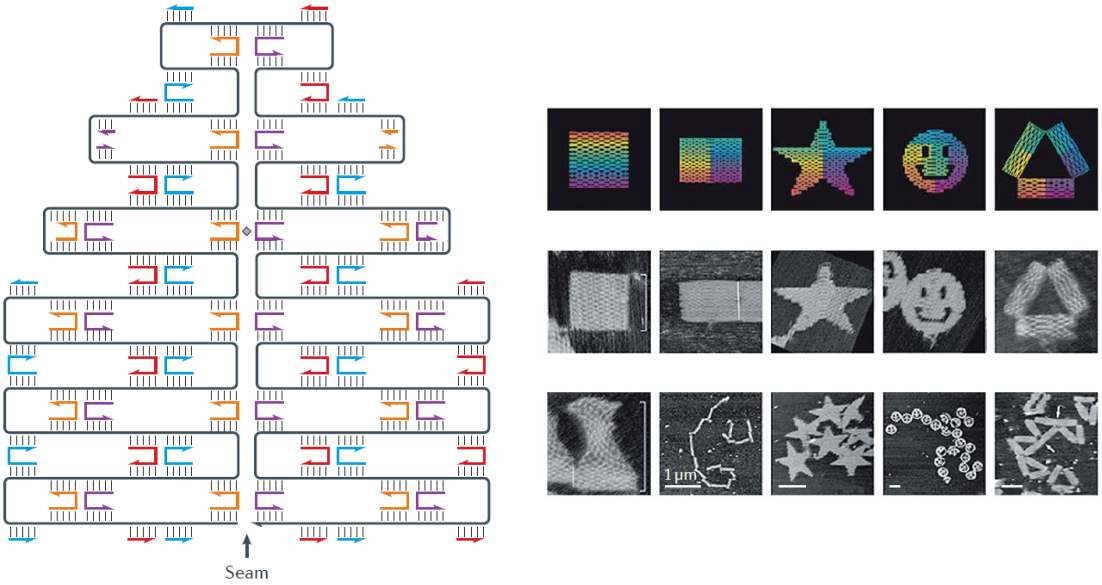


Figure 1.10. DNA nanotechnology.
(a) DNA scaffold as a template for protein crystallization. **(b)** Some DNA tile motifs (top row) and the atomic force microscopy (AFM) images of their assemblies into lattices (bottom row). **(c)** scaffolded DNA origami. The left panel shows a long genomic DNA strand folded with the help of multiple short staple strands to give a desired, computationally designed shape. The right panel shows the AFM images of DNA origami in various shapes with scale bars of 100 nm. Adapted with permission from (56). Copyright 2017, Macmillan Publishers Limited.

Most DNA nanostructures and nanodevices described to date are built with B-DNA helices and involve classical Watson-Crick A-T and G-C base pairs. However, in addition to the iconic double helix, DNAs are known to form other important secondary structures based on alternative pairing schemes, especially the abovementioned triplexes and G-quadruplexes. Since the sequence, structural, and assembly requirements of DNA triplexes and G-quadruplexes are drastically different from their duplex counterparts, they introduce novel resources to the process of engineering with DNA. These alternative helical structures may help to overcome some of the limitations of nanostructures composed solely of W-C base pairs, such as susceptibility to enzymatic degradation; low resistance to heat and to denaturing reagents; relatively high flexibility and deformability; and low sensitivity to chemical stimuli (57).

Taking advantage of the sequence-dictated formation of DNA triplexes, a handful of DNA triplex-based sensing strategies have been proposed in the past decade (19). Many of these approaches are utilizing the pH-sensitive property as well as the strand displacement mechanism of YR*Y triple-helical structures, those containing the CG*C^+ base triple, to construct reconfigurable DNA nanostructures and nanodevices based on

the duplex-to-triplex transition (Figure 1.11). Specifically, the parallel YR*Y triplex has been widely exploited in the design of DNA nanostructures owing to its pH dependence. Rather than being a limitation in application, the pH dependence of the YR*Y motif in fact turns out to be a useful property as it allows fine-tuning the assembly and/or disassembly of the triplex by simply adjusting the solution pH, typically in the range of 5 to 7 (58).

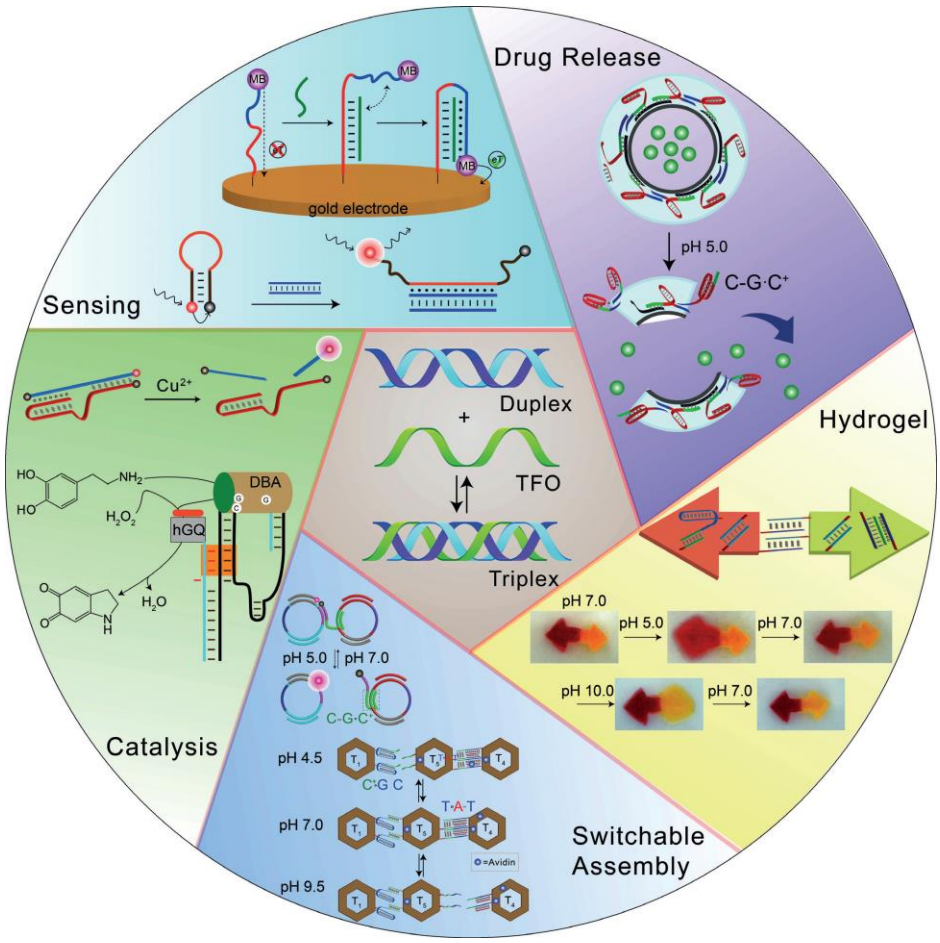


Figure 1.11. DNA triplexes in nanotechnology.
Various DNA triplex-based systems for sensing, controlled drug delivery, switchable catalysis, and the design of shape-memory hydrogels. Adapted with permission from (19). Copyright 2017, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

Last but not least, as another important element in the toolbox for building DNA-based nanomaterials, G-quadruplexes have also been largely exploited in the field toward some remarkable constructions and applications over the years, such as G-wires and DNA origami (Figure 1.12). G-quadruplexes possess a number of unique properties that are particularly useful for nanotechnology, such as cation-sensitivity, structural

polymorphism, high thermal stability, resistance to denaturing conditions, electrical conductivity, and stiffness (57).

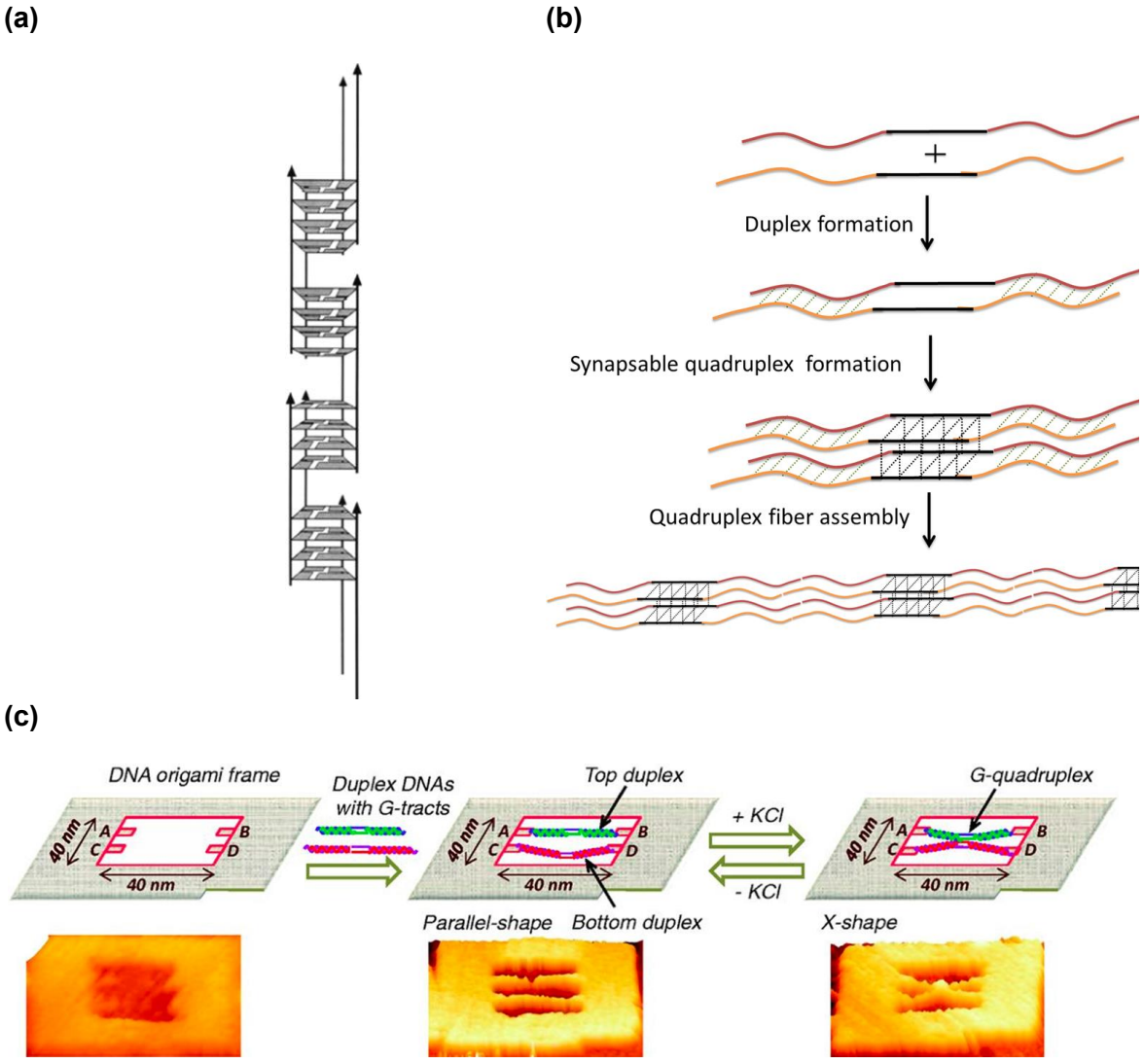


Figure 1.12. DNA G-quadruplexes in nanotechnology.
(a) G-wire assembled with 5'-$G_4T_2G_4$ from *Tetrahymena* telomeric DNA. Adapted with permission from (59). Copyright 1994, American Chemical Society. **(b)** Assembly of synapsable G4-based nanowires. Adapted with permission from(60). Copyright 2014, American Chemical Society. **(c)** Design of a DNA origami frame containing snapsable DNA Adapted from (61)

While recent years have seen a blossoming of nanotechnologies involving unusual nucleic acid structures, most of them are using the classical B-form DNA with the canonical Watson-Crick base pairing scheme; utilization of the non-canonical classes of DNA, on the other hand, still remain relatively rare (60). One should recognize the limitations of the solely B-form DNA-based nanostructures and what advantages the

non-canonical DNA structures can offer to enable greater possibilities in engineering DNA nanostructures as well as in their applications.

Over the decades, the Sen lab has done notable work in designing DNA-based nanostructures and devices such as a “synapsable” duplex(62). In particular, Sen and colleagues introduced a block of guanine mismatches within a Watson-Crick duplex, which permits two such duplexes to “synapse” together at the mismatching site via G-quadruplex formation under specific salt conditions. In 2014, the Sen lab also reported a twisting electronic DNA nanoswitch, with a core of four guanine-rich single strands linking to four DNA double helices. The addition and the removal of K^+ or Sr^{2+} promote alternative conformers of such DNA construct, which shows strikingly distinct electronic properties (63). These simple but also elegantly designed innovations are successful examples of integrating the non-canonical DNA structure into existing nanomaterials, demonstrating how quadruplexes present more possibilities in DNA nanotechnologies.

1.6. Project Overview

More recently, Lat et al. (64) have reported a novel DNA nanostructure, a long and reversibly self-assembling 1-dimensional composite built, for the first time, out of DNA triple helices interspersed with short stretches of G-quadruplexes. Such a “TQ-hybrid” nanostructure requires potassium ions and modestly acidic pH for its formation and can be easily disassembled by changes to either one of these two conditions. The assembly of the monomeric TQ hybrid tiles is achieved by incorporating unique guanine-only sticky ends, which serve as “glue” to connect individual tiles via G-quartet formation.

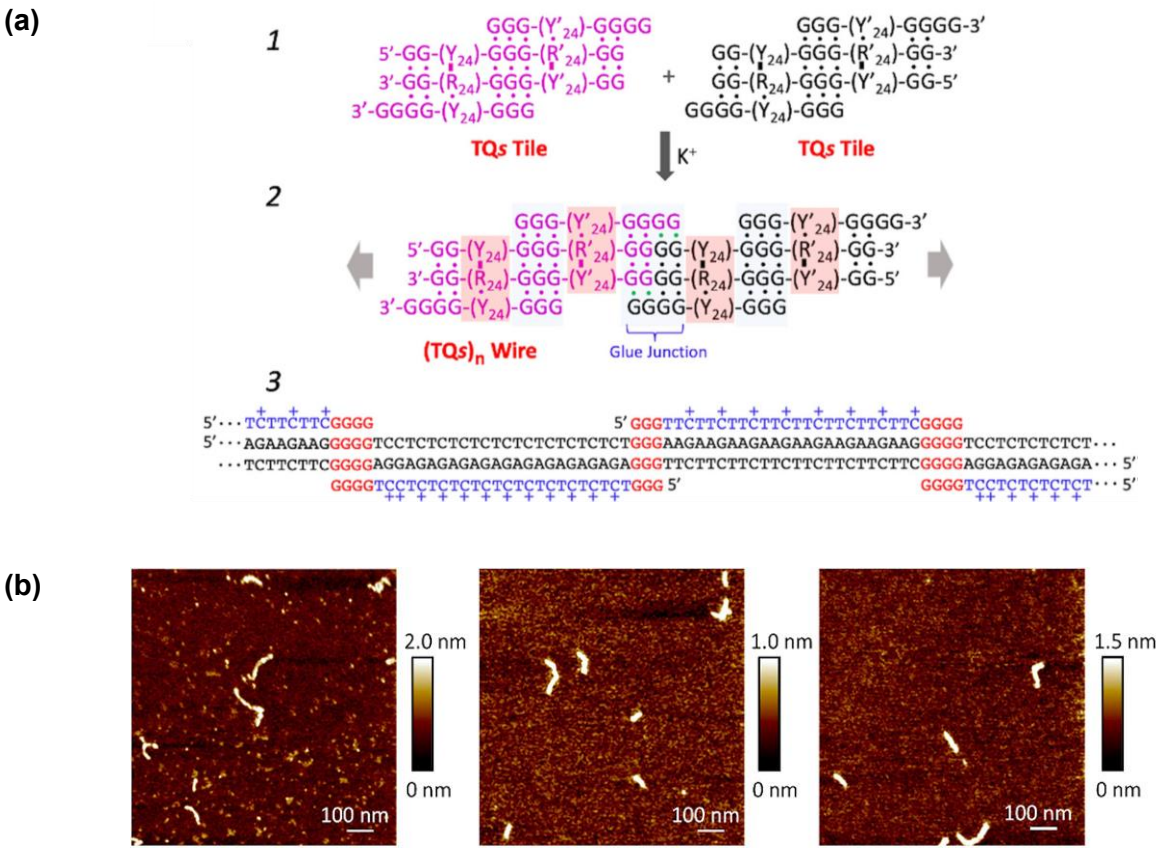


Figure 1.13. 1-dimentional DNA nanostructure built from triplex-quadruplex hybrid composites.

(a) The design of sticky triplex tiles and a triplex-quadruplex hybrid (TQs) 1-dimentional DNA nanostructure. **(b)** AFM images of $(TQs)_n$ nanostructures. Adapted with permission from (64) Copyright 2021, Willey-VCH GmbH.

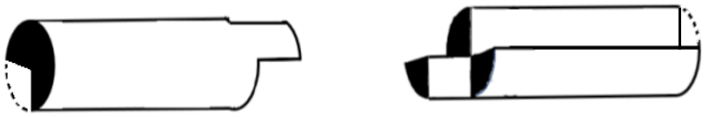
Inspired by this recent work, here we investigated systematically what appears to be a new paradigm for DNA-DNA recognition used in the above TQ nano-assemblies, namely, the use of different arrangements of guanine bases only to enable a kind of "socket-plug" shape complementarity. This study was carried out using triplex-based DNA tiles that have guanine-only sticky ends on one side. I have studied the ways in which, different guanine-rich sockets and plugs are capable of recognizing and binding to each other when initiated by the addition of quadruplex-stabilizing potassium (and, also sodium, rubidium, and strontium) cations. The behaviour of different pairs of TQ hybrid DNA tiles, under various cation and ionic conditions, are explored to define a new set of "self-complementary" and "other-complementary" rules for this unique kind of DNA-DNA recognition and binding.

Chapter 2. Generation of triplex-monomer “tiles” and “socket-plug” triplex-dimers

2.1. Introduction

To carry out a systematic investigation on the DNA-DNA recognition and binding between the triplex-quadruplex hybrids, four different types of guanine-only overhangs were designed and assigned to four different triplex monomer tiles, A, B, C, and D (Figure 2.1). In particular, each triplex tile was designed to have its own unique sequence and size different from the others, to avoid the possibility of third strand exchange between the constituent triplexes within these tiles. Additionally, computational analyses were performed on the sequences to verify that there is neither sequence symmetry nor self-complementarity present in any of the strands, in order to avoid the danger of the individual strands forming undesired secondary structures such as hairpins, with the potential to disrupt the proper triplex formation. All four monomeric tiles contained at least two terminal guanine residues, all on one side of the triplex (on the 3’ side of the purine or “R” strand within any triplex). For tiles A, B, and C, one strand of the three component strands in the triplex contained four overhanging guanines (i.e. two extra terminal Gs protruding beyond the terminal guanines in the remaining strands). The protruding strands were, respectively, the W-C purine (“R”) strand, W-C pyrimidine (“Y”) strand, and the Hoogsteen or third (“hY”) strand in the tiles A, B, and C, while tile D simply contained a blunt end with no protruding Gs in any one of its constituent strands.

a



b

A	B	C	D
5′-(hY_{25})– GGGG	5′-(hY'_{35})– GG	5′-(hY''_{45})– GG	5′-(hY_{25})– GG
5′-(R_{25}) – GG	5′-(R'_{35}) – GGGG	5′-(R''_{45}) – GG	5′-(R_{25}) – GG
3′-(Y_{25}) – GG	3′-(Y'_{35}) – GG	3′-(Y''_{45}) – GGGG	3′-(Y_{25}) – GG

c

5′- TCT TCC TTC TTC TCT CTC TCT CTC T GGGG -3′
5′-AGA AGG AAG AAG AGA GAG AGA GAG A GG -3′
3′- TCT TCC TTC TTC TCT CTC TCT CTC T GG -5′

Triplex A (YR*hY = 25 triples)

5′- CTT CTC TTC TCT TCC TCT TCT CTC TCC TCT CCT CT GG -3′
5′- GAA GAG AAG AGA AGG AGA AGA GAG AGG AGA GGA GA GGGG -3′
3′- CTT CTC TTC TCT TCC TCT TCT CTC TCC TCT CCT CT GG -5′

Triplex B (YR*hY = 35 triples)

5′- TCT CTT CTT CTC TCC TTC CTT CTC TTC CTC TCC TTC TTC CTC CTT GG -3′
5′- AGA GAA GAA GAG AGG AAG GAA GAG AAG GAG AGG AAG AAG GAG GAA GG -3′
3′- TCT CTT CTT CTC TCC TTC CTT CTC TTC CTC TCC TTC TTC CTC CTT GGGG-5′

Triplex C (YR*hY = 45 triples)

5′- TCT TCC TTC TTC TCT CTC TCT CTC T GG-3′
5′-AGA AGG AAG AAG AGA GAG AGA GAG A GG-3′
3′- TCT TCC TTC TTC TCT CTC TCT CTC T GG-5′

Triplex D (YR*hY = 25 triples)

Figure 2.1. Design of four types of monomeric DNA triplex tiles.
(a) Schematic drawings demonstrating the dimerization between two triplex tiles. The dashed lines indicate spatial vacancy within the triplex regions**. (b)** Designs of 4 types of guanine-only sticky ends. R presents the purine-rich component strand in the triplex; Y presents the pyrimidine strand that interacts with the purine strand via Watson-Crick base pairing scheme; hY presents the other pyrimidine strand that interacts with the purine strand via Hoogsteen base pairing. **(c)** The actual sequences of each triplex tile, with the W-C duplex in black, the purine stand highlighted in bold, and hoogsteen strands coloured in blue.

2.2. Materials and methods

2.2.1. DNA oligonucleotides and purification

All DNA oligonucleotides were purchased from Integrated DNA Technologies, Inc. (Coralville, Iowa, United States). The sequences of all the nucleotides used in the project are listed in Table 2.1. The oligonucleotides were purified using 8% denaturing polyacrylamide gels. Visualized by UV shadowing, the desired DNA bands with the correct sizes were excised from gel, eluted overnight into 1X TE buffer (10 mM Tris, 0.1 mM EDTA, pH 7.4), and recovered by ethanol precipitation using 3 M NaOAc. After washing with 70% cold ethanol and air-drying, the DNA pellets were suspended in 1X TE buffer to make up the desired stock concentration.

For native gel analysis and DMS protection assay, oligonucleotides were first treated with freshly prepared 10% (v/v) piperidine at 90℃ for 30 minutes, to eliminate oligonucleotide molecules with lesions formed during chemical synthesis, followed by 5' end radiolabeling with ^{32}P using γ-^{32}P ATP purchased from PerkinElmer Health Sciences Canada, Inc. (Woodbridge, Ontario, Canada) using a standard kinasing procedure. The radiolabeled DNAs were purified again using 8% denaturing polyacrylamide gels.

Table 2.1. List of oligonucleotide sequences used in the assembly of various DNA constructs.

Oligo Name	Sequence (5' → 3')
ssR(A)	AGA AGG AAG AAG AGA GAG AGA GAG AGG
ssY(A)	GGT CTC TCT CTC TCT CTT CTT CCT TCT
ssHY(A)	TCT TCC TTC TTC TCT CTC TCT CTC TGG GG
ssR(B)	GAA GAG AAG AGA AGG AGA AGA GAG AGG AGA GGA GAG GGG
ssY(B)	GGT CTC CTC TCC TCT CTC TTC TCC TTC TCT TCT CTT C
ssHY(B)	CTT CTC TTC TCT TCC TCT TCT CTC TCC TCT CCT CTG G
ssR(C)	AGA GAA GAA GAG AGG AAG GAA GAG AAG GAG AGG AAG AAG GAG GAA GG

ssY(C)	GGG GTT CCT CCT TCT TCC TCT CCT TCT CTT CCT TCC TCT CTT CTT CTC T
ssHY(C)	TCT CTT CTT CTC TCC TTC CTT CTC TTC CTC TCC TTC TTC CTC CTT GG
ssR(D)	AGA AGG AAG AAG AGA GAG AGA GAG AGG
ssY(D)	GGT CTC TCT CTC TCT CTT CTT CCT TCT
ssHY(D)	TCT TCC TTC TTC TCT CTC TCT CTC TGG
Alkyne_ssY(A)	GGT C/i5OctdU/C TCT CTC TCT CTT CTT CCT TCT
Alkyne_ssY(C)	GGG GTT CC/i5OctdU/ CCT TCT TCC TCT CCT TCT CTT CCT TCC TCT CTT CTT CTC T
ssR(AA)	GGT CTC TCT TCT CCT TCT CTT CTC TTC AGA AGG AAG AAG AGA GAG AGA GAG AGG
ssY(AA)	GGT CTC TCT CTC TCT CTT CTT CCT TCT GAA GAG AAG AGA AGG AGA AGA GAG AGG
ssHY1(AA)	CTT CTC TTC TCT TCC TCT TCT CTC TGG GG
ssHY2(AA)	TCT TCC TTC TTC TCT CTC TCT CTC TGG GG
ssR(211_A)	AGA AGG AAG AAG AGA GAG AGA GAG AG
ssY(211_A)	GT CTC TCT CTC TCT CTT CTT CCT TCT
ssHY(211_A)	TCT TCC TTC TTC TCT CTC TCT CTC TGG
ssR(211_C)	AGA GAA GAA GAG AGG AAG GAA GAG AAG GAG AGG AAG AAG GAG GAA G
ssY(211_C)	G GTT CCT CCT TCT TCC TCT CCT TCT CTT CCT TCC TCT CTT CTT CTC T
ssHY(211_C)	TCT CTT CTT CTC TCC TTC CTT CTC TTC CTC TCC TTC TTC CTC CTT G

ssR indicates a single-stranded purine-rich strand of DNA; ssY indicates a single-stranded pyrimidine-rich strand of DNA; ssHY indicates a pyrimidine-rich Hoogsteen strand of DNA that is involved in triplex formation.

2.2.2. Assembly of DNA monomeric triplexes and triplex-dimers

Monomeric DNA triplex monomers were assembled by mixing respective component strands (each at 1 μM concentration) in the presence of 10 mM Li^+ acetate and 3 mM Mg^{2+} acetate at pH 5.2. The DNA mixtures were heat-denatured at 100℃ for 4 minutes and gradually cooled down from 100℃ to 20℃ in a thermocycler using a constant cooling rate of 7.5℃/min.

For formation of triplex-dimers, the pre-assembled constituent monomeric triplexes (each at 1 μM concentration) were mixed at equal volume, followed by the addition of final concentrations of either 50 mM K^+ acetate or Na^+ acetate (pH 5.2) to initiate the dimerization. The DNA solutions were allowed to rest at 22℃ for 18 hours to complete dimer formation.

2.2.3. Native gel analysis

The formation of various DNA complexes (duplexes; monomeric triplexes, and triplex-dimers) was investigated using 7.5% (29:1 acrylamide:bis) native polyacrylamide gels (containing and run in 50 mM Tris acetate and of one of 5 mM K^+, Na^+, or Li^+ acetate at pH 5.2). Prior to loading on gels, the DNA mixtures were mixed with loading buffer (50 mM Tris acetate, 30% glycerol, bromophenol blue and xylene cyanol at pH 5.2). The gels were run at 100 V at 22℃ with efficient cooling. Gels were exposed at 4℃ and scanned using a Typhoon 9410 Phosphorimager (Amersham Biosciences).

2.3. Generation of monomeric DNA triplex tiles

To establish a standard ionic condition in which monomeric DNA triplex tiles can stably form but not triplex-dimers, a lithium concentration titration, ranging from 0 to 20 mM, was performed with triplex A and C, respectively. To generate the triplex monomers, the three component strands were mixed at equal concentration (1μM each) in a reaction tube prior to the addition of LiOAc (pH 5.2) at various concentrations. After 18 hours of incubation at 22℃, the DNA samples were run on a 7.5% native gel containing 50 mM Tris acetate at pH 5.2. The results in Figure 2.2.a-b demonstrate that even at the highest tested lithium concentration (20 mM), more than 50% of DNA remained as duplex rather than triplex after 18 hours of incubation, suggesting that the

presence of Li^+ alone may not be sufficient and that additional ions might be required to further stabilize the monomeric triplexes in order to obtain 100% triplex monomer formation (but not triplex-dimer). Therefore, a magnesium titration in the presence of 10 mM LiOAc was carried out. Figure 2.2.c-d show that at the addition of Mg^{2+} can help stabilize triplex monomers and importantly, at 10 mM LiOAc plus 3 mM $Mg(OAc)_2$, both A and C were able to reach 100% triplex monomer formation without forming any triplex-dimers. Hence, this was chosen to be the standard triplex monomer forming condition for all subsequent experiments.

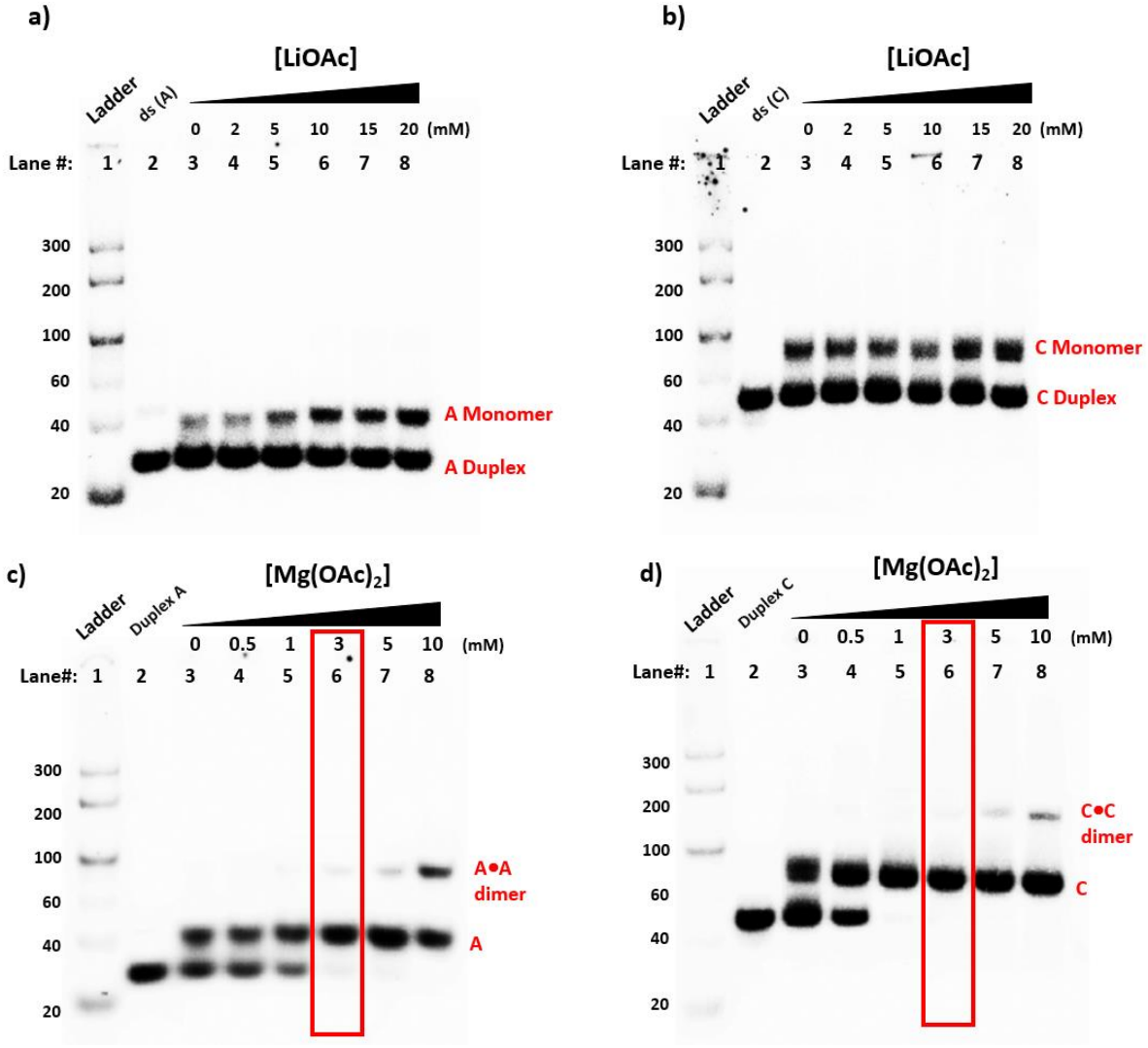


Figure 2.2. Establishing a standard solution for monomeric triplex tiles.
(a, b) [LiOAc] titration for triplex A and C. The constituent strands of a triplex were mixed at 1μM in the presence of LiOAc at various concentrations (0, 2, 5, 10, 15, and 20 mM). The DNAs were allowed to rest at 22 ℃ for 18 hours before running on a 7.5 native gel containing 50 mM Tris acetate at pH 5.2. **(c, d)** [$Mg(OAc)_2$] titration in the presence of 10 mM LiOAc at pH 5.2. The

constituent strands of a triplex were mixed at 1μM in the presence of 10 mM LiOAc plus $Mg(OAc)_2$ various concentrations (0, 0.5, 1, 3, 5, and 10 mM). The DNAs were allowed to rest at 22 ℃ for 18 hours before running on a 7.5 native gel containing 50 mM Tris acetate and 5 mM LiOAc at pH 5.2.

After establishing the standard ion conditions for triplex monomer formation, all four triplex tiles were assembled by mixing the three component strands in the presence of 10 mM Li^+ acetate (LiOAc) and 3 mM Mg^{2+} acetate ($Mg(OAc)_2$) at pH 5.2. Figure 2.3 shows the formation of these monomeric DNA triplex tiles when run on a 7.5% native polyacrylamide gel at pH 5.2 containing and running 50 mM Tris acetate and 5 mM LiOAc. It can be seen that in each case, the addition of Hoogsteen strand to the pre-assembled W-C duplex results in the formation of a triplex, as indicated by an upward shift of the band on the gel compared to the corresponding duplex.

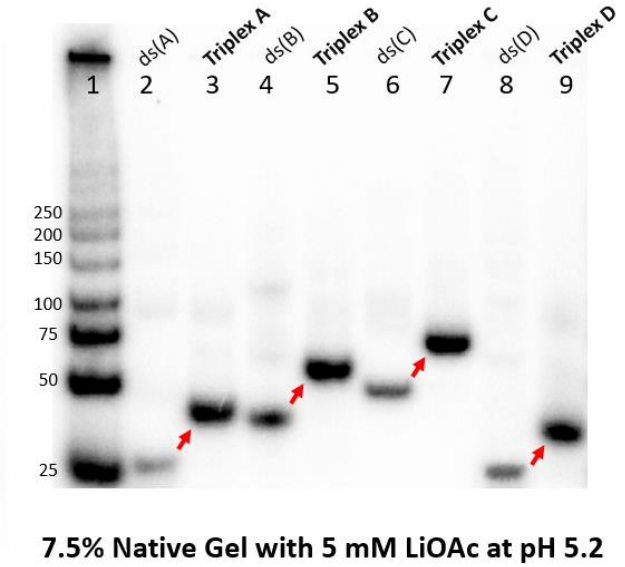


Figure 2.3. Generation of monomeric triplex tiles.
To analyze triplex tile formation, a 7.5% native is run in 50 mM Tris acetate buffer plus 5 mM lithium acetate at pH 5.2. Prior to loading on gels, samples were incubated overnight (18 hours) at room temperature (22℃) in the presence of 10 mM LiOAc and 3 mM $Mg(OAc)_2$, pH 5.2.

2.3.1. pH dependence of monomeric triplex tile formation

To examine the effect of pH on the formation of triplex tiles, the constituent strands for all four triplexes were also assembled individually, in the presence of 10 mM LiCl and 3 mM $MgCl_2$ at *neutral* pH (pH 7.5), followed by running in a native gel containing and running in 50 mM Tris acetate plus 5 mM LiOAc, pH 7.5. In this case, no

gel shift corresponding to triplex formation from duplex was observed (Fig. 2.4). This result indicates that, as expected, a mildly acidic pH environment is required for the triplex formation. At pH 7.5, the cytosine protonation for the required CG*C^+ base triples does not occur sufficiently to enable stable triplexes to form. To examine whether acidic pH was also required for pre-assembled triplex structures to remain intact, triplexes were first assembled at pH 5.2, and then run on a native gel at pH 7.5. Likewise, in this case, no triplex structures were observed, indicating that the acidic environment was required not only for the formation but also the stability of the triplex tiles. We noted that such an absolute requirement for lower than neutral pH could prove to be a powerful tool for reversibly controlling the assembly and disassembly of the triplex structures, and any assemblies built from the triplexes, simply by adjusting the pH.

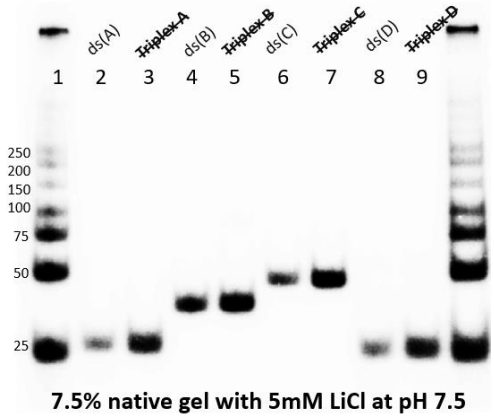


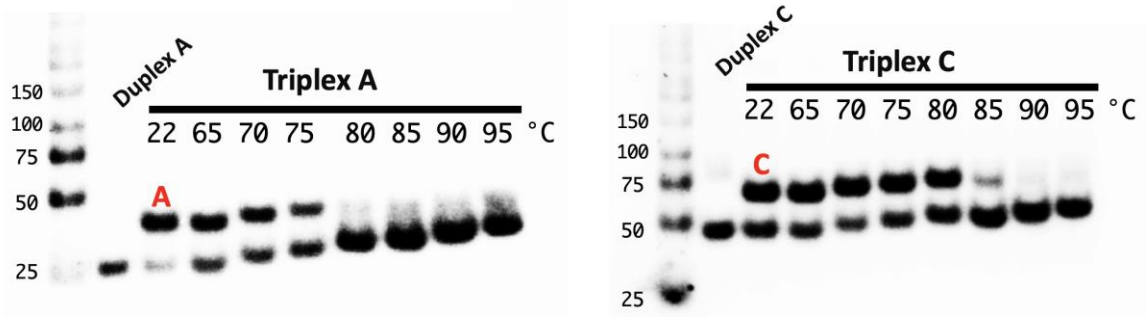
Figure 2.4. Neutral pH prevents triplex formation.
To analyze the effect of pH on triplex formation, a 7.5% native is run in 50 mM Tris borate buffer plus 5 mM lithium chloride at pH 7.5. Prior to loading on gels, samples were incubated overnight (18 hours) at room temperature (22℃) in the presence of 10 mM LiCl and 3 mM $MgCl_2$, pH 7.5.

2.3.2. Stabilities of monomeric triplex tiles

To gain insight into the stabilities of the monomeric triplex tiles, melting experiments were performed by heating pre-assembled triplex tiles to various temperatures in the presence of 50 mM Li^+ (pH 5.2) to obtain estimates of their melting temperatures (T_m). Figure 2.5.a presents the melting behaviours of triplex A and of triplex C, separately, on a 5 mM Li^+-containing native gel (pH 5.2) after being heated at different temperatures ranging from 20 to 95 ℃. The intensities of bands corresponding to triplex monomers were plotted as a function of temperature to obtain normalized melting curves for triplex A and triplex C, respectively (Figure 2.5.b). The results suggest that 50% of triplex A melts roughly at 72.5 ℃, whereas triplex C approximately at 80 ℃.

The melting temperature of triplex C is higher than that of triplex A most likely because each component strand within triplex C has 45 nts (excluding the guanine overhangs), which is 20 nts longer compared to that of triplex A. The long triplex is expected to have a higher melting temperature.

(a)



(b)

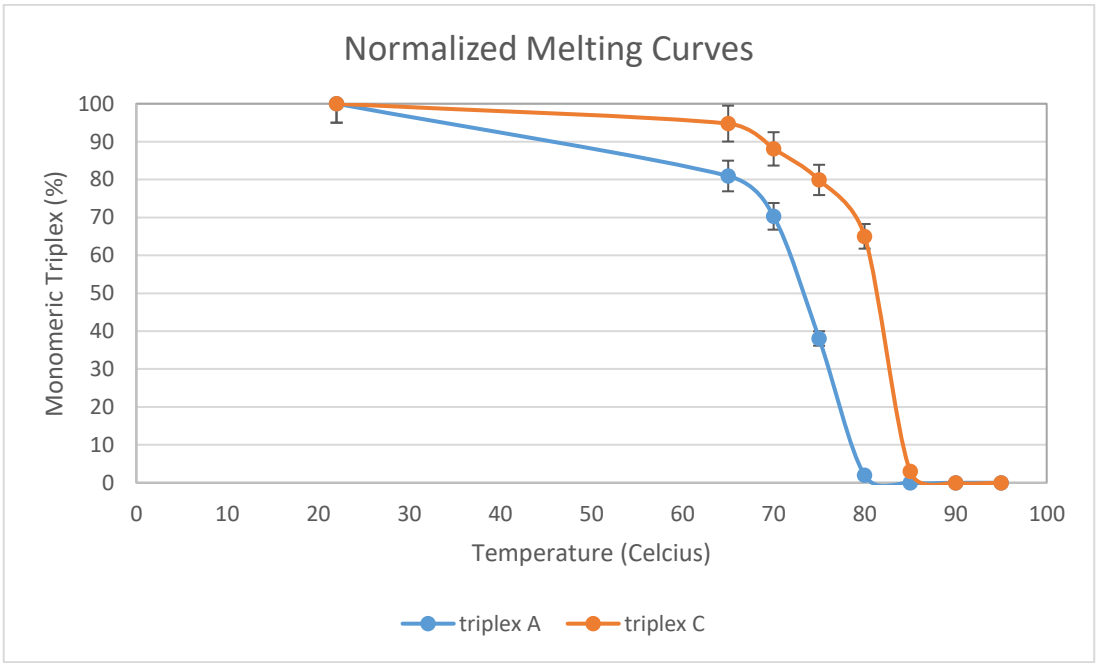


Figure 2.5. Melting behaviours of monomeric triplex tiles A and C in 50 mM Li^+ pH 5.2.

(a) To determine the melting temperatures, pre-assembled triplexes are heated at various temperatures for 4 minutes in a thermocycler, after which were immediately loaded on to a 7.5% native gel with 5 mM LiOAc at pH 5.2. **(b)** Densitometry analyses were carried out to quantify triplex monomer formation using the ImageQuant TL 8.1 software (Cytiva). The percentages of monomeric triplex were calculated as the total intensity of the band corresponding to the monomer divided by the total intensity of the entire lane. The melting temperatures of each monomeric triplex was derived from estimating the temperatures at which 50% of the triplex is melted.

2.4. Generation of "socket-plug" triplex-dimers via G-quadruplex formation

2.4.1. Homo-dimerization: "self-complementary" recognition and binding

After the successful assembly of the triplex tiles, we investigated the possibility of homo-dimerization of the individual triplex tiles, A, B, C, and D, in the presence of K^+ ion, and, separately, Na^+ ion. As shown by Lat *et al.* (54), the addition of K^+ can initiate the formation of G-quadruplexes from unpaired guanines at the ends of the triplex tiles. We were interested in investigating whether tiles A, B, and C (with tile D as a negative control) were comparably able to dimerize to form triplex-homodimers. To carry out this experiment, monomeric triplex tiles A, B, C, and D (dissolved at 1 μM concentration in 10 mM LiOAc plus 3 mM $Mg(OAc)_2$, pH 5.2, in independent tubes) were made up to 50 mM KOAc (still maintaining pH 5.2) each, and allowed to rest overnight at 22 ℃, prior to analysis. Figure 2.6.a shows the pre-assembled monomeric triplexes prior to the addition of K^+ (the native gel in Figure 2.6.a was run in 50 mM Tris acetate plus 5 mM LiOAc, pH 5.2). Following incubation in 50mM KOAc, pH 5.2, all four DNA incubations were run on a native gel containing and run in 50 mM Tris acetate plus 5 mM KOAc at pH 5.2. New sets of products with retarded mobility (with reference to the duplex size ladder run in both Figures 2.6.a and 2.6.b) could be seen in Figure 2.6.b, which were not seen in Figure 2.6.a. Each new product had approximately twice the molecular weight of its corresponding triplex monomer, consistent with their being triplex-dimers of A, B, and C, respectively (i.e. "A•A", "B•B", and "C•C"). The triplex D, containing blunt-ended guanines, did not show any corresponding mobility-retarded band, consistent with triplex D being unable to dimerize under these conditions.

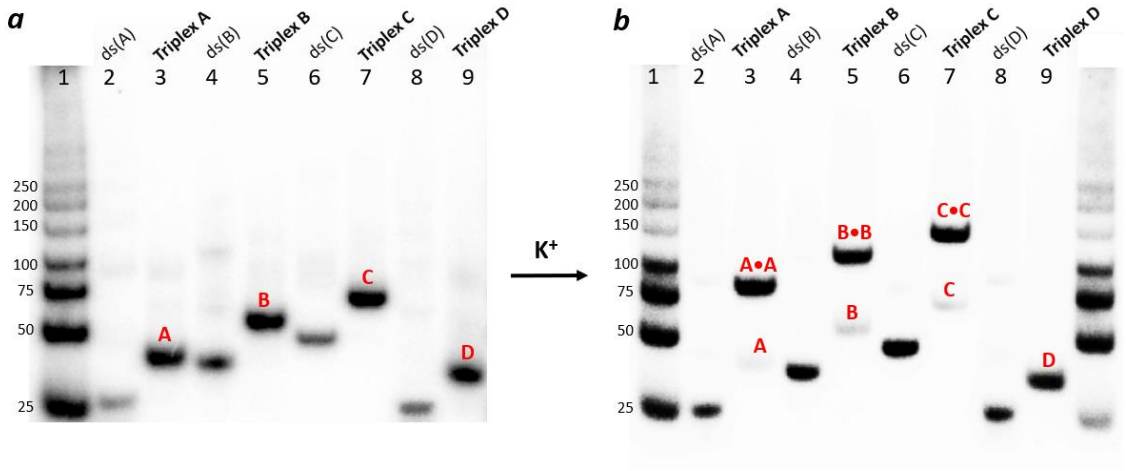


Figure 2.6 Addition of K^+ gives rise to the formation of various triplex-dimers. **(a)** Four types of triplex tiles (A, B, C, and D) were first assembled in 10 mM Li^+ and 3 mM Mg^{2+} at pH 5.2 and run on a 7.5% native gel containing 5 mM Li^+ at pH 5.2. **(b)** Following their assembly, each triplex tile was treated with 50 mM KOAc and analyzed on a 7.5% native gel containing 5 mM K^+ at pH 5.2.

To test for the possibility of homo-dimerization in the presence of Na^+, a less effective G-quadruplex stabilizing metal ion than K^+, the abovementioned experiment was carried out in the same manner, except that the pre-assembled monomeric triplex tiles A, B, C, and D (1 μM DNA concentration in 10 mM LiOAc plus 3 mM $Mg(OAc)_2$, pH 5.2), were incubated, separately, in 50 mM NaOAc, pH 5.2, at 22℃ overnight prior to analysis by native gel electrophoresis. The DNAs were run on a 7.5% native gel containing 50 mM Tris acetate plus 5 mM NaOAc at pH 5.2 (Figure 2.7). Similar to what we have seen when the DNAs were incubated in K^+ (Figure 2.6.b), the incubation in Na^+ also gave rise to three mobility-retarded products, which we expect to be triplex-dimers of A, B, and C, respectively (i.e. “A•A”, “B•B”, and “C•C”). Once again, the absence of such a product within triplex D indicates that blunt-ended triplex monomer was unable to dimerize in Na^+. We noticed that, under the same cation concentration (50 mM), K^+ was able to drive the dimerization of triplex almost to completion, with little triplex monomers left unreacted, whereas in Na^+, a considerable portion of triplex tiles still remained as monomers even after overnight incubation. This observation can be explained by the fact that K^+ is more effective in stabilizing G-quadruplex than Na^+ (28).

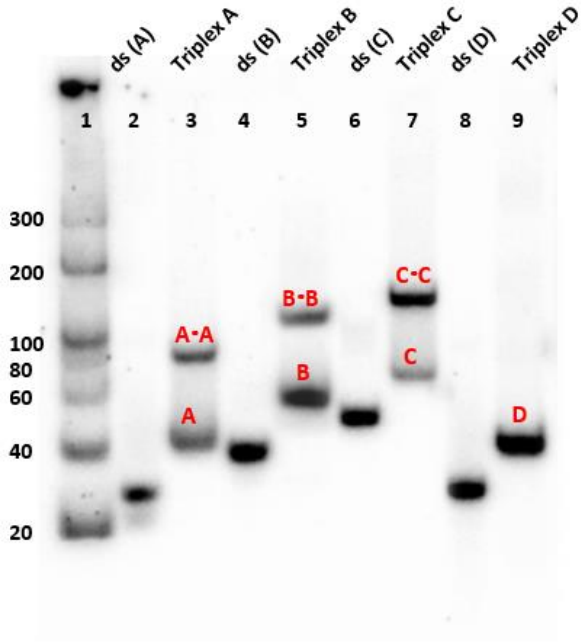


Figure 2.7. Homo-dimerization of triplexes in the presence of 50 mM NaOAc. Four types of monomeric triplex tiles (A, B, C, and D) pre-assembled in 10 mM Li^+ and 3 mM Mg^{2+} at pH 5.2 were incubated in 50 mM NaOAc overnight before running on a native gel containing 50 mM Tris acetate plus 5 mM Na^+ at pH 5.2.

To ensure that the triplex-dimers are formed due to the presence of K^+ or Na^+ added *during* the incubation carried out in solution, rather than during the subsequent gel electrophoresis carried out with low mM of the relevant salt present in the native gel solution and in the running buffer, a control experiment was performed in which the triplex tiles pre-assembled in 10 mM LiOAc and 3 mM $Mg(OAc)_2$ were incubated at 22 ℃ for either 0 or 1 hour, without the addition of either K^+ or Na^+, and loaded onto three native gels containing 50 mM Tris acetate (pH 5.2) plus one of 5 mM Li^+, K^+, or Na^+. The gels in Figure 2.8 shows that both triplex A and C remained as monomers in all the gels, suggesting that the formation of triplex-dimers was a result of the metal cations being present during solution incubation, rather than exposure to the cations present in the gel. In the 5 mM K^+ gel, a trace amount of triplex-dimer A was formed, most likely because K^+ is a very strong G-quadruplex stabilizer. Yet the overall dimer formation was nevertheless very marginal.

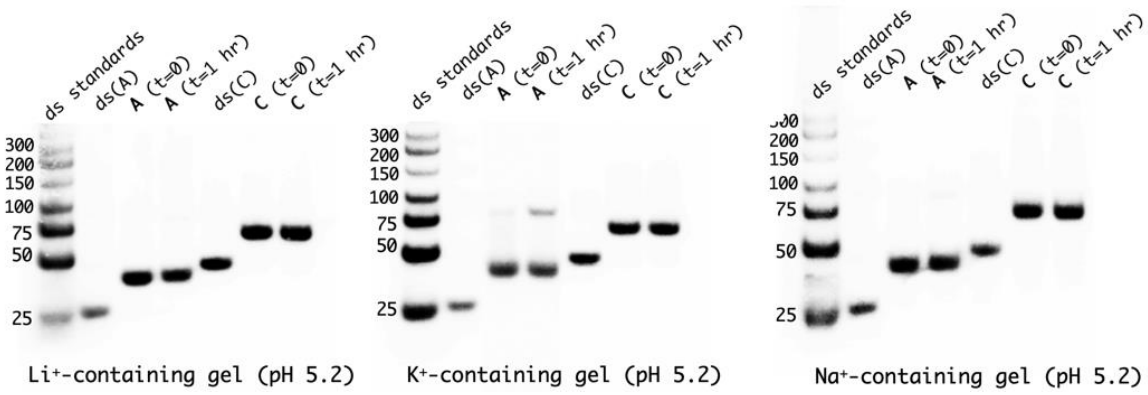


Figure 2.8. The presence of salt in the gel solution and running buffer have little contribution to the overall triplex-dimer formation.

Figure 2.9 shows schematic representations of A•A, B•B, and C•C triplex-dimers. In all three cases, a pair of monomeric tiles (for instance, A) can be arranged in a certain way such that their G-rich sticky ends are perfectly complementary to each other, in a manner of a “socket” and its shape-complementary “plug”, enabling the formation of a complete 4-layer G-quadruplex. The presence of either K^+ or Na^+ ion stabilizes these central G-quartets (the ones formed compositely from the two participating tiles), enabling them to serve as a “glue” to join together, end-to-end, the two participating tiles. Interestingly, all three tiles, A, B, and C, with differing protruding strand in each case, are capable of forming their respective triplex homodimer. In sharp contrast, the blunt-ended triplex D is unable to dimerize owing to its lack of a protruding strand with guanines capable of filling the spatial vacancy in its counterpart, to form a G-quadruplex.

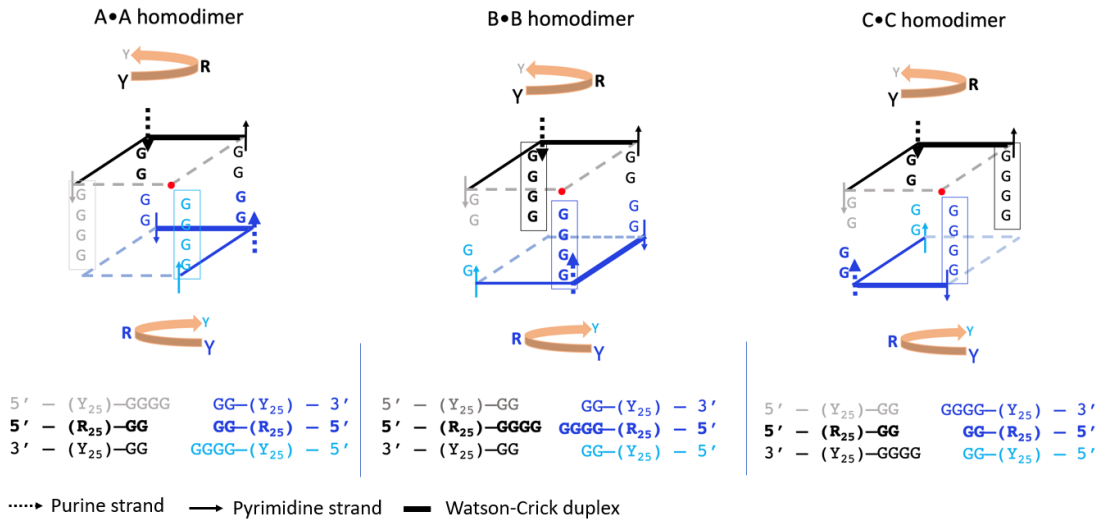
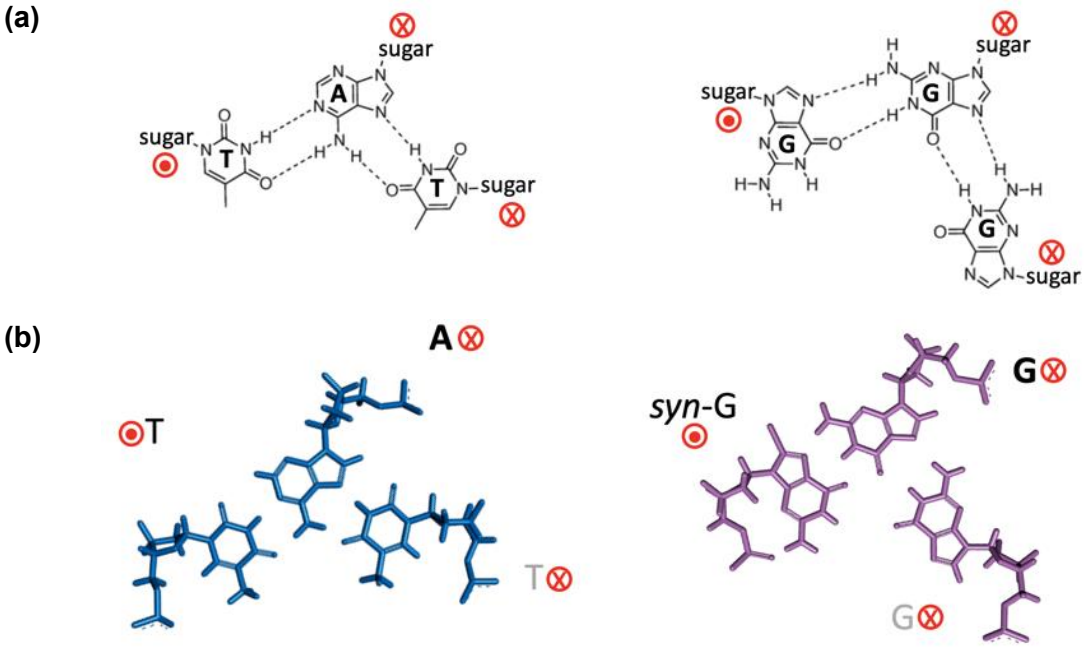


Figure 2.9. Schematic representations of A·A, B·B, and C·C triplex-dimers.
In each drawing, the Watson-Crick purine strand is represented by a thick, dashed line, and the terminal guanines are highlighted in bold. The pyrimidine-rich Hoogsteen strand is represented by a solid line in a lighter colour The overhanging strand, containing four terminal guanines, in each case is shown as boxed, for easy identification. The red dot in each figure represents the vertex closest to the eye of the viewer.

To get some perspective on the geometric transition of the terminus of the YR*Y triplex-helix into a G-quadruplex, we looked at NMR solution structure of a regular TA*T triple-base and compared it to that of a G-triple structural motif that has been observed during the folding process of the G-quadruplex aptamer TBA (65,66). The geometric arrangement of a G-triple is very similar to that of a regular G-quartet, except for missing one guanine base. Figure 2.10.b shows the solution NMR images of a TA*T base triple and a G-triple. By superimposing the NMR images, one can see that while the Watson-Crick duplex components in both structural motifs remain more or less static, while the backbone of the triplex-forming strand in the G-triplex is slightly twisted to a certain extent in comparison to the regular triple, as highlighted by the red circle. We believe that such a modest conformational change is required for the individual triplex tiles to dimerize via G-quadruplex formation.



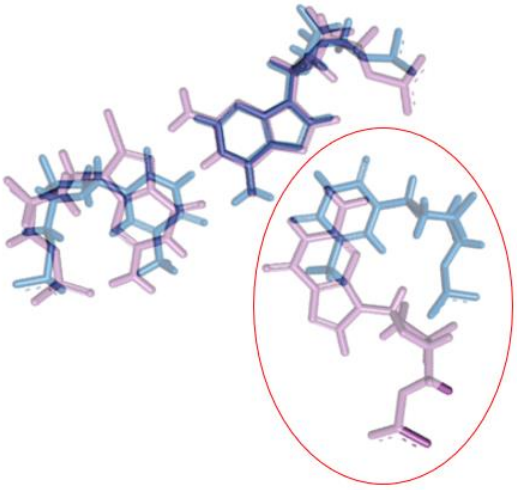


Figure 2.10. Schematic illustrations and NMR structures of a TA*T base triple and a G-triple.
(a) Stick illustration of a regular TA*T base triple and a G-triple. For the TA*T triple, the TA represents the standard Watson-Crick duplex structure, whereas the asterisk sign indicates the base pairing with the Hoogsteen strand that lies in the major groove of the duplex. Adapted with permission from (55). **(b)** solution NMR structures for TA*T (NDB: 1AT4) and G-triple (NDB: 2MBJ).

2.4.2. Hetero-dimerization: other-complementary recognition and binding

To explore the possibility of cross-hybridization between the distinct triplex tiles, A, B, and C, two different kinds of monomeric tiles were mixed together in the presence of varying concentrations of K^+ or Na^+, ranging from 10 – 90 mM. Since all monomeric triplex tiles were designed to be of different lengths, their corresponding homodimers had distinct and characteristic gel mobilities (Figure 2.7) based, ultimately, on the size of the relevant triplex monomer. The formation of a putative heterodimers between two distinct monomeric triplex tiles could therefore be identified by the appearance of a new gel band with intermediate mobility relative to the homodimers formed by the two relevant monomeric triplexes.

Triplexes A and B were found to be able to dimerize to form A•B heterodimer in the presence of either K^+ or Na^+ (Figure 2.11.a-b). However, the formation of the A•B heterodimer was much more significant in K^+ rather than in Na^+, most likely owing to the fact that K^+ is a better G-quadruplex stabilizing agent than Na^+. Triplexes B and C

formed significant heterodimer only in the presence of K^+, but not Na^+ (Figure 2.11.c-d). Most notably, the third pair of triplex tiles, triplexes A and C, showed wholly distinct outcomes in the presence of K^+ versus in Na^+(Figure 2.11.e-f). In the presence of K^+, triplexes A and C formed overwhelmingly the A•C heterodimers, and not the homodimers, A•A and C•C (densitometry shows the yield of A•C at 50 mM KOAc to be more than 6-fold higher than the production of either A•A or C•C). By striking contrast, incubation in Na^+, only the A•A and C•C homodimers were observed, in approximately equal measure, but no trace of the A•C heterodimer could be seen.

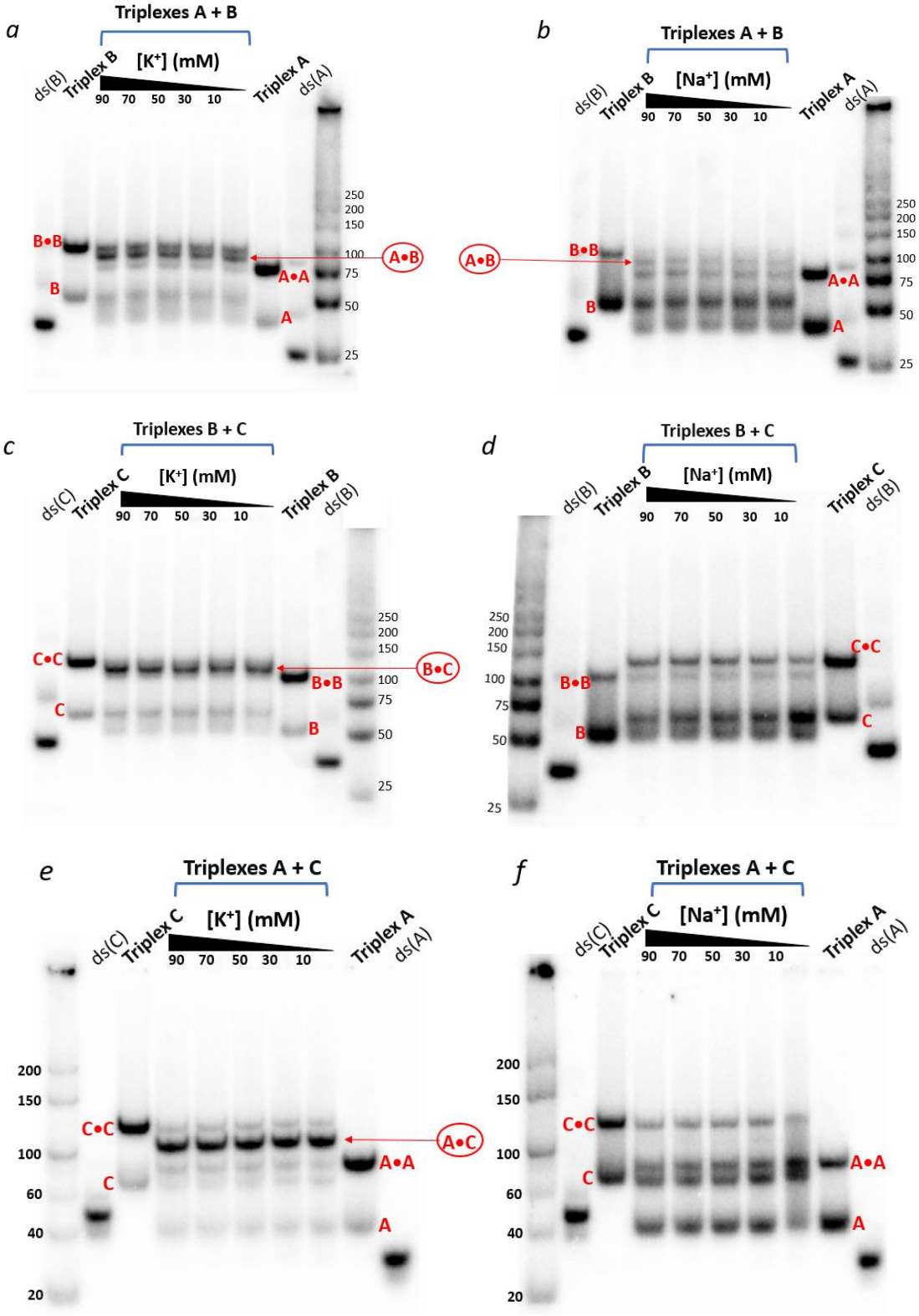


Figure 2.11. Generation of triplex heterodimers in the presence of K^+ versus Na^+. Monomeric triplexes A and B, B and C, and A and C, separately, were mixed in the presence of different concentrations (10 to 90 mM) of KOAc, pH 5.2 **(a, c, and e)** and of NaOAc **(b, d, and f)**

The mixtures were incubated for 18 hours prior to loading on to native gels containing and also run in 50 mM Tris acetate plus 5 mM KOAc (for the K^+ incubation) or 5 mM NaOAc (for Na^+ incubation) at pH 5.2. In each gel, heterodimers, if there were any, are indicated by red arrows.

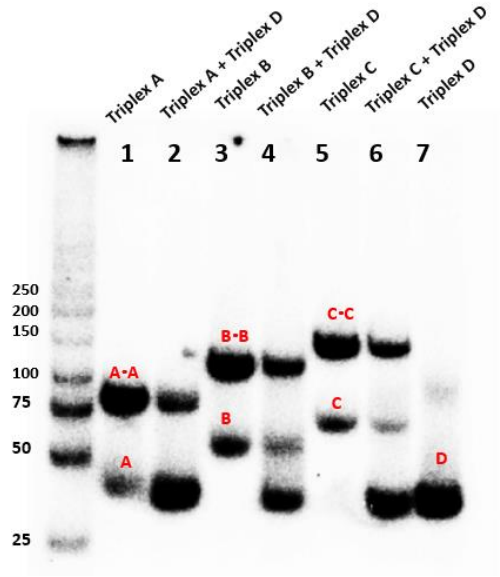


Figure 2.12. Mixing of triplex A, B, or C with triplex D does not generate heterodimers.

Triplex A, B, or C is mixed with triplex D, separately, in the presence of 50 mM KOAc. No heterodimer is observed on the native gel.

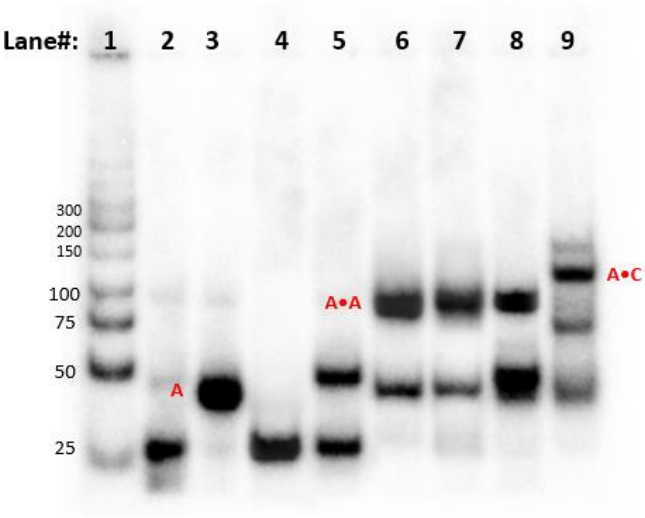
To test whether the blunt-ended monomeric triplex, D, was capable of forming heterodimers, it was mixed, individually, with monomeric triplexes A, B, and C in 50 mM KOAc (pH 5.2). Figure 2.12 shows that, only the A•A, B•B, and C•C homodimers formed, but no trace of any heterodimers containing D could be observed in the gel (Figure 2.12, lane 2, 4, and 6). This result, again suggests that in order for monomeric triplex tiles to dimerize, both triplex participants should have a protruding strand with two overhanging Gs to fill in the vacancy in the partner component, to enable G-quadruplex formation involving the two participating tiles.

To avoid the formal – albeit highly unlikely – possibility of strand exchange between triplex monomers during the dimerization reaction, each triplex was initially designed with its own unique sequence and length. For triplexes A and C, each of their component strands contained 25 and 45 nucleotides, respectively, excluding the guanine overhangs. To verify that no strand exchange occurs, an experiment was carried out of adding the constituent strands that eventually form A•C heterodimer in a reaction tube, one strand at a time, to check the what products formed at each stage. In

Figure 2.13.a, starting from lane 2 which contains the Watson-Crick duplex of A, the addition of the Hoogsteen strand of A (ssHY_A) in the absence of potassium ion gave rise to monomeric triplex A (lane 3). However, the addition of C Hoogsteen strand (ssHY_C) to duplex A, did not result in any triplex monomer formation (lane 4). In lane 5, the four strands of duplex A and duplex C were mixed and the yield was two clean bands corresponding to the two duplexes with their expected gel mobilities, namely, 25 and 45 nts, when compared to the double stranded DNA ladder (lane 1). In lane 6, the three component strands of triplex A were mixed again, but this time in the presence of 50 mM K^+, which led to the formation of homodimer A, represented by the presence of a mobility-retarded band, whose molecular size is approximately twice as big as the triplex monomer (lane 3). In lanes 7 and 8, the Hoogsteen strand and the Watson-Crick duplex of C were mixed, separately, with triplex A strands. The addition of Hoogsteen C (ssHY_C) to the triplex A strands did not result in a different triplex formation. Likewise, the mixing of duplex C with triplex A also did not yield any new products. Finally, the mixing of 6 DNA strands, 3 from triplex A and 3 from triplex C, gave rise to the A•C heterodimer. These results clearly demonstrate that no strand exchange occurred when the component strands of A and C were mixed, and that no non-specific products were formed. Therefore, one can safely assume that the product arising from the mixing of monomeric triplexes A and C is, indeed the triplex hetero-dimer, A•C.

Further to the above, a cross-labeling experiment was also carried out to verify the strand components present in the product of A plus C mixture. In Figure 2.13.b, lanes 5 to 7 contain three independent samples of A and C mixture, each with a different radio-labeled component strand. In lane 5, only the purine strand of A (ssR_A) was radiolabeled with ^{32}P, while C was wholly unlabeled. In lane 7, the purine strand of C (ssR_C) was radiolabeled, but none of the strands within A. In lane 6, both of the respective purine component strands of A and C were labeled with ^{32}P. The results seen are consistent with the designated A•C product containing *both* triplexes A and C.

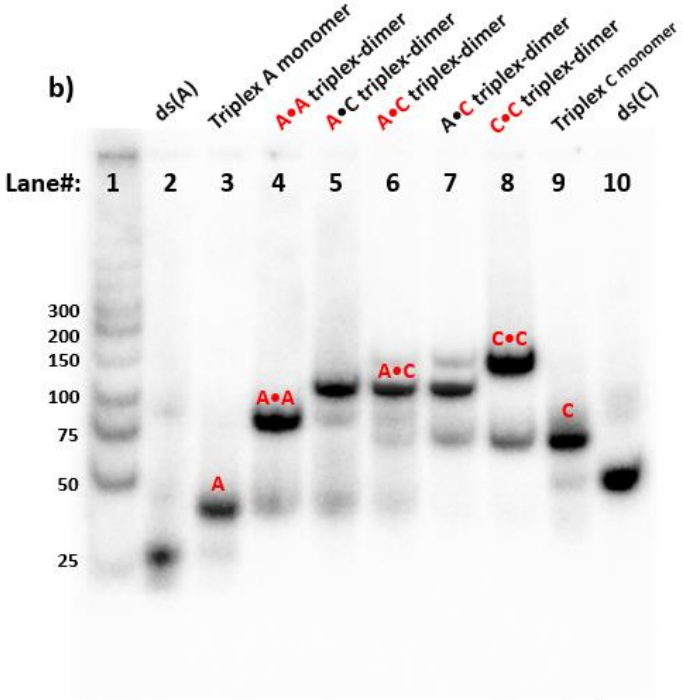
a)



11/23/2021

1: ds ladder
2: ssR(A) + ssY(A) = ds(A)
3: ssR(A) + ssY(A) + ssHY(A) = triplex A
4: ssR(A) + ssY(A) + ssHY(C) = ds(A) + ssHY(C)
5: ssR(A) + ssY(A) + ssR(C) + ssY(C) = ds(A) + ds(C)
6: ssR(A) + ssY(A) + ssHY(A) = triplex-dimer A•A
7: ssR(A) + ssY(A) + ssHY(A) + ssHY(C) = triplex-dimer A•A + ssHY(C)
8: ssR(A) + ssY(A) + ssHY(A) + ssR(C) + ssY(C) = triplex-dimer A•A + ds(C)
9: ssR(A) + ssY(A) + ssHY(A) + ssR(C) + ssY(C) + ssHY(C) = triplex-dimer A•C

b)



11/19/2021

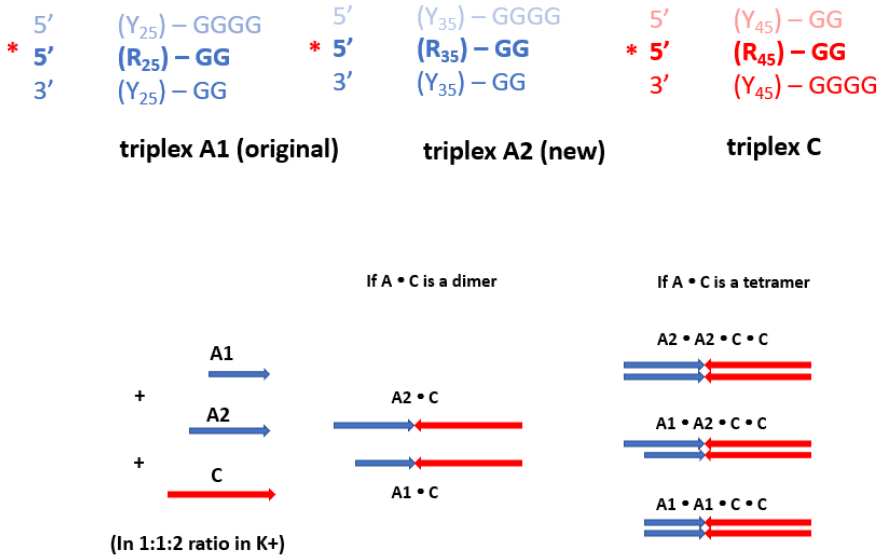
Figure 2.13. Addition of one component strand at a time and cross-labeling of A•C heterodimer.

(a) The purine strand (ssR) of triplex A and C were radiolabeled by P^{32} for visualization. All the DNA samples were assembled in the presence of 10 mM LiOAc plus 3 mM Mg(OAc)2 at pH 5.2. 10 mM of KOAc was then added to each assembled DNA complex, except for lane 3, and allowed to rest at 22℃ for 18 hours before running in a 7.5% native gel with 50 mM Tris acetate at pH 5.2. **(b)** The purine strand (ssR) of triplex A and C were radiolabeled by P^{32} for visualization. The letters highlighted in red represent the radiolabeled DNA.

To verify that A•C is indeed a dimer of triplexes, as opposed to a tetramer or some other composites, we carried out an experiment in which the monomeric triplex C was mixed in a 2:1:1 ratio in KOAc with two different lengths of the monomeric triplex A, ("A1" and "A2"), which have the same guanine overhangs but different-length YR*Y triplex regions (25 base triples for A1 and 35 base triples for A2). As illustrated in Figure 2.14.a, if A•C were a triplex-dimer containing one A tile and one C tile, the two different triplex A monomers should have equal probabilities of hybridizing with triplex C, given that their guanine overhangs are the same. Hence there should be two triplex-dimer products (and gel bands), A1•C and A2•C. On the other hand, if A•C were a triplex-tetramer, the mixing of triplex C with triplex A1 and A2 should lead to three different tetrameric products (and gel bands), namely, A1•A1•C•C, A1•A2•C•C, and A2•A2•C•C.

To test the above hypothesis, the incubated mixture of A1, A2, and C, along with other control samples, was run on a native gel containing and run in 50 mM Tris acetate and 5 mM KOAc, pH 5.2. Two major bands can be observed in lane 3 (Figure 2.14.b), which contained the A1, A2 and C mixture. The lower band has the same gel mobility as A1•C (lane 4), whereas the upper band has the same mobility as that of A2•C in lane 5. This result confirms our hypothesis, that A•C is indeed a triplex-dimer.

a)



b)

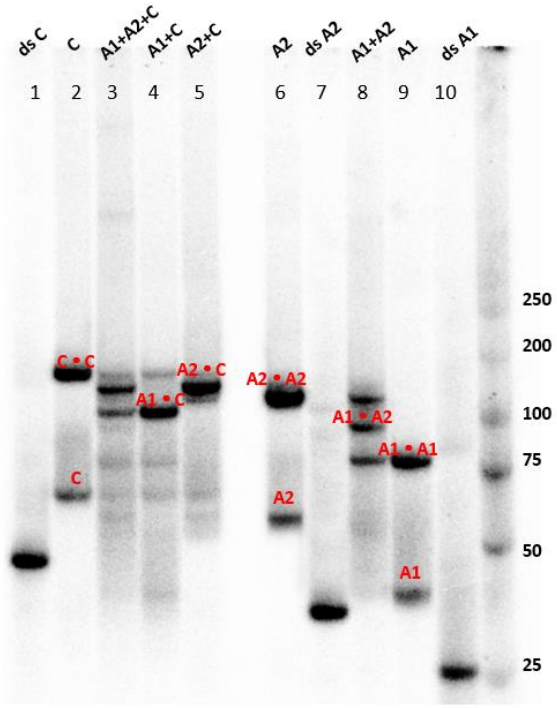


Figure 2.14. Verification that A•C is a triplex-dimer rather than a tetramer.
(a) The design of the experiment and the expected outcomes. Triplexes A1 and A2 have their own unique sequence and length, but share the same guanine sticky end. Pre-assembled A1, A2 and C were allowed to mixed in a 1:1:2 ratio in the presence of 50 mM KOAc. Depending on whether A•C is a dimer or a tetramer, different products are predicted to form. **(b)** Different DNA samples were run on a native gel containing 50 mM Tris acetate and 5 mM KOAc at pH 5.2. analysis.

Chapter 3. Mutually exclusive formation of A•C heterodimer in K^+ solutions versus A•A and C•C homodimers in Na^+ solutions

3.1. Introduction

Out of the three types of triplex heterodimers (A•B, B•C, and A•C) reported above, the A•C heterodimer was particularly striking, for two reasons.

First, the formation of any A•C triplex-heterodimer at all was a completely unexpected result, in as much as it violated the "socket-plug" complementarity rules presumed for our observation of triplex homodimer formation. Based on opposing the G-rich ends of triplexes A and C in head-to-head fashion (as illustrated in Figure 2.9), one might not expect A and C to dimerize together, given that the juxtaposition of overhanging guanines and the guanine-less cavities from these two tiles do not suggest an obvious "socket-plug" complementarity fit (Figure 3.1). That is presuming the absence of some form of distortion or re-arrangement of the strands. The fact that triplexes A and C are able to recognize and bind to each other clearly indicates the possibility of either an unusual geometric arrangement between the guanine residues involved in the tile-bridging G-quadruplex or the formation of a lesser (i.e. less than four full quartets) quadruplex (that was stable in the presence of the highly stabilizing ion K^+ but not with the less stabilizing Na^+). Two potential models were proposed to rationalize the A•C dimer. Figure 3.1 shows one of the possible ways for triplexes A and C to bind each other: either, by pushing the two extra guanines on triplex C's protruding strand out of the way such that the remaining guanine residues highlighted in red boxes could form a 2-layered G-quadruplex to hold two tiles together; alternatively, the extra Gs on the Hoogsteen of triplex A could be pushed away to allow the remaining Gs to form a regular 2-layered quadruplex.

Second – and more puzzling – was the observation that the mixing of triplexes A and triplexes C in the presence of different counterions (K^+ and Na^+) resulted in distinct triplex-dimer products. In K^+, the A•C heterodimer was strongly favoured over the A•A and C•C homodimers, whereas in the case of Na^+, the A•C heterodimer was completely absent, while the homodimers formed in roughly equal measure (Figure2.11.e-f). These striking observations suggested that the G-quadruplex involved in the A•C heterodimer

likely constituted a distinct and perhaps incomplete structure, in terms of guanine-guanine pairing (and perhaps containing less than four complete G-quartets), that nevertheless formed rapidly, faster than the “complete”, four-quartet socket-plug G-quadruplexes likely forming in and holding together the A•A and C•C homodimers. Such a kinetically favoured but possibly structurally less robust G-quadruplex could nevertheless be sufficiently stable in highly stabilizing K^+-containing solutions as to dominate the population of possible triplex-dimers. In Na^+ solutions, however, only thermodynamically robust but relatively slow-forming homodimers would be stable enough to accumulate. An analogous scenario, of alternative G-quadruplex structures (some kinetically favoured and some thermodynamically favoured) forming in K^+ versus Na^+ solutions, was first noted in 1990 (24). In terms of what kind of G-quadruplex(es) might exist within the A•C heterodimer, it is also possible that K^+ (which typically shows a distinctive eight-fold coordination to the keto oxygens of guanines from two adjacent G-quartets) may enable the formation of topologically or geometrically different G-quadruplex that could not easily form with the four-fold coordination properties of Na^+ (which typically binds in the plane of individual G-quartets, coordinating to the four keto oxygens available). For these intriguing reasons, I spent much effort in trying to understand the structure of the mysterious A•C triplex heterodimer.

We used a variety of different techniques to study the properties of the A•C heterodimer. First, in addition to the study of electrophoretic mobilities in native polyacrylamide gels, we used Förster resonance energy transfer (FRET) to validate the formation and dominance of the A•C heterodimer of in K^+ but not in Na^+. Second, to characterize the base pairing details of individual guanines within the G-quadruplex (or quadruplexes) forming in the different triplex-dimers, dimethyl sulfate (DMS) methylation protection assays were carried out for the A•A, A•C, and C•C triplex-dimers. The kinetics of formation and the stabilities (as reported by melting behaviour) of these three triplex-dimers were also investigated using both the gel electrophoresis and the FRET methods.

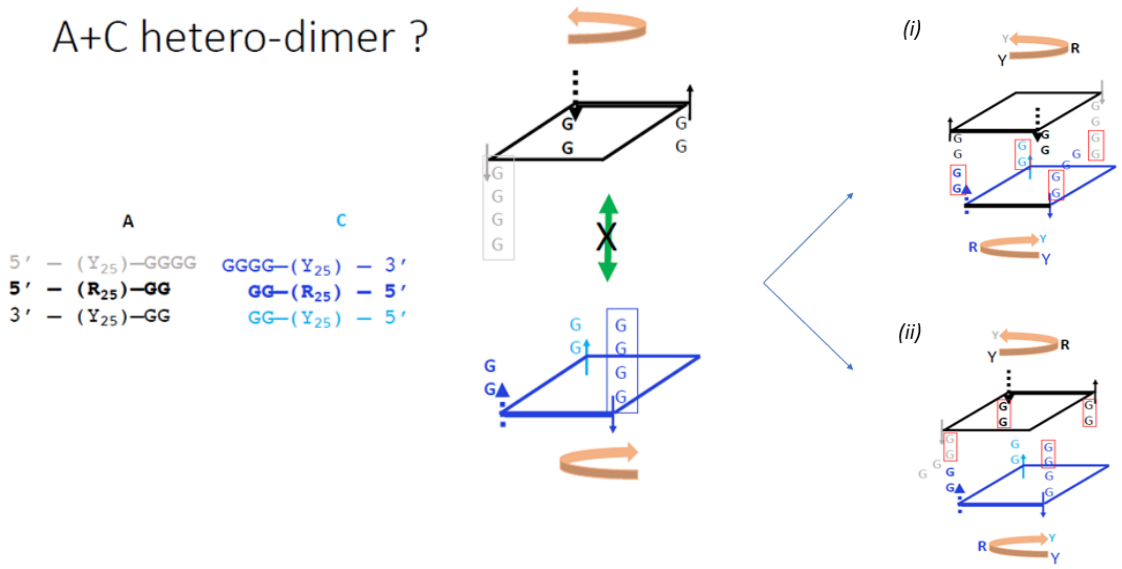


Figure 3.1. Schematic representation and proposed models for the A•C hetero-dimer.

The monomeric triplex tile A is shown in black and the triplex tile C is shown in blue. In standard head-to-head geometric juxtaposition, the G-rich sticky ends of triplex A and C are not complementary to each other in a "socket-plug" sense, to enable the formation of a regular 4-layered G-quadruplex.

3.2. Materials and methods

3.2.1. Alkyne-modified DNAs and fluorophore-azides

Alkyne-incorporating oligonucleotides were purchased from Integrated DNA Technologies, Inc. (Coralville, Iowa, United States). Prior to their use in CuAAC "click" reactions, the oligonucleotides were purified via ethanol precipitation and dissolved in TE buffer. The fluorophore azides, Cy3-azide (90%, HPLC purified) and Cy5-azide (95%, HPLC purified) purchased from Sigma-Alrich Canada Co. (Ontario, Canada), were dissolved in double-distilled H_2O to final concentrations of 1 M.

3.2.2. CuAAC reaction

The CuAAC reaction protocol was adapted and modified from (67). The alkyne-containing DNA oligonucleotides were first heat-denatured at 100 ℃ for 4 minutes in 2% v/v DMSO in water, and immediately transferred to ice. 100 μM fluorophore azide, followed by pre-mixed 100 μM $CuSO_4$ and 500 μM THPTA (final concentrations), were added to the denatured DNA (1 μM). 2.5 mM (final) of freshly prepared sodium

ascorbate solution was then added to the mixture to initiate the CuAAC reaction. The THPTA was added to serve a dual purpose: 1) to maintain the Cu(I) oxidation state of copper sources and 2) to protect biomolecules from oxidative damage during the cycloaddition reaction(68).

The reaction mixtures were briefly vortexed, spun down, and incubated at 22℃ in a thermal cycler for 1 hour, following which the DNA was ethanol precipitated, washed with cold 70% ethanol, air-dried, and dissolved in 10 μL of TE buffer. After mixing with denaturing loading dye, the DNA solution was loaded onto an 8% denaturing gel (containing 7 M urea). The fluorophore-attached DNA bands, visualized using a ChemiDoc™ MP Imager (Bio-Rad Laboratories, Inc.), were excised from the gel, eluted into TE buffer and recovered by ethanol precipitation.

3.2.3. Fluorescence spectroscopy and FRET measurements

Fluorescence spectra were measured on a fluorescence plate reader (Infinite M200 Pro, Tecan). 20 μL of each sample was loaded into an independent well of a black, flat bottomed 384-well Falcon plate (Corning). Samples were photo-excited at 500 nm light and emission spectra were recorded from 520 to 750 nm. All the measurements taken were repeated for three independently made up samples.

3.2.4. Kinetics analysis

Where kinetics of formation were measured using gel electrophoresis, the quantification of triplex-dimer formation was carried out with ImageQuant TL 8.1 software (Cytiva). The peak density of each band in a given lane was manually selected and quantified. Percentages of triplex-dimer formation were defined as the intensity of a band corresponding to a particular triplex-dimer was divided by the total intensity of all bands in that gel lane, The percentages obtained were plotted against time of incubation. For kinetics of 0.5 μM triplex-dimer formation in the first 10 minutes at 22℃ in 1 mM KOAc or 10 mM NaOAc, pH 5.2, solutions, the data were fit to linear regression lines and the graphs were plotted using SigmaPlot software (SPSS Inc.).

Where kinetics of formation were measured using FRET, the intensity of the FRET signal from mixture of A and C (each at 0.5 μM) as a function of time was

recorded in 5-minute intervals over the first hour of incubation at 22℃. The change in the intensity of the fluorescence signal over time was plotted together in a graph along with the first-hour kinetic analysis data obtained from gel electrophoresis using SigmaPlot (above) The two sets of data, from FRET and from gel electrophoresis, were both fit to non-linear regressions and the respective k_{obs} values were generated using the program SigmaPlot.

3.2.5. Triplex-dimer stability analysis

Triplex-dimers were preassembled at 1 μM concentrations and incubated at room temperature (22℃) in the presence of 50 mM KOAc or NaOAc, pH 5.2. Triplex-dimers so formed were heated at various temperatures for 4 minutes in a thermocycler (PerkinElmer), following which the DNA samples were immediately loaded onto a 7.5% native gel containing 50 mM Tris acetate, pH 5.2, and 5 mM of either KOAc or NaOAc. Densitometry analysis was performed to calculate the percentage of dimer remaining unmelted, which was defined as the fraction of triplex-dimer remaining from the total DNA in a given lane. The percentages of unmelted dimer so calculated were plotted against the temperatures to which the dimers were treated. The melting temperature (T_m) was determined as the temperature at which 50% of the dimer remained intact, relative to the starting temperature at 22℃. Each data point on a curve plotted represented the average value obtained from three replicate experiments.

3.2.6. DMS methylation-protection assay

Prior to carrying out the DMS methylation-protection experiments, the DNA components of interest were assembled and allowed to rest at 22℃ for 18 hours to enable the formation of the desired complexes. Subsequently, these DNA complexes were treated with 0.2% freshly prepared dimethyl sulfate (DMS) (Thermo Fisher Scientific Inc.) in 10 mM Tris (pH 5.2). Methylation was allowed to proceed for 30 minutes, after which 1.5 μL of β-mercaptoethanol (Thermo Fisher Scientific Inc.) was added to quench to methylation reaction. Following the above treatment, the DNA solutions were mixed with native gel loading buffer and loaded onto a 7.5% native polyacrylamide gel containing and also run in 50 mM Tris acetate and 5 mM LiOAc (pH 5.2). Bands corresponding to the desired DNA structures were excised from the gel. The

DNAs were eluted overnight from excised gel pieces into TE buffer and recovered via ethanol precipitation. Next, the recovered DNAs were treated with freshly prepared 10% (v/v) piperidine at 90℃ for 30 minutes followed by lyophilization. The DNA pellets obtained were dissolved in denaturing gel loading buffer (containing 95% formamide, 1 mM EDTA, bromophenol blue and xylene cyanol), heat-denatured at 100℃ for 4 minutes, and immediately put on ice prior to loading onto an 8% denaturing polyacrylamide gel (containing and run at 50 mM Tris-borate-EDTA). For different G-ladder references, selected individual single-strands were radiolabeled with ^{32}P at their 5' ends and subsequently treated with DMS and piperidine following the same procedure described above.

3.3. Förster resonance energy transfer: an alternative approach for investigation on triplex-dimers

The copper (Cu^{+})-dependent azide-alkyne cycloaddition ('CuAAC') reaction (Figure 3.3.a), also known as the "Click" reaction, is one of the most useful techniques for the covalent conjugation of two molecular entities containing, respectively a terminal alkyne and an organic azide (69). One of the many applications of azide-alkyne cycloaddition chemistry as applied to nucleic acids is the efficient and cost-effective covalent labeling of oligonucleotides with fluorophores or other pendant labels.

Among the various types of molecular probes available for the conjugation to nucleic acids, labeling oligonucleotides with fluorescent dyes represents perhaps one of the most popular approaches. Prior to the development of CuAAC reaction, labeling oligonucleotides with fluorophores was mostly achieved by coupling between amino groups to carboxylic acids via activated esters. Two of the major drawbacks of such coupling reactions when compared to the azide-alkyne cycloaddition reaction are low efficiency and low stability of the active esters in aqueous solutions (as they can be easily hydrolyzed). Conversely, using the CuAAC reaction to conjugate various azido-fluorophores to alkynated DNA or RNA leads to more efficient and high yielding conjugation. Organic azides and alkynes are generally stable in the aqueous solutions used for the coupling and only react exclusively with each other, thereby generating little in the way of side-products.

Multiple positions on nucleotides are suitable for modification with alkyne or azide moieties. The most commonly used sites of modification include the phosphate group, the 5’ and 3’ positions on the ribose sugar, and the C5 position of pyrimidines (70). Indeed, 5-octadiynyl dU is one of the most widely used alkyne-modified bases, which incorporates an eight-carbon linker terminating in an alkyne functionality.

Cyanine (Cy) dyes are among the most popular organic fluorophores that have found a wide range of application in biochemistry and biotechnology, especially for imaging and quantification purposes. Owing to their high yield of bright and stable fluorescence, the dyes Cy3 and Cy5 constitute the most widely used donor-acceptor pair for the study of various molecular interactions. Cy3 is a bright, water-soluble, and pH-insensitive orange-fluorescent dye with excitation and emission peaks centred at 520 nm and 570 nm, respectively; while, Cy5 is a far-red fluorescent dye with excitation and emission peaks centred at 630 nm and 670 nm, respectively (Figure 3.2). The significant overlapping region between the emission spectrum of Cy3 and the excitation spectrum of Cy5 renders them an excellent choice FRET pair for investigating distances within 10 to 100 Å. Both Cy3 and Cy5 have been modified with terminal azide moieties (Figure 3.3.b), rendering them suitable for coupling with alkynated oligonucleotides using CuAAC.

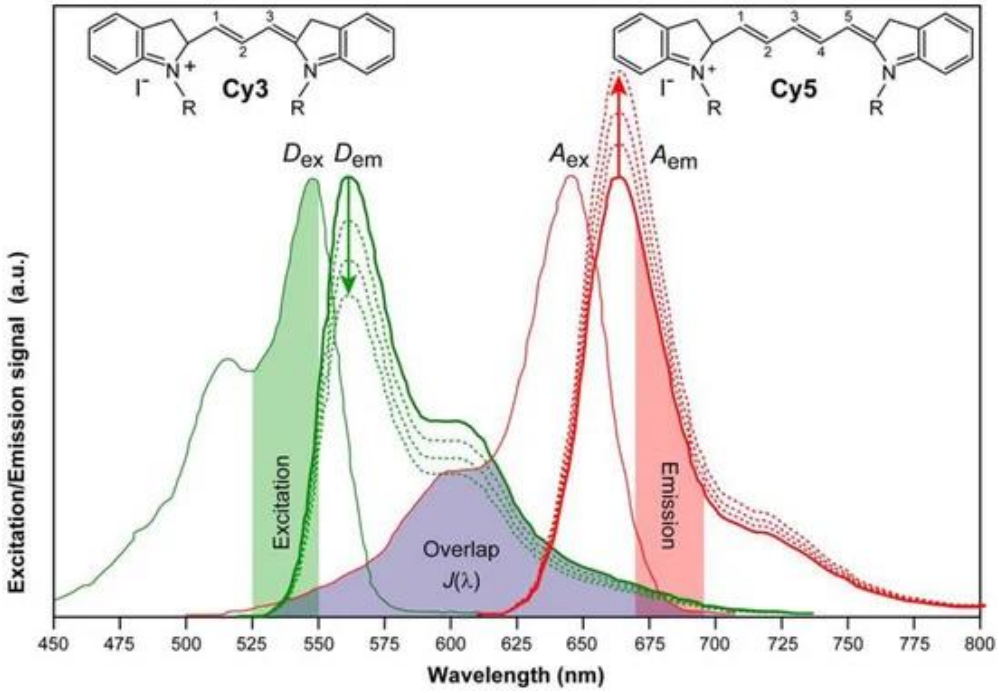


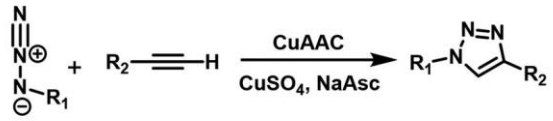
Figure 3.2. Fluorescent energy transfer between Cy3 and Cy5.
The overlap of the Cy3 emission spectrum with Cy5 excitation spectrum. Adapted from (71).

3.4. Fluorescence labeling of alkyne-modified DNA using "click" chemistry

Apart from radiolabeling, fluorescent labeling of nucleic acids is also widely used. Although the latter's sensitivity is slightly lower than the former's, fluorescence labeling of nucleic acids offers some key advantages over radiolabeling: (a) fluorescent dyes can generate signal in real-time with high resolution; (b) the handling of fluorophore dyes is free of radiation hazards (72). This method has found wide application in not only sequencing, but also diagnostics of genetic diseases, gene polymorphism studies, quantitative analysis of gene expression, etc (73). Here, to study the interaction between the triplex tile A and the tile C in the presence of K^+ versus Na^+ counterions, we labeled the respective Watson-Crick Y strands of triplex A and C with, separately, Cy3-azide and Cy5-azide at specific locations. To incorporate an alkyne group into the oligonucleotides, 5-octadiynyl dU, was inserted during chemical synthesis at designated positions close, in each case, to the 5' end of the W-C Y strand (Figure 3.3.c). The site of modification was carefully chosen such that it was in sufficient proximity to the G-rich sticky ends to

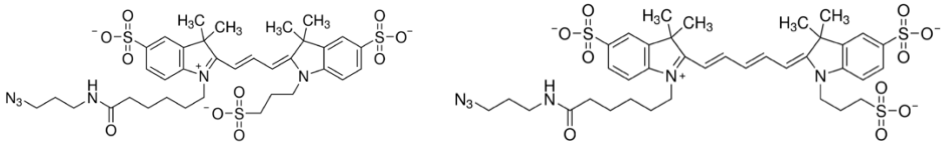
register a significant FRET, yet not so close to the guanines as to hamper their G-quadruplex formation. CuAAC was used to covalently attach the Cy3 and Cy5 dyes to the alkynated olionucleotides.

(a)



Azide Alkyne Triazole

(b)



Cy3-azide Cy5-azide

(c)

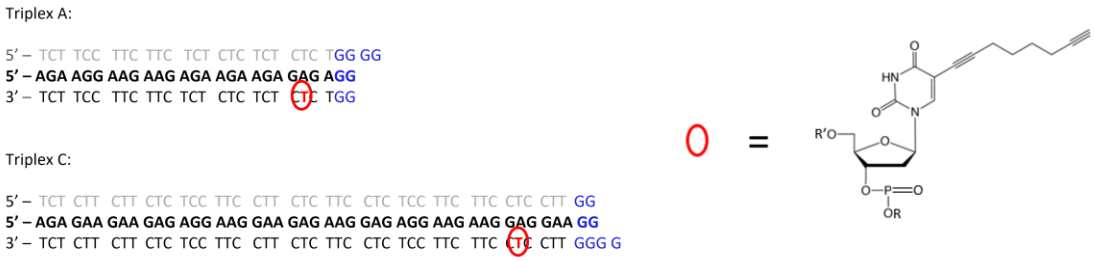


Figure 3.3. CuAAC reaction for conjugation between DNAs and fluorophore dyes.

(a) CuAAC reaction mechanism. Adapted with permission from (74) **(b)** Chemical structures of Cy3-azide and Cy5-azide. **(c)** Incorporation of a modified base 5-Octadiynyl dU with a terminal alkyne to DNA.

3.5. Ion-specificity of triplex-dimers manifested in contrasting fluorescence signals from K^+ versus Na^+ solutions

In addition to our gel mobility-based approach at studying monomeric triplex tile dimerization, we used Förster Resonance Energy Transfer (FRET) to validate the formation of different “socket-plug” hybrids in the presence of K^+ versus Na^+. The rationale was that if triplexes A-Cy3 and C-Cy5 were able to form the heterodimer, A•C, in the presence of K^+, such a dimerization should bring the two fluorophore dyes

together sufficiently close for the observer to detect a FRET signal emitted from Cy5 (670 nm) following excitation of exclusively Cy3 at 500 nm. Conversely, if A-Cy3 and C-Cy5 were unable to dimerize together in the presence of Na^+, exclusive excitation of Cy3 should not result a FRET emission signal from Cy5. Figure 3.4 outlines the design and logic behind such a FRET experiment. Briefly, three independent triplex A and C mixtures were prepared in the presence of the appropriate counterions. The first mixture (Mixture 1) contained both Cy3-labeled A and Cy5-labeled A in the presence of unlabeled C to report on the formation of the A•A homodimer. Similarly, the second mixture (Mixture 2) contained both Cy3-labeled C and Cy5-labeled C, in the presence of unlabeled A, to report the formation of the C•C dimer. Finally, the third reaction mixture (Mixture 3) contained Cy3-labeled A and Cy5-labeled C to report on the formation of the A•C heterodimer.

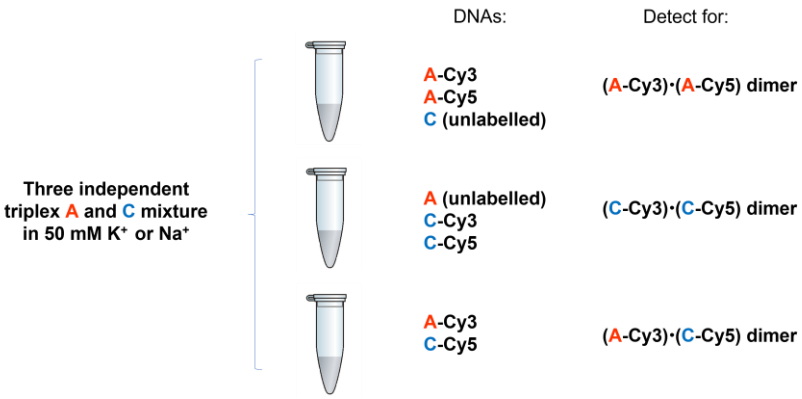


Figure 3.4. FRET experiment design for validation of different "socket-plug" hybrids forming in K^+ versus Na^+.

Figure 3.5 shows the emission spectra of the above three triplex tile mixtures when they were incubated in either 50 mM of KOAc, pH 5.2, or in 50 mM NaOAc, pH 5.2. In the presence of K^+, a FRET signal was observed only in Mixture 3, consistent with the formation of the A(Cy3)•C(Cy5) heterodimer, but no FRET signal was observed in Mixtures 1 and 2, consistent with the formation of neither the A(Cy3)•A(Cy5) homodimer nor the C(Cy3)•C(Cy5) homodimer. By contrast, in the presence of Na^+, Mixtures 1 and 2 (containing, respectively, the A(Cy3)•A(Cy5) homodimer and the C(Cy3)•C(Cy5) homodimer) showed FRET signals; while Mixture 3 showed no FRET signal, consistent with the non-formation in it of the A(Cy3)•C(Cy5) heterodimer. All of these above results were fully consistent with the results we obtained from gel

electrophoresis (Figure 2.11.e-f), providing independent evidence of the sensitivity of the A•C triplex heterodimer formation in the presence of a specific counterion K^+.

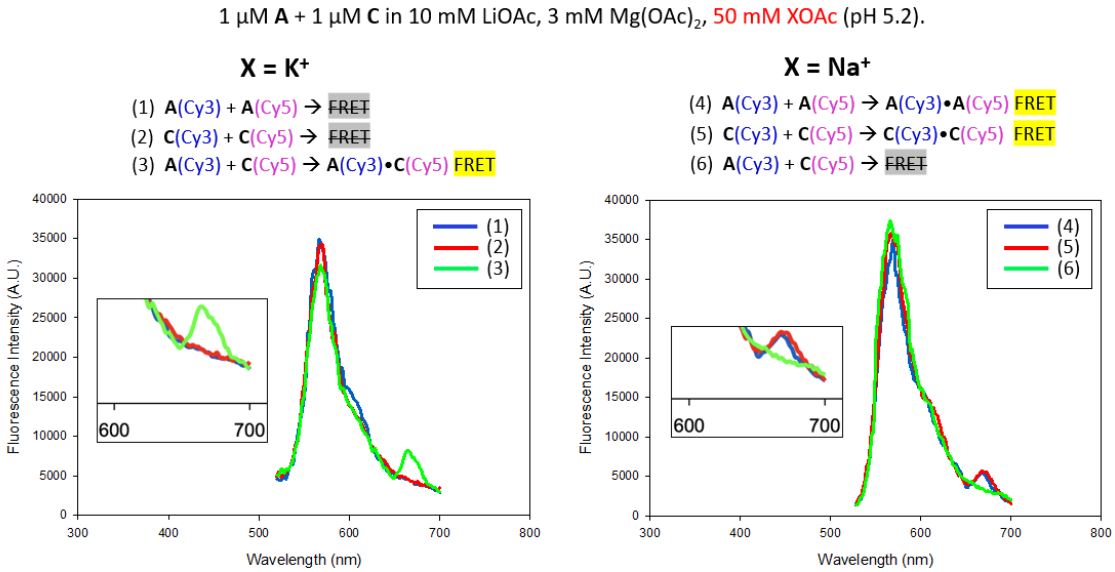


Figure 3.5. Detection of FRET signals for verification of different TQ hybrids in different counterions.
Cy3 (570 nm) and Cy5 (670 nm) dye emission spectra for mixtures of triplex A and C in the presence of K^+ versus Na^+.

3.6. DMS methylation: investigation on the G-quadruplex-junction between dimerized triplex tiles

To investigate whether or not a G-quadruplex structure was shared between the dimerized triplex tiles, DNA footprinting experiments were carried out with dimethyl sulfate (DMS), a chemical reagent that methylates the N7 position of guanines within either single-stranded DNA or DNA helices in which the N7 position is not participating in hydrogen bonding. For example, a guanine in a Watson-Crick duplex would be methylated by DMS, but not in the R strand of YR*Y triplexes nor in G-quadruplexes, because the N7 positions in these two structures are involved in Hoogsteen hydrogen bonding. Our various DNA complexes (with a single DNA strand at a time 5'-labeled with ^{32}P) were first treated with 0.2% DMS, then run and purified from a pH 5.2 native gel. The DMS-treated DNA samples were subsequently pelleted by ethanol precipitation, washed, and subsequently treated at 90℃ for 30 min with 10% v/v freshly prepared aqueous piperidine to break DNA phosphodiester backbone(s) at site(s) of guanine methylation. DNA lyophilized after this treatment were then loaded and analyzed on 8% polyacrylamide sequencing gels, as shown in Figure 3.6.

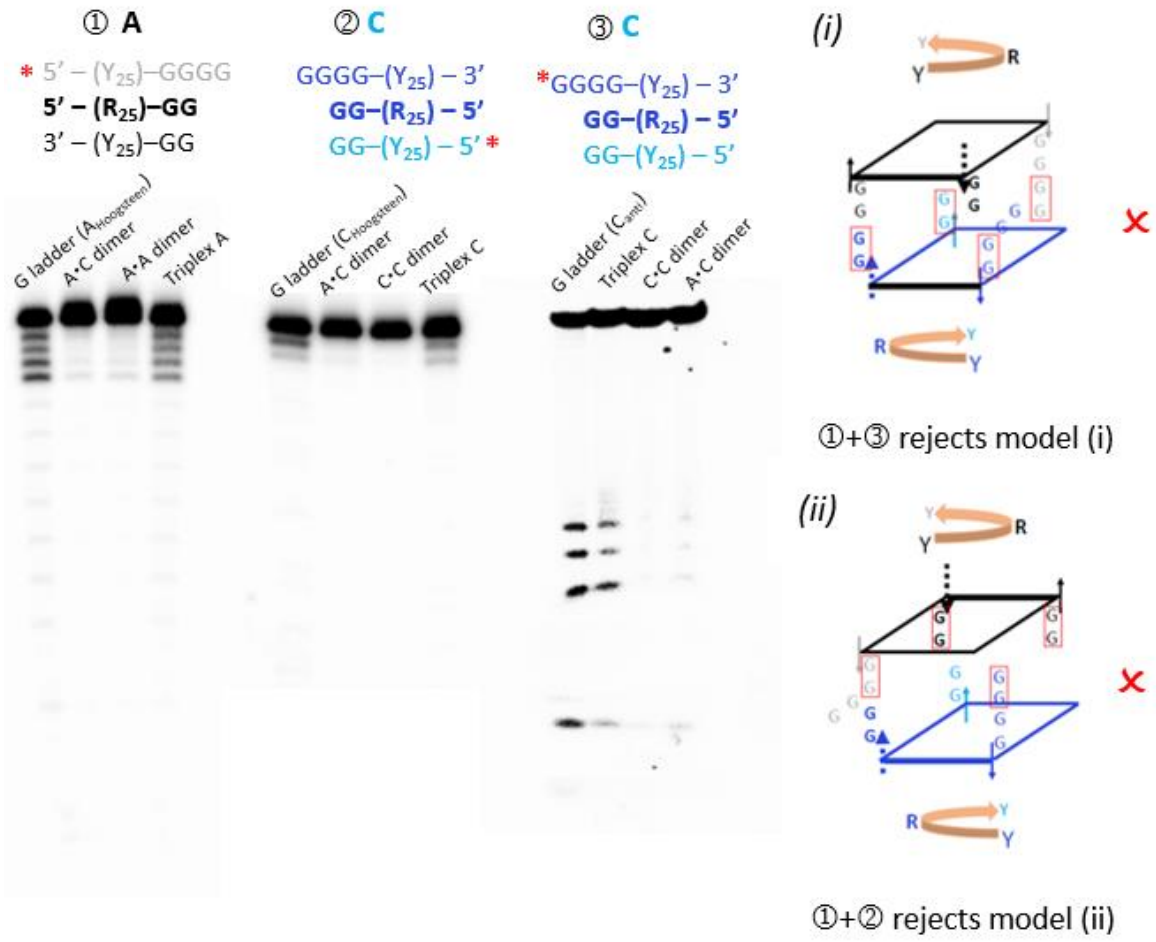


Figure 3.6. DMS footprinting results for various DNA complexes.
Different DMS-methylated DNA complexes were treated with 10% v/v hot aqueous piperidine and lyophilized to dryness prior to running on denaturing gels. In ① and ②, the samples were run on an 8% denaturing gels, while in ③, samples were run on a 20% denaturing gel. The red asterisks indicate the strands that had been 5' end-labelled by ^{32}P, as well as used as reference G-ladders.

As expected, all guanines present within the sequences of ^{32}P-5'-labelled *single-stranded* DNAs ("G ladder" in different gels) were reactive with DMS since their N7 positions were not actively involved in hydrogen bonding. Similarly, the unpaired guanine residues in the monomeric triplex tiles ("Triplex A" and "Triplex C") were also unprotected from DMS methylation because they were freely overhanging at one end of the triplex. By contrast, the guanines in all three triplex-dimers (A•A, C•C, as well as A•C) were unreactive with DMS, indicated by the absence of bands in comparison with the G-ladders and the monomeric triplexes. These results are consistent with *all* of the examined guanines within each of the above triplex-dimers participating in G-quadruplex formation. In other words, these results confirm that the dimerization of triplex tiles is

indeed achieved by the formation of G-quadruplex by the overhanging guanine residues on the ends of individual tiles.

On the other hand, these DMS footprinting results also reject the two hypothesized models for the A•C dimer illustrated in Figure 3.6. If model (i) were the correct model, two of the following premises should hold true: Firstly, two guanines at the very 5' end of the anti-strand in triplex C should be methylated by DMS, since they are pushed away from the central cavity; secondly, the two terminal guanines on the Hoogsteen strand of triplex A should be protected from methylation while the interior two should not – as in this model only the terminal Gs are involved in forming a 2-layered G-quadruplex. The fact that all four guanines in the Hoogsteen strand of triplex A are well-protected from DMS methylation indicates that all of them are involved in Hoogsteen hydrogen bonding, hence no guanines are excluded or pushed away from the centre. In addition, since all four guanines in the anti-sense strand of triplex C are also protected, it means that the G-quadruplex that is bridging tile A and C should contain 4 layers of G-quartet instead of 2. The same logic can be applied to reject model (ii) based on the complete protection from DMS in the Hoogsteen strands of triplexes A and C.

3.7. Kinetics of formation and stabilities of triplex-dimers

To gain a better understanding of this special "socket-plug" complementarity, we further investigated the kinetics of formation as well as stabilities of A•A, C•C, and A•C triplex-dimers in K^+ and in Na^+, respectively. For the initial kinetics experiments (Figure 3.7.a), 50 mM KOAc was added to 0.5 μM of triplex monomers pre-assembled in 10 mM LiOAc plus 3 mM $Mg(OAc)_2$ at various time points, and loaded on a native gel, after which a densitometry analysis was performed to determine the percentage of dimer forming at each time point. Figure 3.7.a-b (each representing three independent experiments) show that for both A and C, the dimerization reaction reached completion within the first hour of incubation in 50 mM K^+. To observe the kinetic behaviours of A and C where the dimerization was occurring in a linear time-dependent fashion, the concentration of potassium ion was reduced down to 1 mM and the time window was adjusted to the first 10 minutes of reaction. Figure 3.7.c shows that in 1 mM K^+, over the initial 10 minutes of incubation, the rate of formation of the A•C hetero-dimer was approximately two-fold faster in comparison to the formation of A•A and C•C homodimers, both of which formed at approximately the same speed. The fact that A•C

was able to form at a faster rate than A•A or C•C in K^+ suggests that the A•C heterodimer is likely to be a kinetically favoured product relative to the homodimers. The kinetics of homo-dimerization of A and C in Na^+ were also characterized and the results were shown in Figure 3.7.d.

Along with our gel electrophoresis-based measurements, we also used our FRET-based approach to study the first-hour kinetics of formation of the A•C heterodimer in the presence of 1 mM KOAc. Figure 3.8 shows results of kinetic analyses using both the gel electrophoresis and, separately, the FRET method (also derived from three independent experiments); it can be seen that the results are highly consistent with each other. In the first 10 minutes of reaction, the FRET signal increases in an approximately linear fashion, following which the intensity of signal continues to increase, albeit with a decreasing rate, until a dimerization plateau is reached in approximately 30 minutes, indicating a reaching of equilibrium for the reaction. For comparison purposes, in the gel method, the band intensity of A•C heterodimer after 1 hour of reaction is assumed to represent the equilibrium value, and the intensities of other time points are compared to that of 1-hour time point (which then represent the percentage of reaction completion at those given time points). Collectively, the results from these two independent experimental approaches suggest that A•C heterodimer has a fast rate of formation in the presence of K^+ counterion, hence a structure that is kinetically favoured.

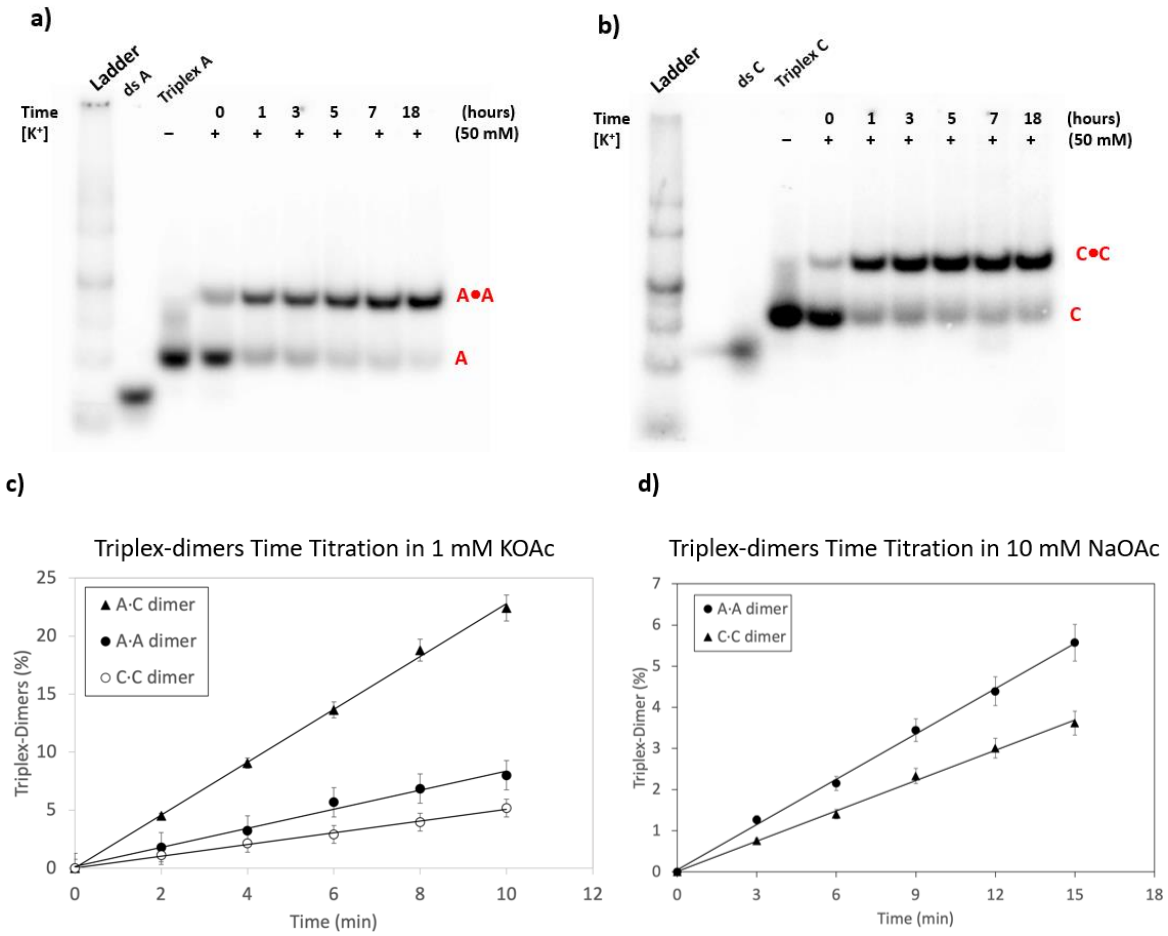


Figure 3.7. Kinetics studies of A•A, A•C, and C•C triplex-dimers in 1 mM KOAc and in 10 mM NaOAc.

In 1 mM K^+, the best-fit regression lines are plotted for A•A (●), C•C(○), and A•C(▲) triplex-dimers with estimated slopes of 0.8, 0.5, 2.3 %/min, respectively. In 10 mM Na^+, the estimated slopes of A•A (●) and C•C(▲) homodimer formation are 0.4 and 0.3 %/min respectively. The error bars represent one standard deviation from the mean value obtained from three independent experiments.

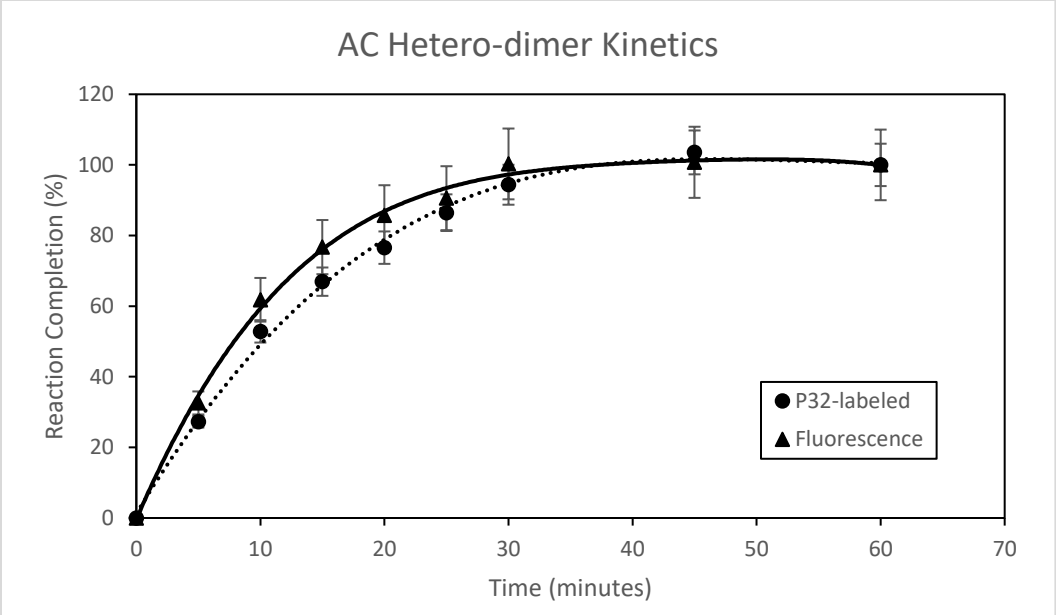


Figure 3.8. First-hour kinetic analyses of A•C heterodimer formation using FRET and gel electrophoresis.

The dimerization reaction was initiated by the addition of 1 mM KOAc to mixtures of 1 μM A and 1 μM C in 10 mM LiOAc and 3 mM $Mg(OAc)_2$, pH 5.2 and incubated for up to 1 hour. Data points collected using FRET(▲) and gel(●) methods are plotted on the same graph. Their k_{obs} values during the first 10 minutes of reaction were estimated to be 8.86 ± 0.12 min^{-1} and 8.99 ± 0.07 min^{-1}, respectively. The error bars represent one standard deviation from the mean value obtained from three independent experiments, each of the gel mobility and FRET experiments..

To examine the thermal stabilities of the various triplex-dimers in K^+ versus in Na^+ (Figure 3.9) melting experiments were performed upon each dimer. Briefly, A•A , A•C, and C•C, each at 1 μM DNA concentration, were assembled in separate tubes in the presence of 10 mM LiOAc, 3 mM $Mg(OAc)_2$, plus 50 mM of either KOAc or NaOAc. Following 18 hours of incubation at 22 ℃, each DNA sample was heated at various temperatures for 4 minutes before immediate loading on a 7.5% native gel running continuously in 50 mM Tris acetate, pH 5.2 (the immediate loading was done to minimize any re-dimerization of the "melted" triplex dimers). Densitometry analysis was then performed to estimate the amount of DNA that remained as triplex-dimer after heating at different temperatures. The melting temperature was calculated as the temperature at which 50% of the triplex-dimer were remained after heating, in comparison to the unheated sample (22℃). In K^+, the melting temperatures for A•A (●), A•C(▲), and C•C(O) triplex-dimers are determined to be 73°, 77°, and 79℃, respectively. Since Na^+ only allows homodimer formation, the melting behaviour of A•C heterodimer could not be investigated in Na^+, while the melting temperature of both homo-dimers was

approximately 52℃. Based on these results, the A•C hetero-dimer appears to have a stability comparable to those of the homo-dimers. Therefore, one cannot rule out the possibility that the A•C dimer, in addition to being kinetically favoured, is also a thermodynamically favoured product that simply forms faster than other, comparably thermodynamically favoured products.

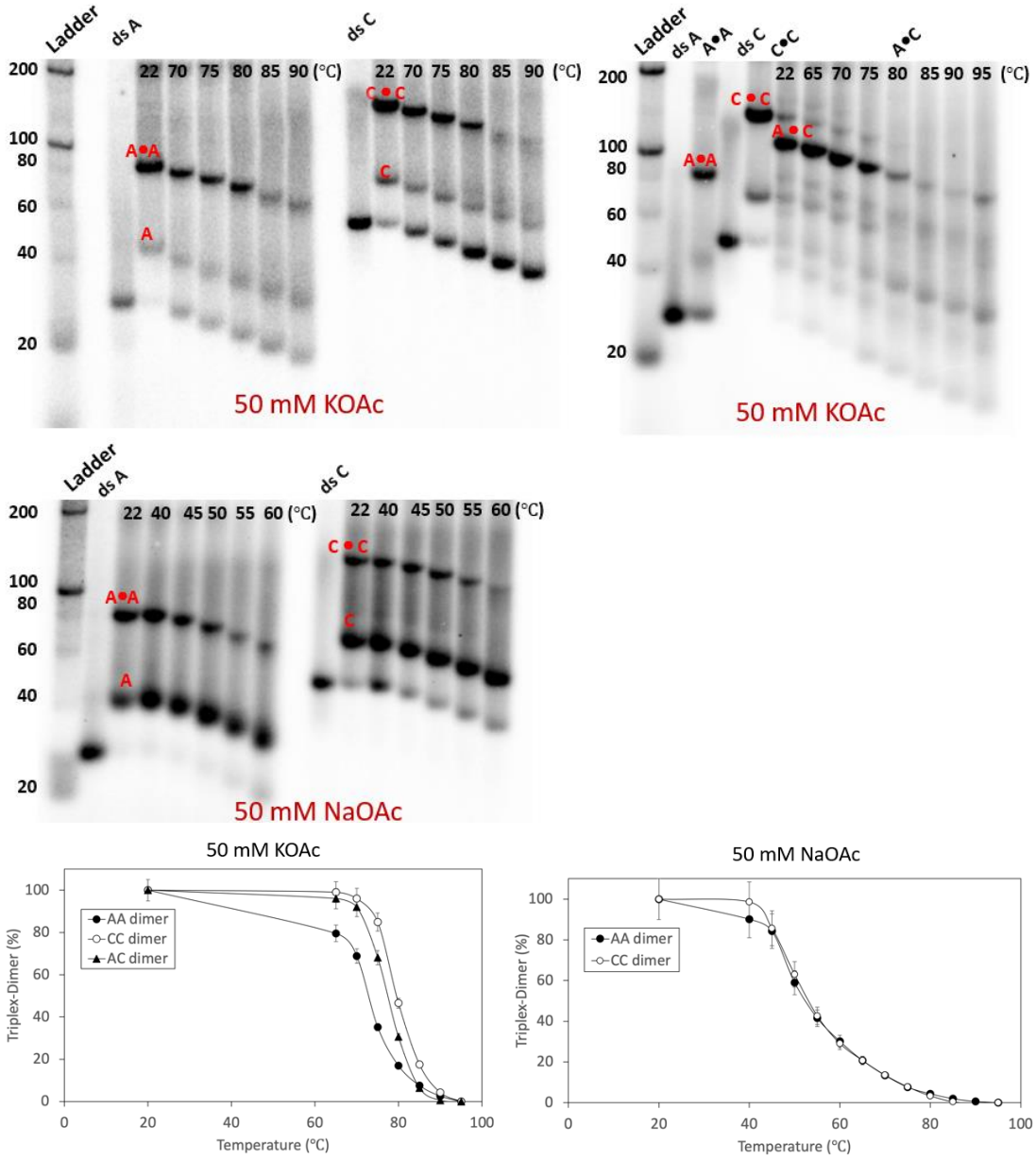


Figure 3.9. Native Gel analysis of melting experiments and the corresponding normalized melting curves for triplex-dimers in K^+ versus in Na^+.

Triplex-dimers were preassembled and rested at room temperature (22℃) for 18 hours in the presence of 50 mM KOAc or NaOAc. Dimers were heated at various temperatures for 4 minutes on a thermocycler, after which the samples were immediately loaded onto a 7.5% native gel

containing 5 mM of KOAc or NaOAc at pH 5.2. The percentages of triplex-dimer that remained unmelted after heating at different temperatures were estimated by densitometry and plotted as a function of temperature. Each data point on the graphs was the average value collected from three independent replicate experiments.

3.8. Formation of nano-scale TQ hybrid tile-based nanocomposites in K^+ and in Na^+

The origins of this project lie in the prior work from our group on the use of guanine-based "sticky ends" at the ends of DNA triplexes to build triplex-quadruplex (TQ) hybrid one-dimensional nanostructures by Lat *et al.* (64). In light of what we have learned about the cation-selectivity of tile homodimerization versus heterodimerization in this work, we investigated whether different nanostructure assembly outcomes would result from the same mixture of tiles, depending on the choice of counterion.

In creating a simple system for testing this idea, we modified triplex A into a double-headed triplex tile that contained the exact same G-rich sticky end (i.e. the sticky ends on either side were both at the 3' end of the respective R strand) overhanging both ends of the tile (Figure 3.10). We then co-incubated this double-headed tile, named "AA", with tile C, with Na^+ counterions and, separately, K^+ counterions, and also under conditions where [AA] >> [C], and where [C] >> [AA]. We found that, once again, the presence of K^+ versus Na^+ led to dramatically different results. Figure 3.11.a-d shows, schematically, our predictions of what kind of tile assemblies might form under these above conditions and the actual gel results corresponding to each scenario are presented on the right to the schematics. Indeed, with the double-headed triplex A alone in the presence of K^+, we were able to join multiple monomeric triplex tiles together and form a DNA TQ hybrid composite on a very large size scale (Figure 3.11.a, lane 3 and c, lane 3), in analogous to the 1-dimensional DNA nanostructure built by Lat *et al.* (64). When the double-headed triplex AA assembled in K^+ is run on a K^+-containing native gel, the composite was unable to enter the gel and instead was restricted to the bottom of the well, as highlighted in Figure 3.11.a and c, suggesting the formation of very high molecular weight species, likely a nanowire. However, when the double-headed triplex AA was incubated in Na^+, the formation of such nanostructure was completely absent (Figure 3.11.b and d). Instead, the bulk of the double-headed triplex tiles remained as monomers, with only a few managing to hybridize with others to form homodimers $(AA)_2$ and, to lower extent, homotrimers $(AA)_3$. The fact that potassium ion can lead to the

assembly of a nano-scale TQ hybrid composite but not sodium is consistent with potassium being a much more effective G-quadruplex-stabilizing alkali cation than sodium, in accordance with most earlier studies on the role of metal cations on G-quadruplex stability (75).

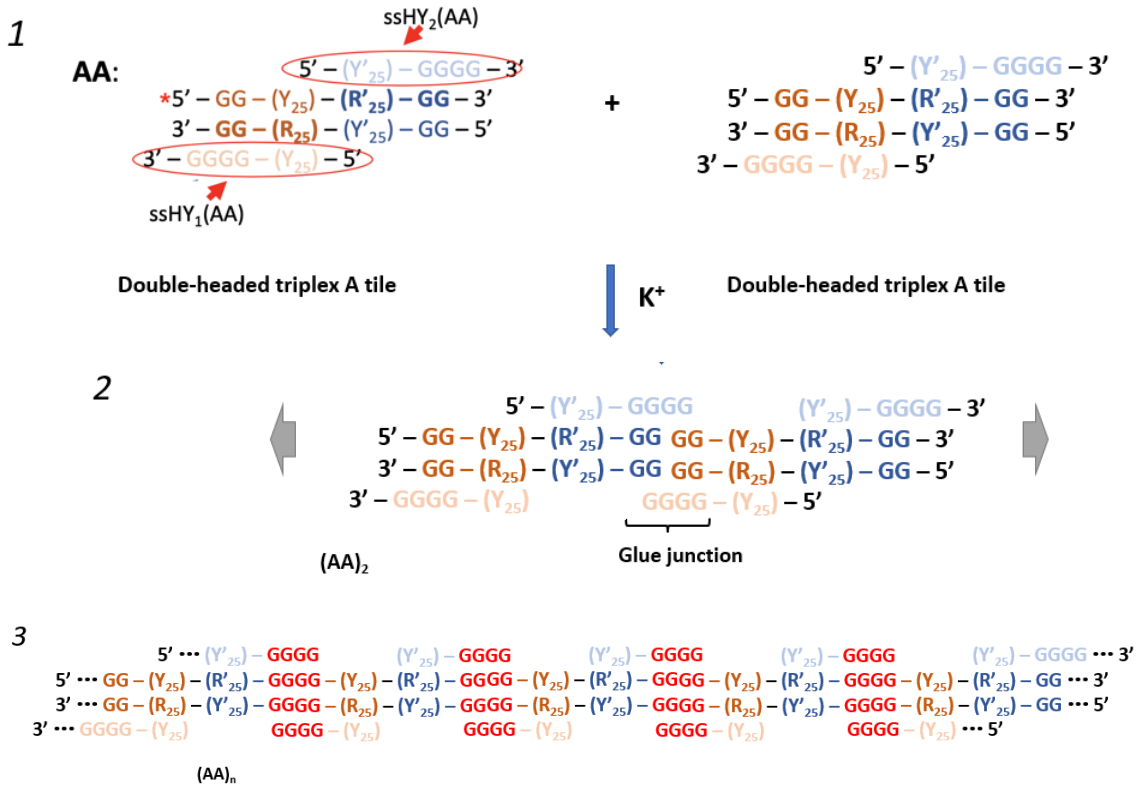
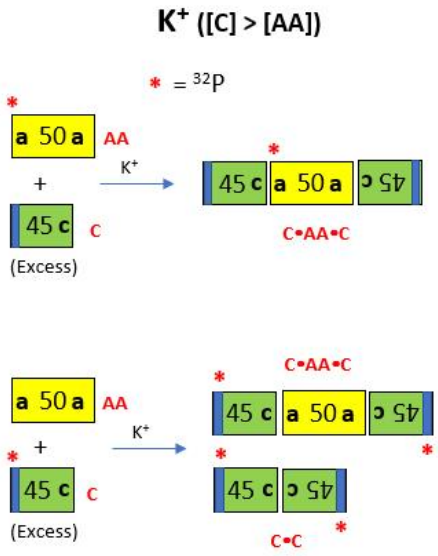


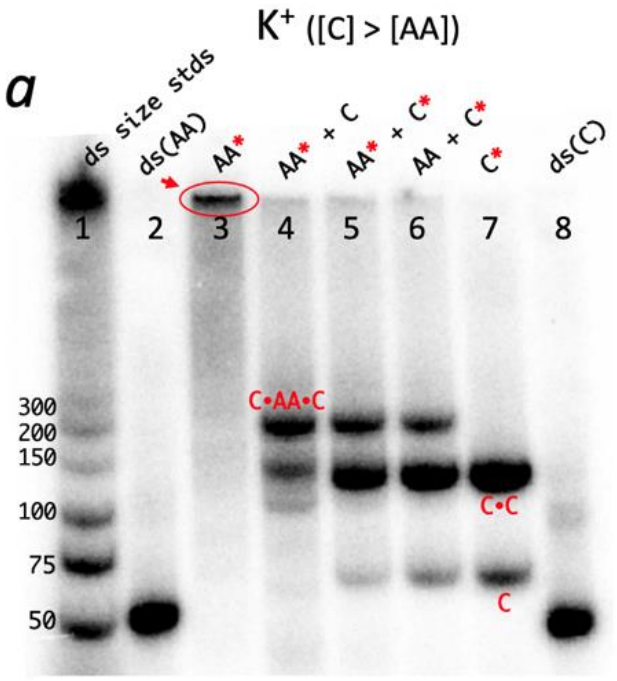
Figure 3.10. Design of double-headed triplex A (“AA”) and its potential to assemble into a 1-dimensional triplex-quadruplex hybrid nanostructure.

In the design of the tile AA, the duplex component of the original single-headed triplex A was extended by 25 nucleotides and two additional guanine residues were added on the blunt-ends. A second Hoogsteen strand with four guanine overhangs is added to allow the formation of a triplex with two identical guanine-only sticky ends on both sides.

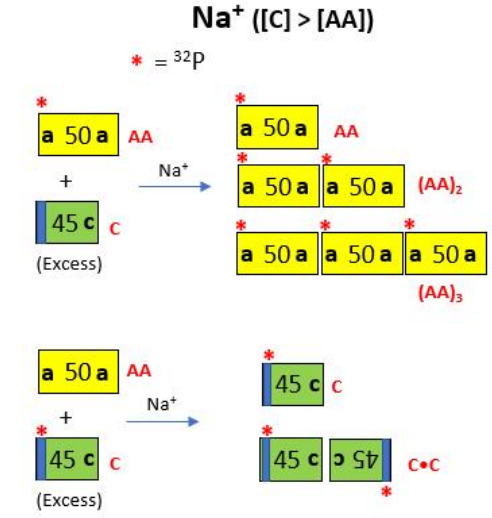
Given the specific ion requirement for the formation of A•C heterodimer, the effect of the different cations is clearly manifest in four different scenarios. Based on our previous finding about the unique property of A•C heterodimer, Figure 3.11 shows the predicted outcomes in each different scenario, side-by-side with the actual experiment results, highly consistent with our predictions.

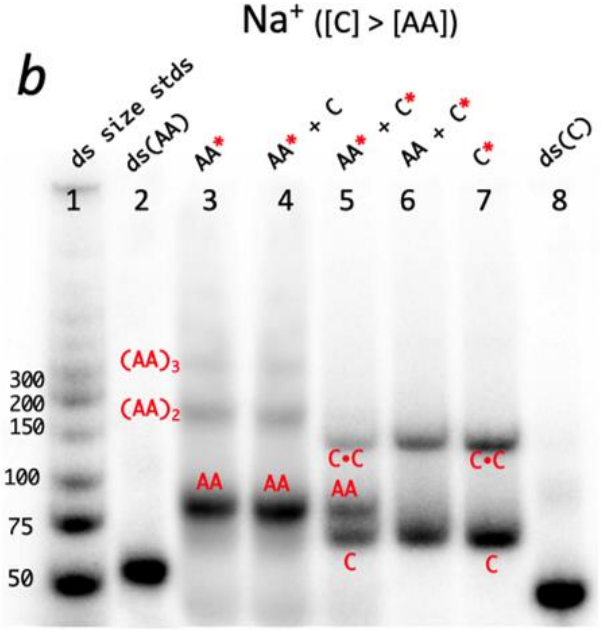
Scenario 1: Excess C with limited AA in K^+





Scenario 2: Excess C with limited AA in Na^+





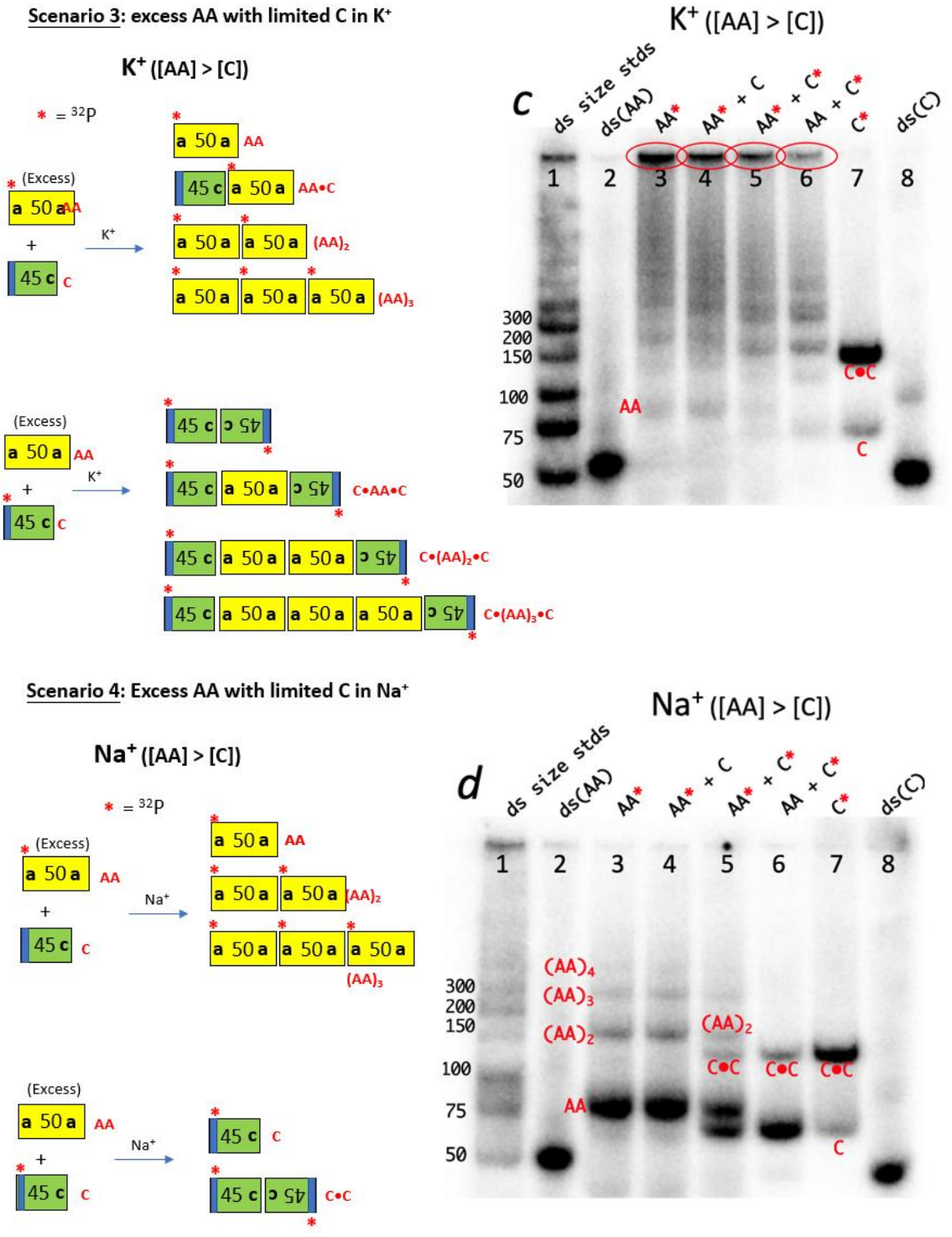


Figure 3.11. Predicted outcomes of mixing double-headed triplex A with single-headed triplex C in four different scenarios.

Double-headed triplex A (“AA”) and single-headed triplex C were mixed in four different conditions: **(a)** [AA] = 1 μM, [C] = 4 μM, KOAc (pH 5.2) = 50 mM; **(b)** [AA] = 1 μM, [C] = 4 μM,

NaOAc (pH 5.2) = 50 mM; **(c)** [AA] = 4 μM, [C] = 1 μM, KOAc (pH 5.2) = 50 mM; **(d)** [AA] = 4 μM, [C] = 1 μM, NaOAc (pH 5.2) = 50 mM. The red asterisk signs indicate radiolabeling by ^{32}P.

In the first scenario, 1 μM of double-headed triplex A was mixed with 4 times more concentrated single-headed triplex C in the presence of K^+([C]>[AA] in K^+). As shown in Figure 2.11.e, heterodimer formation was strongly favoured over homodimers in K^+. For this reason, the excess triplex C was able to "cap" the double-headed AA such that it prevented AA from self-hybridizing to form a continuous long stretch of nanocomposite. Instead, a hetero triplex-trimer was formed, denoted as C•AA•C (Figure 3.11.a Lane 4). Conversely, when [AA] > [C] in K^+, the excess double-headed AA tiles mainly self-hybridized and gave rise to a high molecular weight species that were unable to enter the gels. Yet, in lane 5 and 6 on Figure 3.11.c, one can also see a few bands within in the smear, indicating that a small amount of double-headed A was capped by triplex C, the limiting agent in this scenario. In the case of Na^+, however, whether [AA] > [C] (Figure 3.11.b) or [C] > [AA] (Figure 3.11.d), there was no great difference in terms of outcome. In both Na^+ scenarios, AA largely remained as monomer, with a few homodimers and homotrimers forming. When mixed with triplex C, no heterotrimer was observed, which is consistent with the previous results when single-headed triplex A and C were mixed in Na^+.

To test whether or not the actual assembling method had an effect on the formation of the composites, we used two different ways to assemble these DNA nanocomposites. In one case, the monomeric triplexes AA and C were first assembled separately in 10 mM LiOAc plus 3 mM $Mg(OAc)_2$, These two assembled triplex monomers were then combined together in one tube before the addition of 50 mM KOAc or NaOAc to initiate the formation of nanocomposites from them. In the second method, all the component DNA single strands of triplexes AA and C were simply mixed together in one reaction tube, containing 10 mM LiOAc plus 3 mM $Mg(OAc)_2$, following which 50 mM of either KOAc or NaOAc was added in, followed by incubation. The native gel results presented in Figure 3.12 suggest that there is no significant difference in the outcomes between the two assembly methods, which suggests that the assembly of such DNA composites is highly robust.

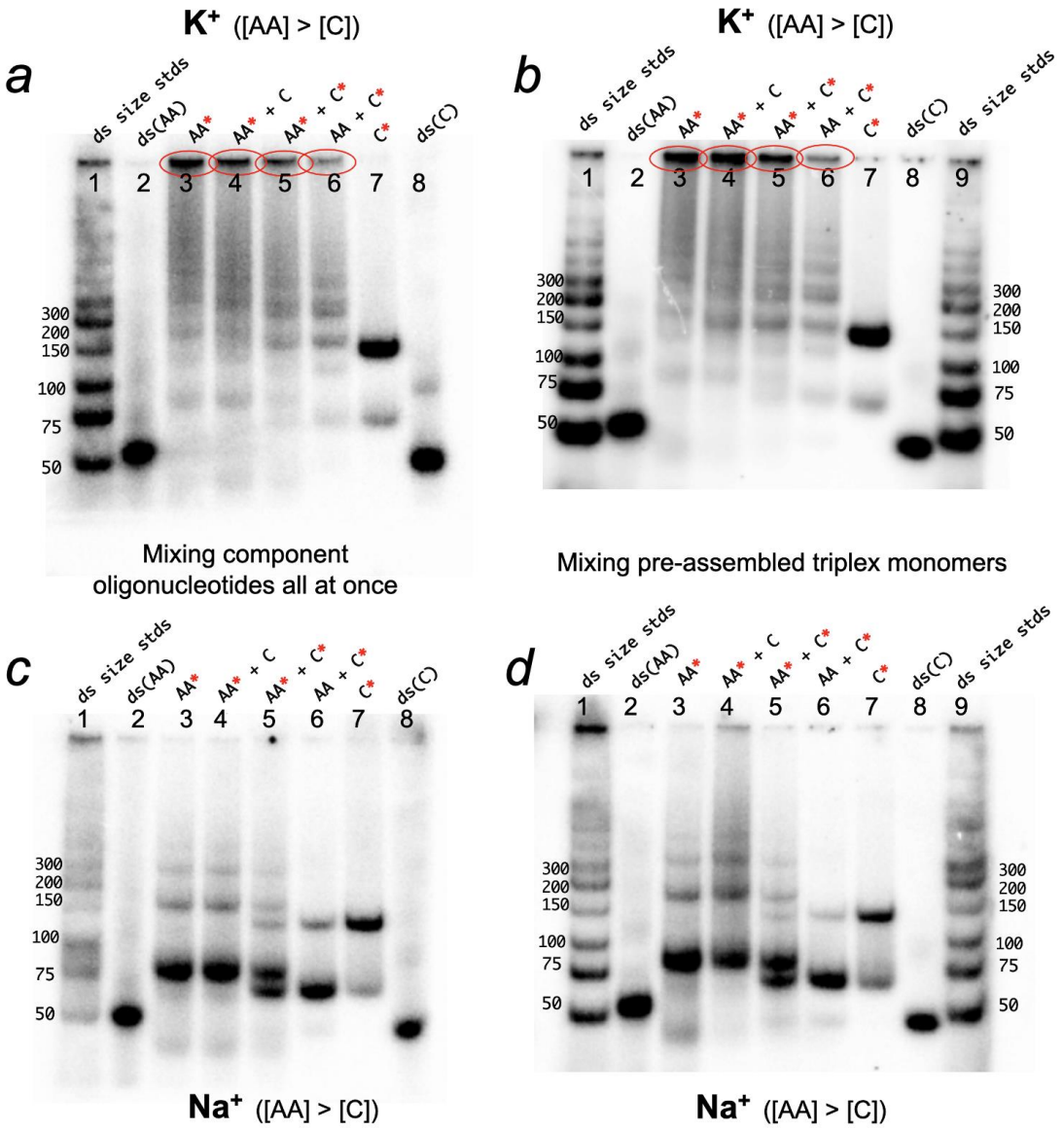


Figure 3.12. Two different methods for assembling DNA nanocomposites lead to the same outcome.

Different DNA composites were assembled either by **(a, c)** mixing component oligonucleotides all at once or, **(b, d)** by mixing the pre-assembled triplex monomers together before adding in K^+/Na^+.

The significance of these experiment was not merely to demonstrate the A•C heterodimerization's requirement for K^+ counterions. More importantly, the fact that the double-headed triplex AA was able to assemble into a nanoscale TQ hybrid composite and the fine-tuning of the ratio between [AA] and [C] could lead to strikingly different products, especially in K^+ but not in Na^+, provides a versatility that should find wide application in the field of nanotechnology. One can potentially control the size of the

desired TQ hybrid nanostructure simply by adjusting the ratio between the concentrations of [AA] and [C], and the appropriate ionic conditions.

3.9. An exploration into the possibility of ion- or pH-dependent reversibility of TQ hybrids

Reversibility is a highly useful property for a nanomaterial as it allows one to have control over the formation and/or dissolution of nanocomposites as desired. We conducted experiments to explore whether these DNA triplex-quadruplex hybrid composites possess any reversibility property, either by switching/sequestering the cations or by altering pH in the environment.

Since potassium acts as the initiator of triplex-dimerization by promoting and stabilizing the formation of the G-quadruplex joining the two triplex tiles together, the removal of potassium ion should, in theory, trigger the collapse of the G-quadruplex structures and thereby disrupt the dimerization. 18-crown-6 is a cyclic organic compound that functions as a ligand for a number of metal cations (76). Like all other crown ethers, 18-crown-6 has a central cavity that can accommodate the cation, which is stabilized by interaction with the lone pairs of electrons on the 6 oxygen atoms within the ring. The availability of crown ethers with cavities of different sizes allows specific cations to be bound with a high degree of selectivity. The ion selectivity of the crown ether depends both on the size of the cations relative to the cavity of the crown ether and on the delicate balance between the cation-crown and the cation-solvation interaction (77).18-crown-6 is known to have a particularly high affinity for potassium ions with a binding constant of 10^6 M^{-1} in methanol. The relative free energy of solvation of K^+ in methanol and the relative free energy of binding of K^+ to 18-crown-6 in methanol was estimated to be 19.6 kcal/mol and -3.5 kcal/mol (78). Therefore, it has been widely used as a phase transfer catalyst or separation agent (79) to sequester potassium ions in the solution.

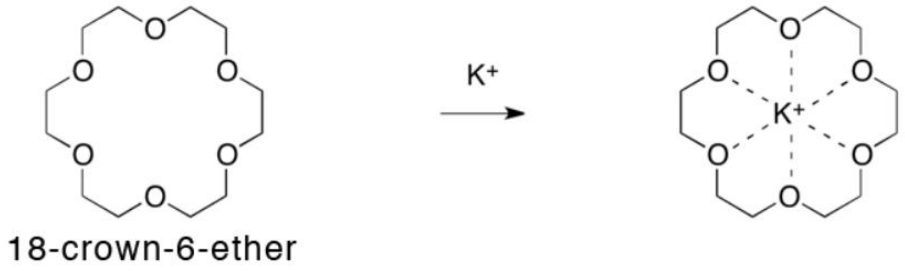


Figure 3.13. 18-crown-6 ether selectively binds to potassium.
The potassium ion fits well within the central cavity of 18-crown-6-ether and coordinates with the 6 oxygen atoms within the ring. Adapted from (66).

To test whether the triplex tile dimerization can be reversed by the removal of potassium ion, 5 mM of freshly prepared 18-crown-6 was added to 1 μM of the triplex heterodimer A•C pre-assembled in 1 mM KOAc. After 18 hours of incubation at 22℃, the DNA sample was run on a native gel containing 50 mM Tris acetate with 5 mM LiOAc. However, no significant difference was observed in the percentage yield of triplex-dimer between the 18-crown-6 ether-treated versus untreated samples (lanes 6 and lane 7 in Figure 3.13.a). This result suggests that the 18-crown-6 ether, although present in a 5-fold excess over the potassium ions, failed to remove the latter that were tightly bound within the G-quadruplex. Hence the dimerization could not be reversed. To test whether increasing the incubation temperature would help the crown ether remove the potassium ions coordinated in the central cavities, another sample of A•C with 18-crown-6 added was incubated at 40℃ in a thermocycler for 18 hours before gel analysis. Again, the gel result (Figure 3.13.b) shows that even raising the temperature did not have a significant effect on destabilizing the triplex-dimer.

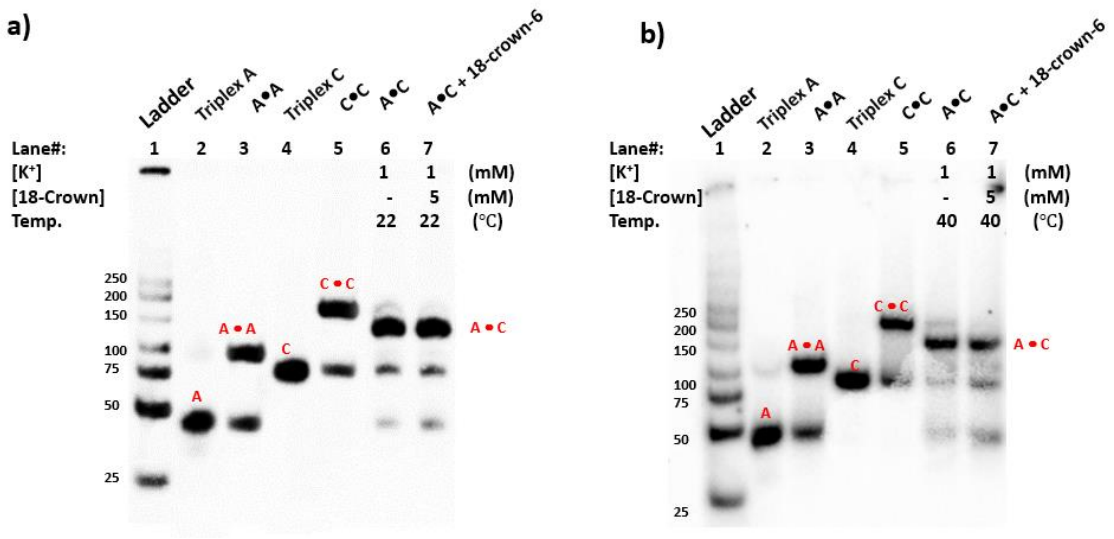


Figure 3.14. 18-crown-6 ether has little effect on reversing the formation of the A•C heterodimer.

5 mM of freshly prepared 18-crown-6 was added 1 μM of A•C heterodimer made in 10 mM LiOAc, 3 mM $Mg(OAc)_2$, and 1 mM KOAc. After addition of crown ether, the samples were allowed to rest at either 22 ℃ **(a)** or 40℃ **(b)** for 18 hours prior to running on native gels with 50 mM Tris acetate and 5 mM LiOAc at pH 5.2.

Since the formation of the A•C heterodimer appears to be strongly disfavoured in the presence of Na^+, we wondered if adding excess of Na^+ to A•C pre-formed in a low K^+ concentration could break down the heterodimer and promote the formation of the homodimers A•A and C•C. A time-titration experiment was conducted in which 50 mM of NaOAc was added to 1μM of A•C heterodimer pre-formed in 1 mM KOAc (after 18 hours of incubation at 22℃) at various time points to test if the heterodimer would breakdown and homodimers would form. The results shown in Figure 3.14 suggest that even adding Na^+ 50 times in excess to K^+ could not breakdown the heterodimer, indicating that the G-quadruplex structure linking the two monomeric triplex tiles in A•C is highly stable and the potassium ions are well-coordinated within the quartet layers with little exchange of ions with the surrounding solution over the experimental timescale.

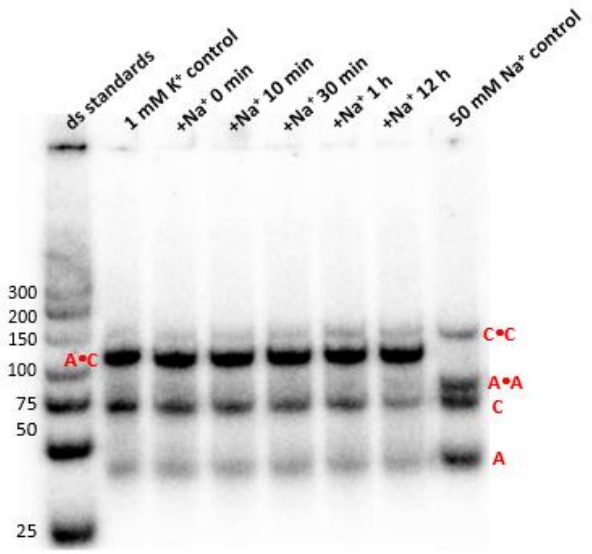


Figure 3.15. Addition of excess Na^+ at various time points to A•C heterodimer pre-formed in K^+. did not reverse the dimerization.

Apart from sequestration of cations in solution, we also investigated whether changing the pH, another important factor in the formation of triplex monomers as well as triplex-dimers, could potentially be utilized to reverse the dimerization reaction or break these structures down altogether. To carry out the experiment, A•A, A•C, and C•C triplex-dimers were first assembled independently at pH 5.2 (in 10 mM LiOAc, 3 mM $Mg(OAc)_2$, and 50 mM KOAc). After 18 hours of rest at 22°C, the samples were run on a native gel at pH 7.5, containing 50 mM Tris borate (TB) and 5 mM LiCl. In contrast to the native gels at pH 5.2, the mobility-retarded bands that correspond to various triplex-dimers as well as triplex monomers were no longer present (Figure 3.15). The bands remaining in the lanes are showed the same gel-mobility as those of duplex A and duplex C, respectively. This result indicates that by increasing the pH from 5.2 to 7.5, all triplex-dimers as well as the monomeric tiles were broken down to their constituent duplexes. This is because both the triplex tiles and the triplex-dimers contain CGC^+ base triples, whose formation and stability require a mildly acidic environment. The loss of cytosine protonation as a result of the pH increase is sufficient to completely destroy the various triplex-containing structures. Therefore, we have demonstrated that pH can be used as an important tool to disassemble the triplex-dimers down to their constituent Watson-Crick duplexes.

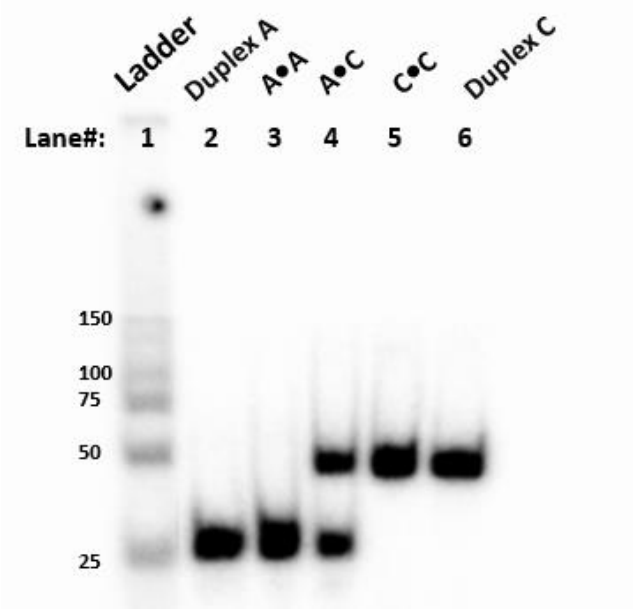


Figure 3.16. Raising the pH of the solution leads to complete breakdown of triplex-dimers as well as monomeric triplex tiles down to Watson-Crick duplexes.

1 μM each of the A•A, A•C, and C•C triplex-dimers were first assembled independently at pH 5.2 (in 10 mM LiOAc, 3 mM $Mg(OAc)_2$, and 50 mM KOAc). After 18 hours of rest at 22℃, the samples were run on a native gel at pH 7.5, containing 50 mM Tris borate (TB) and 5 mM LiCl.

Chapter 4. Investigation of the impact of alternative metal cations and of the number of G-quartets required for triplex dimerization.

4.1. Introduction

To further characterize the “glue” G-quadruplex that enables the dimerization of triplex tiles, we carried out investigation on two key features of G-quadruplexes that are essential for GQ stability: coordination with counterions and the number of G-quartets necessary for a robust “glue” G-quadruplex.

Depending on their binding locations, the interaction between cations and G-quadruplex can be classified as non-specific – where hydrated cations bind to the negatively charged phosphate backbone of DNA; or, as specific – where dehydrated cations bind in the central channel of G-quadruplex and coordinate with multiple guanine keto (O6) groups. Early studies using molecular dynamics stimulations have demonstrated that the coordination of cations at the centre of the quartet is of critical importance for the stability of G-quadruplex (80,81). An absence of coordinated cations at the electronegative central cavity significantly destabilizes the G quadruplex structure; the bound ions make a substantial contribution towards reducing electrostatic repulsions, thereby promoting the folding and stability of GQs. To date, most studies have focused on the roles of potassium or sodium cations on G-quadruplexes formation and stability, mainly because they are biologically relevant cations. However, a number of other monovalent as well as divalent cations have also been reported to play specific roles in G-quadruplex formation and stability. So far, the list of cations known to specifically favour GQ structures includes, in addition to Na^+ and K^+, the monovalent cations Rb^+, NH_4^+, and Tl^+(Cs^+ to a lesser degree); and the divalent cations Sr^{2+} and Pb^{2+} (Ba^{2+} to a lesser degree) (28). Studies with telomere-related sequences have shown that both K^+ and Sr^{2+} facilitate intermolecular G-quadruplex formation, with the divalent cation being the much more effective of the two (82). Thus, Sr^{2+}-induced intermolecular G-quadruplexes are thermodynamically more stable than K^+-induced quadruplexes. This has been rationalized in terms of a charge density for Sr^{2+}, relative to K^+, with enthalpic effects dominating (28,82). K^+ cations are known to coordinate between successive G-quartets, with near octahedral geometry to satisfy the hexacoordinate stereochemistry

(83). K^+ typically binds between every set of two adjacent quartets. On the other hand, Deng *et al.* (84) have demonstrated that Sr^{2+} ions are coordinated between every other pair of stacked quartets (i.e. they exclude one cavity space between adjacent quartets for every one site occupied). Every bound Sr^{2+} associates, like K^+, with eight carbonyl oxygen atoms from adjacent G-quartets (Figure 4.1).

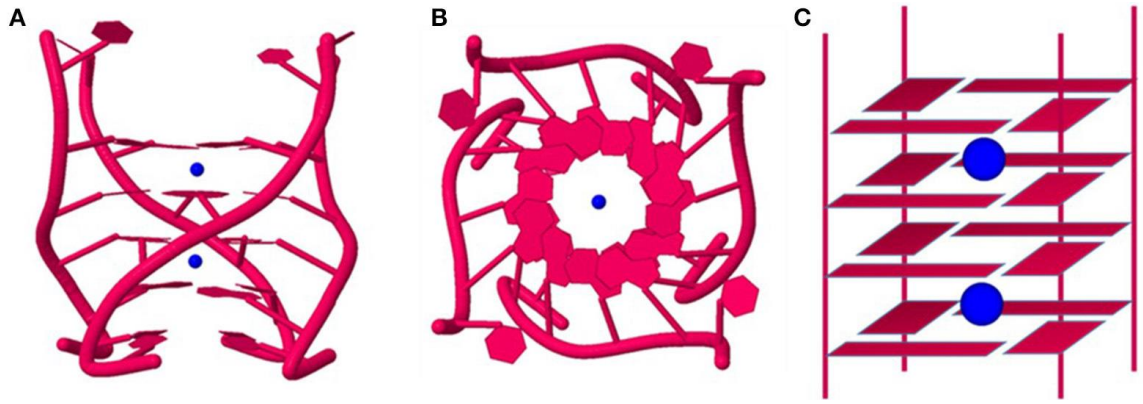


Figure 4.1. Crystal structure of RNA G-quadruplex formed by the sequence $(UG_4U)_4$ in Sr^{2+}.
The **(a)** side view and **(b)** top view of the G-quadruplex. The red ribbons represent the phosphate backbones, and the blue circles represent the Sr^{2+} ions. **(c)** A schematic illustration to show that each Sr^{2+} ion is sandwiched by every other two G-quartets. Adapted from (28).

In early 2000s, Wlodarczyk *et al.* (85) studied the effect of different metal ions on the formation and stability of G-quadruplex structures in the human telomeric sequence $d(TTAGGG)_4$. In their study, the melting curves of the 24-mer DNA sequence in the presence of K^+, Na^+, Rb^+, Li^+, and Sr^{2+} ions were recorded by circular dichroism (CD) spectroscopy and the melting temperatures were listed in Table 4.1. Based on the CD melting curves, at 100 mM of K^+, Na^+, Rb^+, and Li^+, the T_m was 59°, 50°, 40°, and 32 ℃, respectively. Even at lower salt concentrations (10 mM), the G-quadruplex was able to remain reasonably stable in K^+ (T_m = 44.3 ℃), and to a slightly lower extent in Rb^+ (T_m = 26.4 ℃). The effect of Sr^{2+} ions on the formation of G-quadruplex was the strongest among those of the ions studied. Their analysis of the saturation curves at 2 ℃ suggested that the presence of a single Sr^{2+} ion was sufficient to promote the formation of a monomeric quadruplex. The melting temperatures of the GQ in the presence of 0.1 mM and 10 mM $SrCl_2$ were 47.6° and 73.2 ℃, respectively. Together, the results of their study suggested the efficiency ordering of ions on the formation and stabilization of monomeric quadruplex structure of the $d(TTAGGG)_4$ sequence was: $Sr^{2+} > K^+ > Na^+ \geq$

$Rb^+ > Li^+$, which was in good agreement with the general trends suggested by a number of studies in the literature (28).

Table 4.1. Thermodynamic parameters of the quadruplex: random coil thermal transition of d(TTAGGG)$_4$.

salt	concn [mM]	λ_{max} [nm]	T_m [°C]	$\Delta H°$ [kcal/mol]	$\Delta S°$ [cal/(mol·K)]	$\Delta G°$ [a] [kcal/mol]
$SrCl_2$	0.1	301	47.6 ± 0.1	29.5 ± 0.2	92.0 ± 0.6	2.08 ± 0.02
	0.2	301	61.0 ± 0.1	27.7 ± 0.2	82.9 ± 0.6	2.98 ± 0.02
	10	301	73.2 ± 0.3	25.8 ± 0.2	74.6 ± 0.9	3.60 ± 0.05
NaCl	50	295	42.4 ± 0.2	38.8 ± 0.1	122.9 ± 0.4	2.14 ± 0.01
	100	295	50.2 ± 0.3	40.6 ± 0.1	125.5 ± 0.6	3.16 ± 0.01
	154	295	55.1 ± 0.3	42.3 ± 0.2	128.8 ± 0.6	3.87 ± 0.02
	200	295	57.7 ± 0.3	44.1 ± 0.2	133.4 ± 0.7	4.36 ± 0.02
	250	295	60.3 ± 0.3	45.4 ± 0.2	136.2 ± 0.8	4.81 ± 0.02
	303	295	62.2 ± 0.3	47.0 ± 0.2	140.2 ± 0.9	5.21 ± 0.03
KCl	10	290	44.3 ± 0.5	41.1 ± 0.3	129.5 ± 0.9	2.50 ± 0.02
	49	290	50.2 ± 0.2	46.5 ± 0.2	143.8 ± 0.6	3.63 ± 0.02
	104	290	59.0 ± 0.4	48.4 ± 0.3	145.8 ± 1.1	4.96 ± 0.04
	148	290	63.9 ± 0.4	45.7 ± 0.3	135.7 ± 1.1	5.28 ± 0.04
	190	290	66.4 ± 0.4	51.3 ± 0.3	151.1 ± 1.1	6.25 ± 0.04
	230	290	68.6 ± 0.4	51.5 ± 0.4	150.7 ± 1.3	6.57 ± 0.05
	268	290	70.0 ± 0.4	50.9 ± 0.4	148.5 ± 1.3	6.68 ± 0.05
	305	290	71.2 ± 0.4	53.2 ± 0.4	154.4 ± 1.3	7.13 ± 0.05
	340	290	72.2 ± 0.5	54.0 ± 0.4	156.3 ± 1.5	7.38 ±0.06
RbCl	10	295	26.4 ± 0.3	33.1 ± 0.4	110.3 ± 1.3	0.15 ± 0.01
	49	295	34.8 ± 0.1	37.5 ± 0.5	121.6 ± 1.9	1.20 ± 0.01
	90	295	38.5 ± 0.1	40.0 ± 0.5	127.8 ± 0.9	1.73 ± 0.01
	104	295	40.2 ± 0.1	39.8 ± 0.4	127.1 ± 1.7	1.94 ± 0.01
LiCl	10	301	23.0 ± 0.2	15.7 ± 0.3	52.8 ± 0.9	0.11 ± 0.01
	49	301	27.8 ± 0.2	26.9 ± 0.3	89.4 ± 1.0	0.25 ± 0.02
	90	301	31.2 ± 0.2	28.6 ± 0.3	93.8 ± 0.8	0.59 ± 0.01
	104	301	32.0 ± 0.2	28.6 ± 0.3	93.6 ± 0.8	0.66 ± 0.02

Adapted with permission from (85). Copyright 2005, American Chemical Society.

A second important feature for the formation kinetics and stability of a G-quadruplex is the number of G-quartets forming within it. The G-quartet is the fundamental structural unit within G-quadruplexes, with quartets stacking on top of one another, forming a G-stem. The number of guanines in each individual G-block (number of contiguous G's) is directly related to the number of G-quartets formed in the final folded quadruplex structure(86,87). Assuredly, a larger number of stacked G-quartets contributes to a greater stability of a GQ. Generally, in DNA, a stable G-quadruplex typically requires three or more layers of contiguous quartets. Based on previous studies, the association of a minimum of two quartets by $\pi - \pi$ stacking is required to form a stable G-quadruplex (88). Determining the minimum number of quartets required to form the triplex-quadruplex hybrid composite would allow us to gain insight on the stability of this "socket-plug" interaction.

4.2. Materials and methods

4.2.1. Assembly of DNA triplex monomers and dimers in Rb^+ and in Sr^{2+}

The oligonucleotides used for the "211" triplexes were purchased from Integrated DNA Technologies, Inc. (Coralville, Iowa, United States) and their sequences were listed in Table 1.1. All oligonucleotides were purified following the same protocol as described in Section 2.2.1.

Standard monomeric triplex ("A" and "C"), as well as "211" triplex tiles were assembled by mixing their respective component strands (each at 1 μM concentration) in the presence of 10 mM Li^+ acetate and 3 mM Mg^{2+} acetate at pH 5.2 (as described in Section 2.2.2). The DNA mixtures were heat-denatured at 100℃ for 4 minutes and gradually cooled down from 100℃ to 20℃ in a thermocycler at a constant cooling rate of 7.5℃/min.

For assembly of triplex-dimers, the pre-assembled constituent monomeric triplex tiles (each at 1 μM concentration) were mixed at equal volume, followed by the addition of either 50 mM RbCl or 1 mM $SrCl_2$ pre-dissolved in 10 mM Tris acetate(pH 5.2), following which the samples were allowed to rest at 22 ℃ for 18 hours.

4.2.2. Native gel and densitometry analyses

The formation of various DNA constructs was investigated using 7.5% (29:1) native polyacrylamide gel (containing and run in 50 mM Tris acetate and 5 mM Rb^+ or Sr^{2+} chloride pre-dissolved in 10 mM Tris acetate at pH 5.2. Prior to loading on gels, the DNA mixtures were mixed with loading buffer (containing 50 mM Tris acetate, 30% glycerol, bromophenol blue and xylene cyanol, at pH 5.2). The gels were run at 100 V at 22℃ with efficient cooling. Gels were exposed in 4℃ and scanned by a Typhoon 9410 Phosphorimager (Amersham Biosciences).

Following gel phosphorimagery, densitometry analyses were carried out to quantify triplex-dimer formation using the ImageQuant TL 8.1 software (Cytiva). The peak densities in each band in a lane were manually selected and quantified. The

percentages of triplex-dimer formation were calculated as the total intensity of the band corresponding to the dimer divided by the total intensity of DNA in the entire lane.

4.3. Triplex-dimer formation in Rb^+ and in Sr^{2+}

To explore the possibility of forming triplex-dimer in the presence of alternative metal cations known to be capable of stabilizing G-quadruplex structures, triplex A and C were incubated both alone and together in Rb^+, and in Sr^{2+}. Since the efficiency of Rb^+ in promoting G-quadruplex structure is often considered comparable to that of Na^+, we expected that Rb^+ at least should be able to promote the formation of homodimers, but not necessarily the heterodimer. As predicted, when the triplexes were incubated individually, in the presence of 50 mM Rb^+, A•A and C•C homodimers, respectively, were observed (Figure 4.2.a). However, to our surprise, the mixing of monomeric triplex A and C in Rb^+ generated the A•C heterodimer in addition to the A•A and C•C homodimers, with the heterodimer forming at approximately twice the level as the individual homodimers (see Table 4.2, below). Notably, in this case, while the heterodimer continues to be the major triplex-dimer product, it is no longer as strongly dominating as it is in K^+(Figure 2.11.e).

As suggested from the earlier work done by the Sen lab (30) as well as other reports in the literature (85), Sr^{2+} is a highly effective G-quadruplex-stabilizing ion whose efficiency is even higher than that of K^+. Therefore, we predicted that both homo- and hetero-dimers should be able to form even in the presence of low mM Sr^{2+} concentrations. Consistent with our expectation, 1 mM Sr^{2+} was sufficient to drive the formation of A•A, C•C and A•C dimers (Figure 4.2.b). The results were strikingly analogous to those found with K^+, i.e. when monomeric triplexes A and C were mixed, the A•C heterodimer was again the strongly favoured product relative to the homodimers (by a 6:1:1 ratio, Table 4.2).

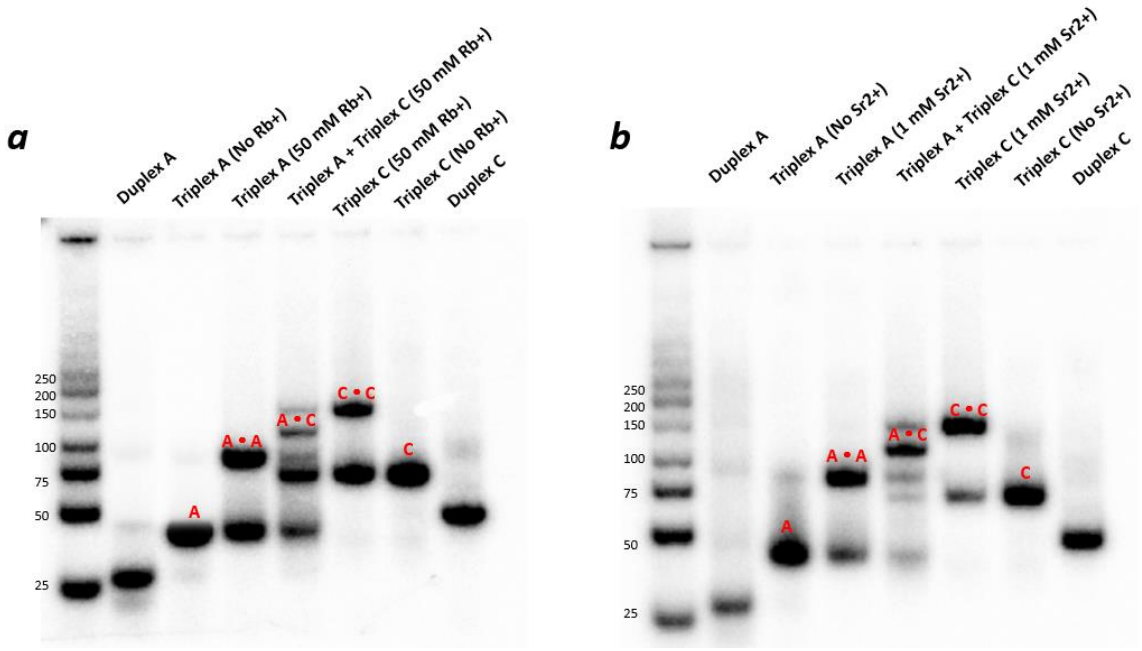


Figure 4.2. Triplex-dimer formation in Rb^+ versus in Sr^{2+}.
Triplex A and C were incubated separately and together for 18 hours in either **(a)** 50 mM RbCl in 10 mM Tris acetate (pH 5.2) and loaded on a native gel containing and running at 50 mM Tris acetate (pH 5.2) and 5 mM Rb^+; or **(b)** 1 mM $SrCl_2$ in 10 mM Tris acetate (pH 5.2) and run on a native gel containing 50 mM Tris acetate (pH 5.2) and 0.1 mM Sr^{2+}.

To compare the efficiency of various metal cations in promoting G-quadruplex structures, and hence triplex dimerization, the percentage yields of all three triplex-dimers in the presence of all the ions studied (K^+, Na^+, Rb^+, and Sr^{2+}) were compared (They are listed in Table 4.2). While all the cations listed promote self-dimerization of triplex A alone and triplex C alone, hetero-dimerization of A and C can only be observed in K^+, Rb^+, and Sr^{2+}, but not in Na^+. This result suggests that Na^+ is insufficient to facilitate the formation of G-quadruplex involved in the A•C heterodimer. One possible factor for the absence of heterodimer formation in Na^+ may be the size of the cation, – a major factor that contributes to both the topology and stability of the G-quadruplex structure formed. The ionic radius of Na^+ (0.95 Å) is considerably smaller than that of K^+ (1.33 Å), Rb^+ (1.48 Å), and Sr^{2+} (1.13 Å) (28). Most interestingly, although Rb^+, does promote A•C heterodimer formation, the ratio between the hetero- and homo-dimers formed from mixtures of A and C tiles is approximately 2:1, significantly lower than the 6:1 ratios formed in either K^+ or Sr^{2+}. Notably, the 1:2:1 ratio of A•A, A•C, and C•C formed suggests a simple binomial relationship; in other words, that Rb^+ supports homodimer and heterodimer formation *comparably*. By contrast, in both K^+ and Sr^{2+}, the

formation of heterodimer is strongly favoured over the homodimers, with heterodimer being approximately 6-fold higher at equilibrium than that of the homodimers (Table 4.2).

Table 4.2. Percentage yields of triplex-dimers in different cations.

Triplex-dimer	A•A	A•C	C•C	Approximate Ratio
50 mM K^+	12. 0 ± 1.0 %	61.6 ± 0.4 %	11.8 ±0.3 %	1 : 6 : 1
50 mM Na^+	17.0 ± 1.0 %	N/A	18.9 ± 0.6 %	1 : 1
50 mM Rb^+	15.6 ± 1.0 %	25.2 ± 0.6 %	12.7 ± 0.7 %	1 : 2 : 1
1 mM Sr^{2+}	10.0 ± 0.3 %	58. 0 ± 0.9 %	10.0 ± 0.8 %	1 : 6 : 1

The percentage yields are calculated based on densitometry data obtained from native gel analyses. Each data point is the averaged value obtained from at least three replicate experiments.

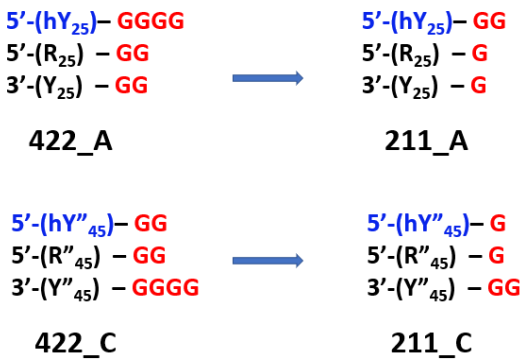
4.4. "211" triplexes: is it possible to dimerize triplex tiles by forming a two-quartet G-quadruplex?

The triplexes that we originally designed can be described as "422 triplexes", since 4, 2, and 2 are the number of non-YR•Y triplex forming guanines added to the three constituent strands of a triplex tile at one end (the 3' end of the R strand). In the arrangement explored this far in the thesis, two of the three strands contain two unmatched guanines at the end, while the third strand contains four. But, is this 422 arrangement, leading to four complete G-quadruplexes in the tile homodimers, strictly necessary? To investigate if a smaller number of G-quartets may be sufficient to stably link two individual triplex tiles, we reduced the number of the mismatched guanine residues at the 3' end of the monomeric triplex from "422" to "211" (Figure 4.3.a).

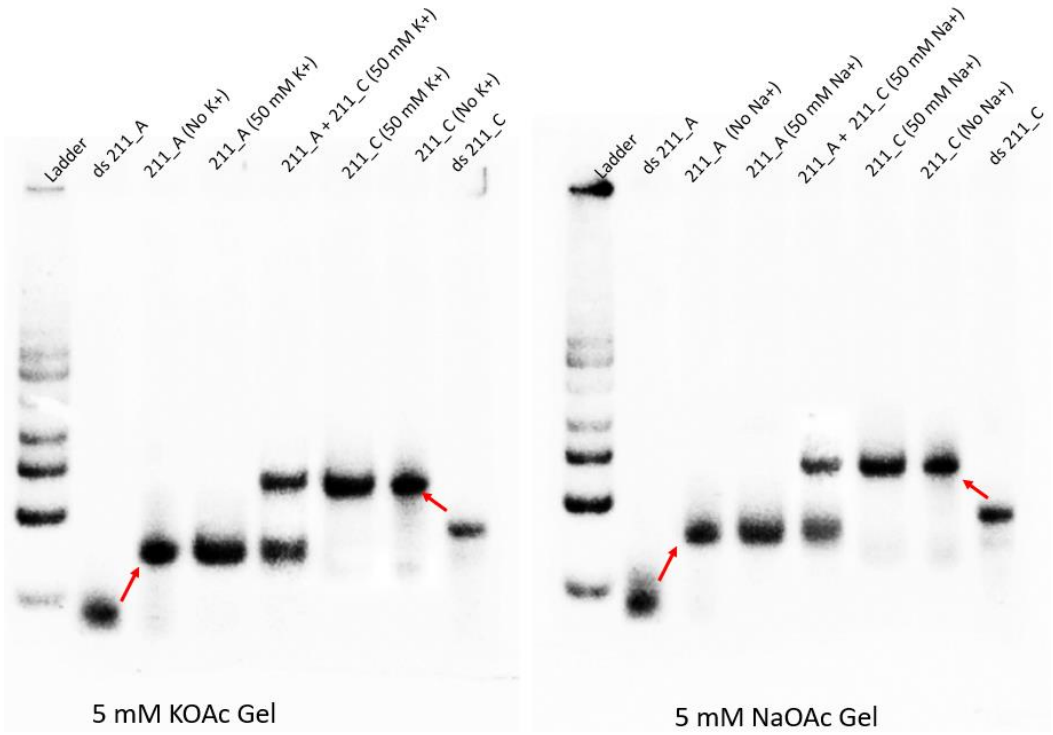
As before, we assembled the monomeric "211" triplex tiles A and C (called "211A" and "211C") in the presence of either 50 mM KOAc or NaOAc, at pH 5.2, under the standard conditions. Following an 18-hour incubation at 22°C, the DNA solutions were run on 7.5% native gels containing the respective counterions to investigate the formation of triplex-dimers by these modified tiles. Figure 4.3.b shows that while the monomeric triplex tiles were nicely formed, as evidenced by the characteristic upward shift of bands relative to the respective parent duplexes, neither the addition of K^+ nor Na^+ gave rise to any triplex-dimer formation. Both 211A and 211C remained as triplex monomers. In addition, the mixing of triplexes 211A and 211C in the presence of K^+ did

not lead to the formation of a putative 211A•211C heterodimer, which in the case of the earlier (422) A and C was strongly favoured in K^+. We considered the that reduction in the number of quartets would make for less stable GQ structures; therefore, given that Sr^{2+} is generally a stronger GQ-stabilizer than K^+ is, we incubated the "211" triplexes under two different Sr^{2+} concentration regimes (1 and 50 mM) to see whether the powerfully stabilizing Sr^{2+}, especially at a high concentration (50 mM), did enable the formation of triplex-dimers. Figure 4.3.c shows that, even with incubation at that high Sr^{2+} concentration (50 mM), neither homo- nor hetero-dimerization of 211 triplex tiles was observed. Together, these results suggest that a two-quartet joining quadruplex does not form with sufficient speed and/or stability to enable the triplex-dimerization.

(a)



(b)



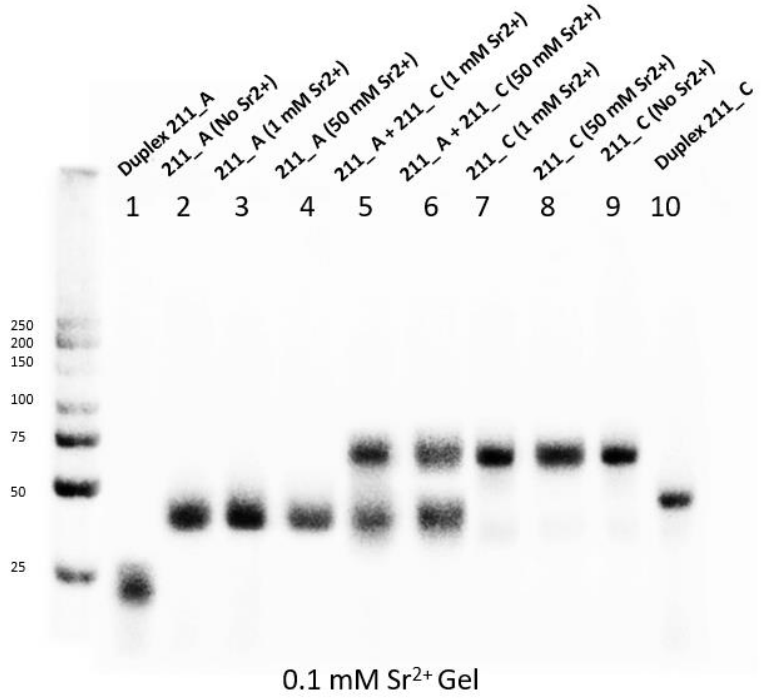


Figure 4.3. The design of "211" triplex tiles and native gel analyses.
(a) The numbers of guanine overhangs in the original triplex A and C are reduced from "422" to "211". **(b)** Various DNA constructs built using the "211" triplexes under different salt conditions are run and analyzed on different 7.5% native gels at pH 5.2.

Chapter 5. Conclusions and future directions

In this project, we carried out an extensive investigation of a unique, “socket-plug” type of DNA-DNA interaction, or recognition paradigm, within a triplex-quadruplex hybrid system. Utilizing our ingeniously designed, purely guanine-based sticky ends, we created a set of triplex-based DNA composites that manifest self- and other-complementary recognition properties, which in turn are different in the presence of different counter-cations. The striking observation that distinct TQ hybrid products are yielded in the presence of different counterions suggests that the interaction between the “socket-plug” pairs is highly ion-specific. In other words, the recognition and binding preferences are strongly subject to change based on the choice of the counterions. Such a sensitivity to specific counterions is a highly useful property (and unprecedented, since it is not found in conventional, Watson-Crick rules-based hybridization), which should allow these TQ hybrids to be used not only for nanotechnology but also in analytical chemistry and many other fields. Depending on the need, one can with ease design and assemble various TQ hybrids by mixing-and-matching different triplex tiles available in our toolbox, in a robust manner. Another key feature of our triplex-quadruplex hybrid system that is worth exploiting is pH dependence. The fact that one of the cytosines in every triple-base unit in the parallel triplex requires an acidic environment to remain protonated allows us to easily reverse the formation of TQ hybrids by altering the pH. Such a pH-dependence property presents the potential of these TQ hybrids, or nanostructures constructed using these hybrids to be developed as pH sensors.

Of the three monomeric DNA tiles examined by us, triplexes A and C, show the almost exclusive property of formation of heterodimer in the presence of K^+ versus homodimers in Na^+. Utilizing such unique property, we have shown that the mixing of a doubly sticky-ended triplex A (“AA”) with single-headed triplex C in K^+ can give rise to a continuous long stretch of TQ hybrid DNA wire on a nanometer scale, which is completely absent in the presence of Na^+. More importantly, having one triplex tile in excess over the other also leads to drastically different yet wholly predictable outcomes, especially in the presence of K^+. Therefore, one could potentially assemble wires of finite and pre-determined length by carefully fine-tuning the concentration ratios between the two participating tiles.

In terms of the possible biological applications of this unique TQ hybrid system, one could potentially develop an ion-sensing nanodevice by exploiting the sensitivity of A•C heterodimer formation to the presence of K^+. For instance, recently, Pirovano *et al.* (89) have developed a wearable sensor for sampling and measuring the concentration of Na^+ and K^+ in human sweat during physical exercise. The fact that the A•C heterodimer was able to stably form with as little as 1 mM K^+ demonstrates its high sensitivity to K^+, which should be a strong asset that one can bring to the development of K^+-sensing devices. Furthermore, pH-dependent reversibility is another important feature of our A•C TQ hybrid that could potentially find wide applications in DNA nanotechnology. Previously, the Krishnan group (90) has developed a DNA nanomachine called the I-switch, which functions as a pH sensor that can efficiently map the spatial and temporal pH changes inside living cells with a high dynamic range between pH 5.8 and 7. Since the A•C TQ hybrid system is built based on parallel triplexes, whose formation and stability are largely dependent on acidic pH, it can also be utilized, in conjunction with fluorophore dyes, as a pH reporter based on fluorescence resonance energy transfer (FRET).

Given their unique properties and high potential for applications in different fields, it is important to continue characterization and discovery on further hybridization possibilities of this TQ hybrid system. One of the many possible future directions is to investigate the orientation of the guanine-based sticky ends. In our single-sided sticky-ended triplex design, the guanine sticky ends are appended to the 3' ends of the triplex tiles (using the R or polypurine component strand as the reference strand). It is worthwhile to test whether there might be any difference if the sticky ends were placed at a tile's 5' end instead. To answer this question, I carried out preliminary experiments, which show results that are both interesting and unexpected. Figure 5.1.a shows our design of a triplex tile, A': while it differs from the original triplex A both in the sequence of its YR•Y component and its overall length, it does contain the same type of 422 guanine-based sticky end (in both cases, the Hoogsteen strands contain extra two Gs relative to the other two component strands). However, crucially, in triplex A', the sticky end is located at the tile's 5' end instead of the 3', as is the case with A. Likewise, triplex C' also has its sticky end located at its 5' end. To compare the similarities and differences in dimerization between having 5'- versus 3'- sticky ends, these four triplex tiles (A; C; A' and C') were incubated individually or in pairs in the presence of 50 mM

KOAc or NaOAc for 18 hours before analysis on native gels. While the 3' sticky-ended monomeric triplexes (triplexes A and C) were able to self-dimerize in both K^+ and Na^+ as reported consistently, above, the 5' sticky-ended triplexes (triplexes A' and C') did not dimerize at all, irrespective of the salt condition (Figure 5.1.b). This result suggests that the orientation of the sticky end is critical to the dimerization of 422 triplex tiles. The reason and logic behind this are yet to be elucidated. We believe that deeper investigations on the different hybridization properties between the two different guanine sticky end orientations should provide us with greater insights into the fundamental properties of this kind of non-canonical DNA secondary structures.

Another interesting area for further exploration is on varying the G-quartet layers, by changing the numbers of "sticky" guanine residues at the end of each component strand within 3'-sticky ended triplex tiles. In this study, we focused on 422 (and, to some extent, 211) TQ tiles, dimerizing with self or other to form "joining" G-quadruplexes consisting of 4 G-quartets. While we have shown that having 2-quartet G-quadruplexes is insufficient to dimerize the individual triplex tiles, the possibilities of 3-quartet quadruplexes has not been explored. In a 3-quartets setting, there are many different possible combinations of Gs that one could potentially try out, such as "322", "321", "311", and so on.

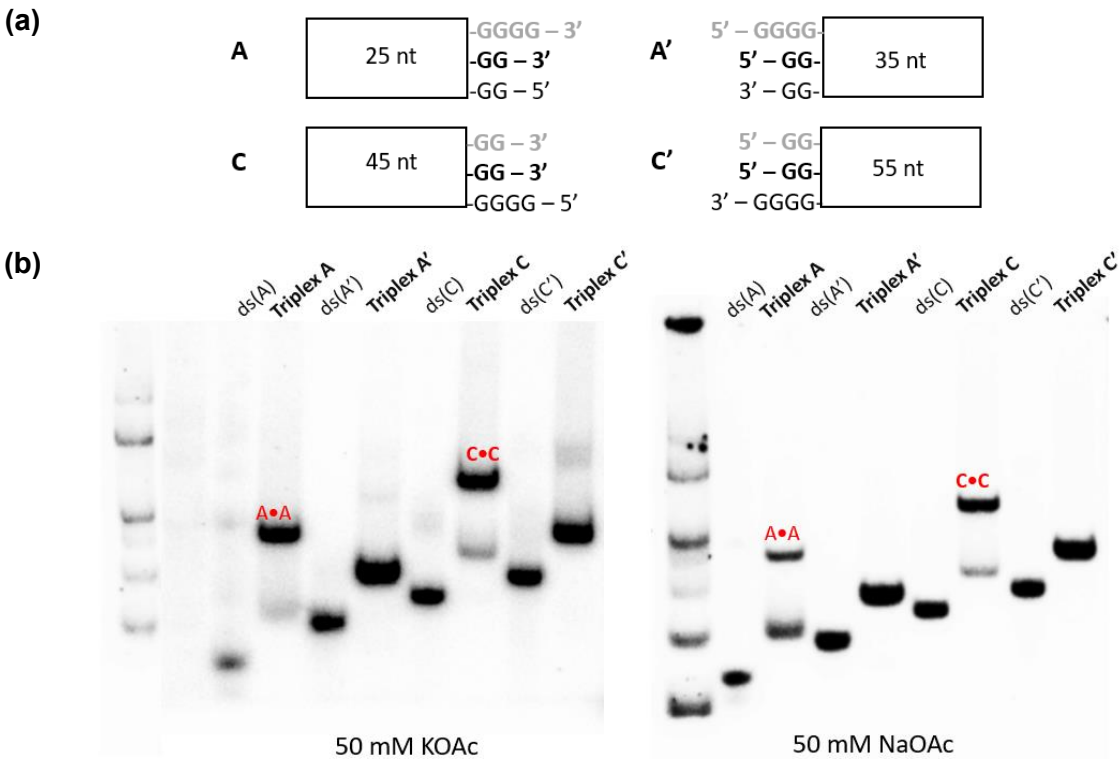
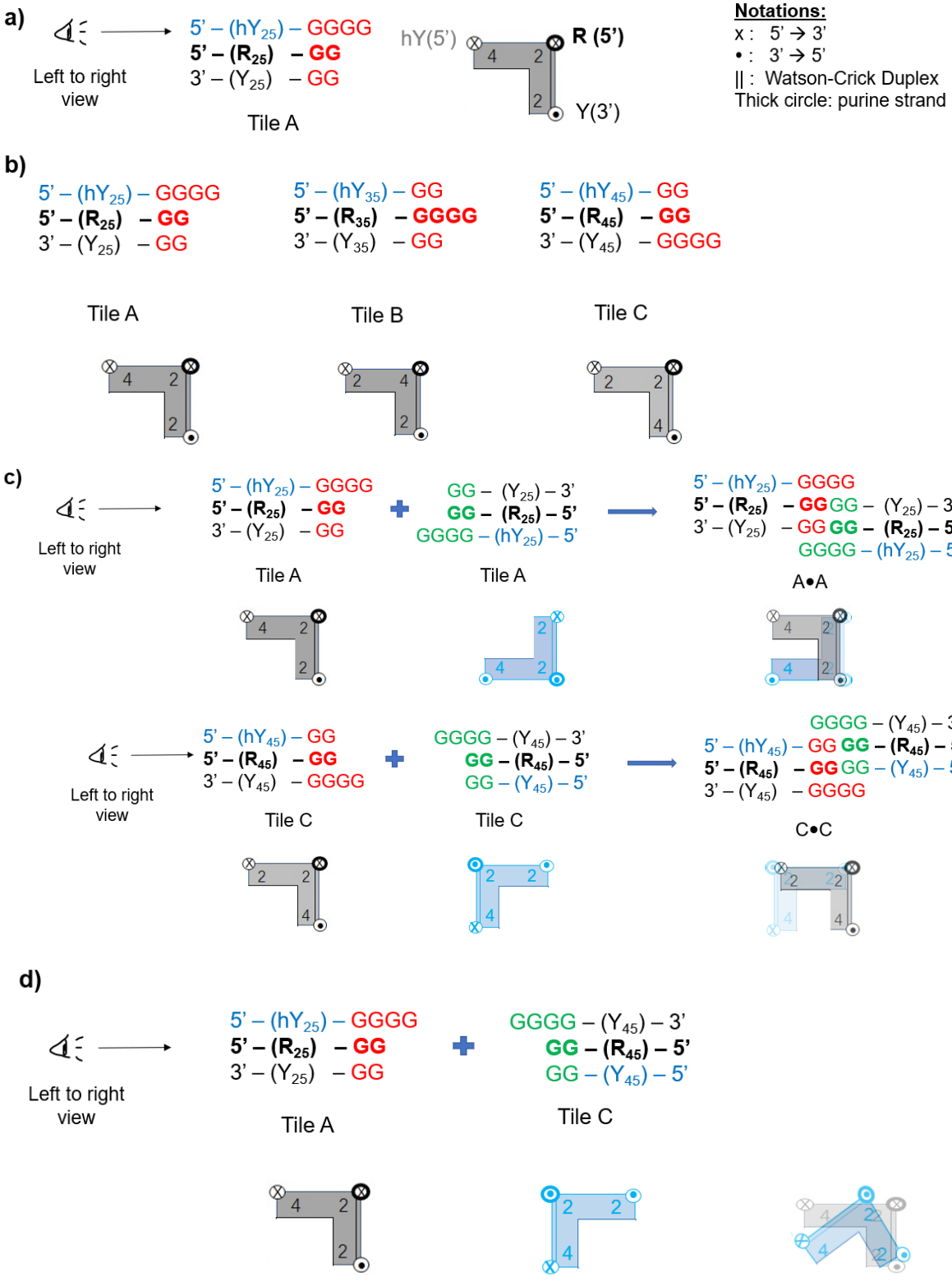


Figure 5.1. Design of 5' "sticky-ended" triplex A' and C' and investigation on their abilities of dimerization.

(a) Schematic drawings demonstrating the design of 5' "sticky-ended" triplex A' and C' in comparison to the original 3' "sticky ended" triplex A and C. **(b)** Native gel analyses to test the ability of triplex A' and C' to dimerize in the presence of 50 mM KOAc (left) or NaOAc (right).

While the mysterious 3-dimentional structure of A•C is yet to be elucidated, we created an end-on view version of schematics to represent the ends of triplex tiles, in an attempt to help us visualize what might be happening at point of junction that allows the formation of a G-quadruplex, and hopefully, find clues to solve the A•C puzzle. Figure 5.2.a shows the design of our end-on view model. In particular, the boomerang shape represents the triplex structure, each circle at the corner represents one triplex component strand, with the purine strand in bold, and connected to its pyrimidine-rich W-C counterpart by two straight lines. The orientation of the strands is represented by either a "X" or a dot, meaning either 5' or 3', respectively. The numbers indicate how many guanine residues are associated with each strand. Using this model, we represented each of our triplex tiles (Figure 5.2.b) and tried to visualize the dimerization reaction by rotating the boomerangs and superimposing one on top of the other. For example, in Figure 5.2.c, when two A triplex monomers are superimposed tail-to-tail, the numbers at each corner add up to 4, indicating the formation of a 4-quartet G-

quadruplex. Likewise, two C triplex monomers can also be arranged and superimposed in a certain way that leads to the formation of 4 complete G-quartets. Notice that in the superimposed images of both A•A and C•C homodimers, one arm of a boomerang (black) will perfectly align with an arm from the second boomerang (blue), meaning that two of the three component strands within a triplex monomer will align with two strands from the other triplex monomer. However, when we used this schematic model to represent A•C heterodimer (Figure 5.2.d), we found that one plausible way to generate complete 4 planar quartet could be to have one triplex slightly tilted away from the other. This way, the numbers at each of the four corners can also be added up to 4, indicating the formation of a complete 4-layered G-quadruplex, which is in agreement with our DMS foot-printing experiment that suggests all the guanine residues in A•C heterodimer are involved in quartet formation. Although it might appear that the planar quartet in this case is distorted, such minor distortion should be easily tolerated by the strong tendency of the G-rich sticky ends to form a G-quadruplex, at least in the presence of potassium ion. One unique feature about A•C heterodimer is that the protruding strands from triplexes A and C are in parallel orientation; on the other hand, in both A•A and C•C homodimers, these two strands are antiparallel to each other. We hypothesize that, these two protruding strands can first rapidly align to form a parallel G dimer (Figure 5.2.e), analogous to the model proposed by Sen and Gilbert (24). Given the strong tendency for G-rich sequences to form G-quadruplex structures, once the dimer forms, it will strongly drive the formation of a G-quadruplex by quickly recruiting the remaining strands to complete the quartets. Strong G4-stablizing agents such as potassium, will be able to promote such rapid conversion from a dimer to a stable quadruplex, which explains why A•C heterodimer is strongly favoured product in potassium, both kinetically and thermodynamically. Sodium, on the other hand, is not sufficient to support the formation of such a structurally less robust G-quadruplex.


a)
Left to right view
5' – (hY25) – GGGG
5' – (R25) – GG
3' – (Y25) – GG
Tile A
hY(5')
R (5')
Y(3')
Notations:
x : 5' → 3'
• : 3' → 5'
|| : Watson-Crick Duplex
Thick circle: purine strand
b)
5' – (hY25) – GGGG
5' – (R25) – GG
3' – (Y25) – GG
5' – (hY35) – GG
5' – (R35) – GGGG
3' – (Y35) – GG
5' – (hY45) – GG
5' – (R45) – GG
3' – (Y45) – GGGG
Tile A
Tile B
Tile C
c)
Left to right view
5' – (hY25) – GGGG
5' – (R25) – GG
3' – (Y25) – GG
Tile A
GG – (Y25) – 3'
GG – (R25) – 5'
GGGG – (hY25) – 5'
Tile A
5' – (hY25) – GGGG
5' – (R25) – GGGG – (Y25) – 3'
3' – (Y25) – GG GG – (R25) – 5'
GGGG – (hY25) – 5'
A•A
Left to right view
5' – (hY45) – GG
5' – (R45) – GG
3' – (Y45) – GGGG
Tile C
GGGG – (Y45) – 3'
GG – (R45) – 5'
GG – (Y45) – 5'
Tile C
GGGG – (Y45) – 3'
5' – (hY45) – GG GG – (R45) – 5'
5' – (R45) – GG GG – (Y45) – 5'
3' – (Y45) – GGGG
C•C
d)
Left to right view
5' – (hY25) – GGGG
5' – (R25) – GG
3' – (Y25) – GG
Tile A
GGGG – (Y45) – 3'
GG – (R45) – 5'
GG – (Y45) – 5'
Tile C

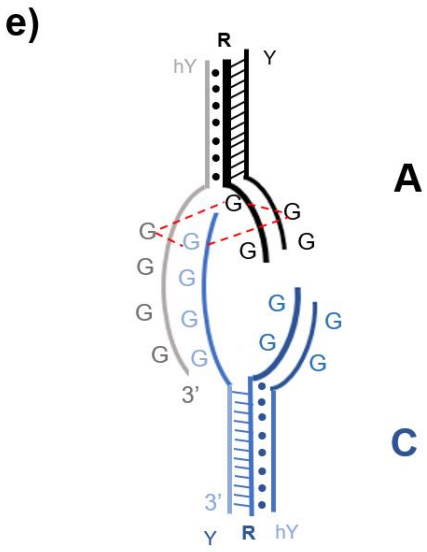


Figure 5.2. End-on view model of triplex dimers.
(a) Notations of this model: the boomerang represents a triplex; the circles represent triplex component strands; the orientation of strands are indicated by an "X" (5') or a dot (3'); the numbers indicate the overhanging guanine residues at the end of a strand. The purine-rich strand in the W-C duplex is highlighted in bold. **(b)** Representations of triplex A, B, and C using this end-on scheme. **(c)** Homo-dimerization of triplexes A and C, respectively. (d) Heterodimerization of triplexes A and C. **(d)** Proposed model of A•C heterodimer.

To build on the groundwork established by this project, the abovementioned possible future research directions should help us define more clearly and comprehensively the rules that operate for this novel type of "socket-plug" DNA complementarity. Ultimately, we hope to compose a "rulebook" which provides guidance on using these innovative TQ hybrids as handy and powerful tools for the construction of various nanocomposites that can be widely applicable in different fields.

References

1. Dahm R. Friedrich Miescher and the discovery of DNA. Vol. 278, Developmental Biology. Academic Press Inc.; 2005. p. 274–88.

2. Avery OT, MacLeod CM, McCarty M. STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES : INDUCTION OF TRANSFORMATION BY A DESOXYRIBONUCLEIC ACID FRACTION ISOLATED FROM PNEUMOCOCCUS TYPE III. Journal of Experimental Medicine [Internet]. 1944 Feb 1;79(2):137–58. Available from: https://doi.org/10.1084/jem.79.2.137

3. Chargaff E, York N. Chemical Specificity of Nucleic Acids and Mechanism of their Enzymatic Degradation 1. EXPERIENTIA. 1950.

4. WATSON JD, CRICK FHC. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. Nature [Internet]. 1953;171(4356):737–8. Available from: https://doi.org/10.1038/171737a0

5. Franklin RE, Goslino XRG. The Structure of Sodium Thymonucleate Fibres. I. The Influence of Water Content. Vol. 6, Acta Cryst. 1953.

6. Xu Y, McSally J, Andricioaei I, Al-Hashimi HM. Modulation of Hoogsteen dynamics on DNA recognition. Nature Communications [Internet]. 2018;9(1):1473. Available from: https://doi.org/10.1038/s41467-018-03516-1

7. Whelan DR, Hiscox TJ, Rood JI, Bambery KR, McNaughton D, Wood BR. Detection of an en masse and reversible B-to A-DNA conformational transition in prokaryotes in response to desiccation. Journal of the Royal Society Interface. 2014;11(97):20140454.

8. Lu X-J, Shakked Z, Olson WK. A-form Conformational Motifs in Ligand-bound DNA Structures. Journal of Molecular Biology [Internet]. 2000;300(4):819–40. Available from: https://www.sciencedirect.com/science/article/pii/S0022283600936908

9. Timsit Y. DNA structure and polymerase fidelity11Edited by T. Richmond. Journal of Molecular Biology [Internet]. 1999;293(4):835–53. Available from: https://www.sciencedirect.com/science/article/pii/S0022283699931996

10. Pohl FM, Jovin TM. Salt-induced co-operative conformational change of a synthetic DNA: Equilibrium and kinetic studies with poly(dG-dC). Journal of Molecular Biology [Internet]. 1972;67(3):375–96. Available from: https://www.sciencedirect.com/science/article/pii/0022283672904573

11. Herbert A. Z-DNA and Z-RNA in human disease. Communications Biology [Internet]. 2019;2(1):7. Available from: https://doi.org/10.1038/s42003-018-0237-x

12. Heinemann U, Roske Y. Symmetry in nucleic-acid double helices. Vol. 12, Symmetry. MDPI AG; 2020.

13. Hoogsteen K. The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. Acta Crystallographica [Internet]. 1959 Oct 1;12(10):822–3. Available from: https://doi.org/10.1107/S0365110X59002389

14. Courtois Y, Fromageot P, Guschlbauer W. Protonated Polynucleotide Structures. European Journal of Biochemistry [Internet]. 1968 Dec 1;6(4):493–501. Available from: https://doi.org/10.1111/j.1432-1033.1968.tb00472.x

15. Nikolova EN, Zhou H, Gottardo FL, Alvey HS, Kimsey IJ, Al-Hashimi HM. A historical account of hoogsteen base-pairs in duplex DNA. Vol. 99, Biopolymers. 2013. p. 955–68.

16. Ghosal G, Muniyappa K. Hoogsteen base-pairing revisited: Resolving a role in normal biological processes and human diseases. Vol. 343, Biochemical and Biophysical Research Communications. 2006. p. 1–7.

17. Felsenfeld G, Davies DR, Rich A. FORMATION OF A THREE-STRANDED POLYNUCLEOTIDE MOLECULE. J Am Chem Soc [Internet]. 1957 Apr 1;79(8):2023–4. Available from: https://doi.org/10.1021/ja01565a074

18. Asensio JL, Brown T, Lane AN. Solution conformation of a parallel DNA triple helix with 5′ and 3′ triplex–duplex junctions. Structure [Internet]. 1999;7(1):1–11.

Available from:
https://www.sciencedirect.com/science/article/pii/S0969212699800045

19. Hu Y, Cecconello A, Idili A, Ricci F, Willner I. Triplex-DNA-Nanostrukturen: von grundlegenden Eigenschaften zu Anwendungen. Angewandte Chemie. 2017 Nov 27;129(48):15410–34.

20. Bacolla A, Wang G, Vasquez KM. New Perspectives on DNA and RNA Triplexes As Effectors of Biological Activity. PLOS Genetics [Internet]. 2015 Dec 23;11(12):e1005696-. Available from: https://doi.org/10.1371/journal.pgen.1005696

21. Frank-Kamenetskii MD, Mirkin SM. TRIPLEX DNA STRUCTURES. Annual Review of Biochemistry [Internet]. 1995 Jun 1;64(1):65–95. Available from: https://doi.org/10.1146/annurev.bi.64.070195.000433

22. GELLERT M, LIPSETT MN, DAVIES DR. Helix formation by guanylic acid. Proc Natl Acad Sci U S A [Internet]. 1962 Dec 15;48(12):2013–8. Available from: https://pubmed.ncbi.nlm.nih.gov/13947099

23. Sen D, Gilbert W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature [Internet]. 1988;334(6180):364–6. Available from: https://doi.org/10.1038/334364a0

24. Sen D, Gilbert W. A sodium-potassium switch in the formation of four-stranded G4-DNA. Nature. 1990;344(6265):410–4.

25. Sundquist WI, Klug A. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. nature. 1989;342(6251):825–9.

26. Williamson JR, Raghuraman MK, Cech TR. Monovalent cation-induced structure of telomeric DNA: the G-quartet model. Cell. 1989;59(5):871–80.

27. Burge S, Parkinson GN, Hazel P, Todd AK, Neidle S. Quadruplex DNA: Sequence, topology and structure. Nucleic Acids Research. 2006 Nov;34(19):5402–15.

28. Bhattacharyya D, Mirihana Arachchilage G, Basu S. Metal Cations in G-Quadruplex Folding and Stability. Frontiers in Chemistry [Internet]. 2016;4:38. Available from: https://www.frontiersin.org/article/10.3389/fchem.2016.00038

29. Subramanian S, Kaufman BT, Thompson ST, Stellwagen E, Cass KH, Wierenga RK, et al. Cation-Dependent Transition between the Quadruplex and Watson-Crick Hairpin Forms of d(CGCG3GCG)t [Internet]. Vol. 31, Proc. Natl. Acad. Sci. U.S.A. 1992. Available from: https://pubs.acs.org/sharingguidelines

30. Venczel EA, Sen D. Parallel and Antiparallel G-DNA Structures from a Complex Telomeric Sequence* [Internet]. Vol. 32, Biochemistry. 1993. Available from: https://pubs.acs.org/sharingguidelines

31. Sigel A, Sigel H, Sigel RKO. Metal Ions in Life Sciences 16 The Alkali Metal Ions: Their Role for Life [Internet]. Available from: http://www.springer.com/series/8385

32. Gu J, Leszczynski J. Origin of Na+/K+ Selectivity of the Guanine Tetraplexes in Water: The Theoretical Rationale. The Journal of Physical Chemistry A [Internet]. 2002 Jan 1;106(3):529–32. Available from: https://doi.org/10.1021/jp012739g

33. Hud N v, Smith FW, Anet FAL, Feigon J. The Selectivity for K+ versus Na+ in DNA Quadruplexes Is Dominated by Relative Free Energies of Hydration: A Thermodynamic Analysis by 1H NMR. Biochemistry [Internet]. 1996 Jan 1;35(48):15383–90. Available from: https://doi.org/10.1021/bi9620565

34. Cotton-Wilkinson - Advanced Inorganic Chemistry_file1.

35. Malgowska M, Czajczynska K, Gudanis D, Tworak A, Gdaniec Z. Overview of the RNA G-quadruplex structures. Vol. 63, Acta Biochimica Polonica. Polskie Towarzystwo Biochemiczne; 2016. p. 609–21.

36. Rhodes D, Lipps HJ. G-quadruplexes and their regulatory roles in biology. Nucleic Acids Research [Internet]. 2015 Oct 15;43(18):8627–37. Available from: https://doi.org/10.1093/nar/gkv862

37. Huppert JL, Balasubramanian S. Prevalence of quadruplexes in the human genome. Nucleic Acids Res [Internet]. 2005 May 24;33(9):2908–16. Available from: https://pubmed.ncbi.nlm.nih.gov/15914667

38. Kudlicki AS. G-Quadruplexes Involving Both Strands of Genomic DNA Are Highly Abundant and Colocalize with Functional Sites in the Human Genome. PLoS One [Internet]. 2016 Jan 4;11(1):e0146174–e0146174. Available from: https://pubmed.ncbi.nlm.nih.gov/26727593

39. Varizhuk A, Ischenko D, Tsvetkov V, Novikov R, Kulemin N, Kaluzhny D, et al. The expanding repertoire of G4 DNA structures. Biochimie [Internet]. 2017;135:54–62. Available from: https://www.sciencedirect.com/science/article/pii/S0300908416303182

40. Bedrat A, Lacroix L, Mergny J-L. Re-evaluation of G-quadruplex propensity with G4Hunter. Nucleic Acids Res [Internet]. 2016/01/20. 2016 Feb 29;44(4):1746–59. Available from: https://pubmed.ncbi.nlm.nih.gov/26792894

41. Puig Lombardi E, Londoño-Vallejo A. A guide to computational methods for G-quadruplex prediction. Nucleic Acids Research [Internet]. 2020 Jan 10;48(1):1–15. Available from: https://doi.org/10.1093/nar/gkz1097

42. Marsico G, Chambers VS, Sahakyan AB, McCauley P, Boutell JM, Antonio M di, et al. Whole genome experimental maps of DNA G-quadruplexes in multiple species. Nucleic Acids Research [Internet]. 2019 May 7;47(8):3862–74. Available from: https://doi.org/10.1093/nar/gkz179

43. Ambrus A, Chen D, Dai J, Bialis T, Jones RA, Yang D. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. Nucleic Acids Res [Internet]. 2006 May 19;34(9):2723–35. Available from: https://pubmed.ncbi.nlm.nih.gov/16714449

44. McElligott R, Wellinger RJ. The terminal DNA structure of mammalian chromosomes. EMBO J [Internet]. 1997 Jun 16;16(12):3705–14. Available from: https://pubmed.ncbi.nlm.nih.gov/9218811

45. Healy KC. Telomere dynamics and telomerase activation in tumor progression: prospects for prognosis and therapy. Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics. 1995;7(3–4):121–30.

46. Lat PK, Liu K, Kumar DN, Wong KKL, Verheyen EM, Sen D. High specificity and tight spatial restriction of self-biotinylation by DNA and RNA G-Quadruplexes complexed in vitro and in vivo with Heme. Nucleic Acids Research [Internet]. 2020 Jun 4;48(10):5254–67. Available from: https://doi.org/10.1093/nar/gkaa281

47. Summers PA, Lewis BW, Gonzalez-Garcia J, Porreca RM, Lim AHM, Cadinu P, et al. Visualising G-quadruplex DNA dynamics in live cells by fluorescence lifetime imaging microscopy. Nature Communications [Internet]. 2021;12(1):162. Available from: https://doi.org/10.1038/s41467-020-20414-7

48. Alexis F, Rhee J-W, Richie JP, Radovic-Moreno AF, Langer R, Farokhzad OC. New frontiers in nanotechnology for cancer treatment. Urologic Oncology: Seminars and Original Investigations [Internet]. 2008;26(1):74–85. Available from: https://www.sciencedirect.com/science/article/pii/S1078143907000920

49. Beissenhirtz MK, Willner I. DNA-based machines. Org Biomol Chem. 2006;4(18):3392–401.

50. Liu X, Lu C-H, Willner I. Switchable reconfiguration of nucleic acid nanostructures by stimuli-responsive DNA machines. Acc Chem Res. 2014;47(6):1673–80.

51. Seeman NC. From genes to machines: DNA nanomechanical devices. Trends Biochem Sci. 2005;30(3):119–25.

52. Seeman NC. Nucleic acid junctions and lattices. Journal of Theoretical Biology [Internet]. 1982;99(2):237–47. Available from: https://www.sciencedirect.com/science/article/pii/0022519382900029

53. Fu T-J, Seeman NC. DNA Double-Crossover Molecules1" [Internet]. Vol. 32, Biochemistry. 1993. Available from: https://pubs.acs.org/sharingguidelines

54. Li X, Yang X, Qi J, Seeman NC. Antiparallel DNA Double Crossover Molecules As Components for Nanoconstruction. J Am Chem Soc [Internet]. 1996 Jan 1;118(26):6131–40. Available from: https://doi.org/10.1021/ja960162o

55. Rothemund PWK. Folding DNA to create nanoscale shapes and patterns. Nature [Internet]. 2006;440(7082):297–302. Available from: https://doi.org/10.1038/nature04586

56. Seeman NC, Sleiman HF. DNA nanotechnology. Vol. 3, Nature Reviews Materials. Nature Publishing Group; 2017.

57. Mergny JL, Sen D. DNA quadruple helices in nanotechnology. Vol. 119, Chemical Reviews. American Chemical Society; 2019. p. 6290–325.

58. Chandrasekaran AR, Rusling DA. Triplex-forming oligonucleotides: a third strand for DNA nanotechnology. Nucleic Acids Research [Internet]. 2018 Feb 16;46(3):1021–37. Available from: https://doi.org/10.1093/nar/gkx1230

59. Marsh TC, Henderson E. G-Wires: Self-Assembly of a Telomeric Oligonucleotide, d(GGGGTTGGGG), into Large Superstructures [Internet]. Vol. 33, Biochemistry. 1994. Available from: https://pubs.acs.org/sharingguidelines

60. Yatsunyk LA, Mendoza O, Mergny JL. “Nano-oddities”: Unusual nucleic acid assemblies for DNA-based nanostructures and nanodevices. Accounts of Chemical Research. 2014 Jun 17;47(6):1836–44.

61. Rajendran A, Endo M, Hidaka K, Lan Thao Tran P, Mergny J-L, Sugiyama H. Controlling the stoichiometry and strand polarity of a tetramolecular G-quadruplex structure by using a DNA origami frame. Nucleic Acids Research [Internet]. 2013 Oct 1;41(18):8738–47. Available from: https://doi.org/10.1093/nar/gkt592

62. Venczel EA, Sen D. Synapsable DNA. Journal of Molecular Biology [Internet]. 1996;257(2):219–24. Available from: https://www.sciencedirect.com/science/article/pii/S0022283696901576

63. a twisting electronic nanoswitch.

64. Lat PK, Schultz CW, Yu HZ, Sen D. A Long and Reversibly Self-Assembling 1D DNA Nanostructure Built from Triplex and Quadruplex Hybrid Tiles. Angewandte Chemie - International Edition. 2021 Apr 12;60(16):8722–7.

65. Limongelli V, de Tito S, Cerofolini L, Fragai M, Pagano B, Trotta R, et al. The G-Triplex DNA. Angewandte Chemie. 2013 Feb 18;125(8):2325–9.

66. Bartley JP, Brown T, Lane AN. Solution Conformation of an Intramolecular DNA Triplex Containing a Nonnucleotide Linker: Comparison with the DNA Duplex. Biochemistry [Internet]. 1997 Nov 1 [cited 2021 Dec 10];36(47):14502–11. Available from: https://pubs.acs.org/doi/10.1021/bi970710q

67. Liu K, Lat PK, Yu HZ, Sen D. CLICK-17, a DNA enzyme that harnesses ultra-low concentrations of either Cu+or Cu2+to catalyze the azide-alkyne “click” reaction in water. Nucleic Acids Research. 2020 Jul 27;48(13):7356–70.

68. Gomez-Navarro N, Maldutyte J, Poljak K, Peak-Chew S-Y, Orme J, Bisnett BJ, et al. Selective inhibition of protein secretion by abrogating receptor-coat interactions during ER export. Available from: https://doi.org/10.1101/2022.01.31.478454

69. Rostovtsev V v, Green LG, Fokin V v, Sharpless KB. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective “Ligation” of Azides and Terminal Alkynes. Angewandte Chemie International Edition [Internet]. 2002 Jul 15;41(14):2596–9. Available from: https://doi.org/10.1002/1521-3773(20020715)41:14<2596::AID-ANIE2596>3.0.CO

70. Fantoni NZ, El-Sagheer AH, Brown T. A Hitchhiker’s Guide to Click-Chemistry with Nucleic Acids. Chemical Reviews [Internet]. 2021 Jun 23;121(12):7122–54. Available from: https://doi.org/10.1021/acs.chemrev.0c00928

71. Ishikawa-Ankerhold HC, Ankerhold R, Drummen GPC. Advanced fluorescence microscopy techniques-FRAP, FLIP, FLAP, FRET and FLIM. Vol. 17, Molecules. 2012. p. 4047–132.

72. Proudnikov D, Mirzabekov A. Chemical Methods of DNA and RNA Fluorescent Labeling. Nucleic Acids Research [Internet]. 1996 Nov 1;24(22):4535–42. Available from: https://doi.org/10.1093/nar/24.22.4535

73. Mirzabekov AD. DNA sequencing by hybridization — a megasequencing method and a diagnostic tool? Trends in Biotechnology [Internet]. 1994;12(1):27–32. Available from: https://www.sciencedirect.com/science/article/pii/0167779994900086

74. Bhattacharya S, Shunmugam R. Polymer based gels and their applications in remediation of dyes from textile effluents. Vol. 57, Journal of Macromolecular Science, Part A: Pure and Applied Chemistry. Taylor and Francis Inc.; 2020. p. 906–26.

75. Hardin CC, Watson T, Corregan M, Bailey C. Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d(CGCG3GCG). Biochemistry. 1992;31(3):833–41.

76. Feller D. Ab Initio Study of M + :18-Crown-6 Microsolvation [Internet]. 1997. Available from: https://pubs.acs.org/sharingguidelines

77. Dang LX. Mechanism and Thermodynamics of Ion Selectivity in Aqueous Solutions of 18-Crown-6 Ether: A Molecular Dynamics Study [Internet]. Vol. 117, J. Am. Chem. Soc. 1995. Available from: https://pubs.acs.org/sharingguidelines

78. Pedersen CJ, Tembe BL, Mccammon JA, Harvey SC, Lybrand TP, Wipff G, et al. Molecular Recognition in Nonaqueous Solvents: Na+, K+, and 18-Crown-6 in Methanol [Internet]. Vol. 111, Proc. Natl. Acad. Sci. U.S.A. Cambridge University Press; 1989. Available from: https://pubs.acs.org/sharingguidelines

79. Schulz WW, Bray LA. Solvent extraction recovery of byproduct 137Cs and 90Sr from HNO3 solutions—a technology review and assessment. Separation Science and Technology. 1987;22(2–3):191–214.

80. Chowdhury S, Bansal M. G-quadruplex structure can be stable with only some coordination sites being occupied by cations: a six-nanosecond molecular dynamics study. The Journal of Physical Chemistry B. 2001;105(31):7572–8.

81. Spackova N, Berger I, Sponer J. Nanosecond molecular dynamics simulations of parallel and antiparallel guanine quadruplex DNA molecules. J Am Chem Soc

[Internet]. 1999;121(23):5519–34. Available from: https://www.cheric.org/research/tech/periodicals/view.php?seq=324290

82. Chen FM. Strontium (2+) facilitates intermolecular G-quadruplex formation of telomeric sequences. Biochemistry. 1992;31(15):3769–76.

83. Largy E, Mergny J-L, Gabelica V. Role of Alkali Metal Ions in G-Quadruplex Nucleic Acid Structure and Stability. In: Sigel A, Sigel H, Sigel RKO, editors. The Alkali Metal Ions: Their Role for Life [Internet]. Cham: Springer International Publishing; 2016. p. 203–58. Available from: https://doi.org/10.1007/978-3-319-21756-7_7

84. Deng J, Xiong Y, Sundaralingam M. X-ray analysis of an RNA tetraplex (UGGGGU) 4 with divalent Sr2+ ions at subatomic resolution (0.61 Å). Proceedings of the National Academy of Sciences. 2001;98(24):13665–70.

85. Włodarczyk A, Grzybowski P, Patkowski A, Dobek A. Effect of ions on the polymorphism, effective charge, and stability of human telomeric DNA. Photon correlation spectroscopy and circular dichroism studies. Journal of Physical Chemistry B. 2005 Mar 3;109(8):3594–605.

86. Lim KW, Alberti P, Guédin A, Lacroix L, Riou J-F, Royle NJ, et al. Sequence variant (CTAGGG)n in the human telomere favors a G-quadruplex structure containing a G.C.G.C tetrad. Nucleic Acids Res [Internet]. 2009/08/19. 2009 Oct;37(18):6239–48. Available from: https://pubmed.ncbi.nlm.nih.gov/19692585

87. Reshetnikov R v, Kopylov AM, Golovin A v. Classification of g-quadruplex DNA on the basis of the quadruplex twist angle and planarity of g-quartets. Acta Naturae [Internet]. 2010 Oct;2(4):72–81. Available from: https://pubmed.ncbi.nlm.nih.gov/22649667

88. Jayapal P, Mayer G, Heckel A, Wennmohs F. Structure–activity relationships of a caged thrombin binding DNA aptamer: Insight gained from molecular dynamics simulation studies. Journal of Structural Biology. 2009 Jun 1;166(3):241–50.

89. Pirovano P, Dorrian M, Shinde A, Donohoe A, Brady AJ, Moyna NM, et al. A wearable sensor for the detection of sodium and potassium in human sweat during exercise. Talanta. 2020 Nov 1;219.

90. Modi S, M. G. S, Goswami D, Gupta GD, Mayor S, Krishnan Y. A DNA nanomachine that maps spatial and temporal pH changes inside living cells. Nature Nanotechnology [Internet]. 2009;4(5):325–30. Available from: https://doi.org/10.1038/nnano.2009.83