Direct observation of the formation of DNA triplexes by single-molecule FRET measurements

Il Buem Lee a, Ja Yil Lee a,1, Nam-Kyung Lee b, Seok-Cheol Hong a,c,*

a Department of Physics, Korea University, Seoul 136-713, Republic of Korea
b Department of Physics, Sejong University, Seoul 143-743, Republic of Korea
c School of Computational Sciences, KIAS, Seoul 130-722, Republic of Korea

A R T I C L E   I N F O

Keywords:
DNA triplex
H-DNA
Single-molecule FRET
pH
Strand alignment

A B S T R A C T

In this report we investigated the effects of various biological and chemical factors (DNA sequence, pH, ions, and molecularity) on the formation of DNA triplexes through single-molecule FRET technique. Using this method, we determined how the third strand bound to a DNA duplex and how stable the triplex structure was under various conditions. From this study, we not only verified a variety of well-known features of DNA triplex but also discovered or experimentally supported several interesting behaviors: at neutral pH, a pyrimidine-motif triplex can be formed; the parallel arrangement was not only possible but also dominant over the antiparallel arrangement for a purine-motif triplex. We demonstrated that our method is a versatile analytical tool in studying structural aspects of nucleic acids, particularly non-classical DNA structures, and provides insights into physical mechanism of such structures.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

It is well-known that the genetic material of living organisms, deoxyribonucleic acids (DNA), takes a right-handed double helical structure. This canonical structure is called B-form DNA. Prior to the double helical structure of DNA proposed by Watson and Crick in 1953 [1], Pauling and Corey proposed a triple helix as the structure of DNA [2]. Although the original proposal for the basic structure of DNA turned out to be wrong, the existence of triple helix was reported and confirmed later [3].

From several decades’ research, much has been learned about DNA triplex [4]. In the triplex, three oligonucleotides wind around each other and the third strand binds to a B-form DNA duplex via a Hoogsteen or a reverse Hoogsteen base pair [5] (Fig. 1). The triplex structure mainly forms on a homopurine/homopyrimidine duplex by binding either homopyrimidine or homopurine single-stranded DNA to the homopurine strand of the duplex. In the former type (pyrimidine-motif triplex), the donated single strand (dubbed triplex-forming oligonucleotide (TFO)) is aligned parallel to the homopurine target strand and thus such a triplex is called parallel pyrimidine-motif triplex. In the latter type (purine-motif triplex), the homopurine TFO is aligned antiparallel to the homopurine target strand, and thus such a triplex is called antiparallel purine-motif triplex. In each base pair plane, the base from the TFO is located at the major groove of the duplex and usually forms a Hoogsteen or reverse Hoogsteen base pair [6]. From the fact that the Hoogsteen hydrogen bond is sensitive to the pH of the sample (the protonation of cytosine in acidic environment is critical for the stability of the bond), the parallel duplex as well as the triplex would be a promising molecular component for in situ biological pH sensors [8]. The formation of DNA triplex is also very interesting from the biological viewpoint. Friedreich ataxia, the most common inherited ataxia, is thought to be caused by a severe reduction in the levels of a protein called frataxin, which plays a crucial role in iron homeostasis in mitochondria, because the triplex structure formed in the sequence of the frataxin gene arrests its transcription and consequently suppresses the level of the protein [9].

Past several decades, numerous methods have been applied to gain structural and thermodynamic details of DNA triplex. Gel-based techniques such as chemical and enzymatic probing were...
commonly used to characterize the formation of DNA triplex. Although the methods have been well established, structural information is only acquired after multi-step, destructive processes such as chemical or enzymatic treatments. The gel methods also provide indirect information that requires interpretation or model-based analysis. Traditional optical and thermal techniques have similar limitations. Being more straightforward, the fluorescence resonance energy transfer (FRET) technique, which detects the change in the distance between two positions at the nanometer scale, was used to investigate the strand arrangements of a triplex. Unfortunately, such an ensemble technique is seriously limited in the case of heterogeneous population and dynamic conformational changes.

Here, we investigated the formation of DNA triplex by single-molecule FRET measurements. Taking advantage of the technique, we were able to clearly distinguish two different orientations of the third strand. We also observed that divalent cations stabilize the triplex structure and the acidic condition facilitates the triplex formation in the use of a homopyrimidine TFO. We found that a homopyrimidine TFO can yield a triplex even at neutral pH and this triplex undergoes dynamic transitions between a folded and an unfolded state. Intriguingly, when the third homopurine strand was provided as a separate molecule, the orientation of the TFO was opposite to the well-known antiparallel orientation: in the absence of geometrical restriction, the parallel arrangement thought to be less stable in fact appears more stable than the antiparallel counterpart. This work demonstrates that our single-molecule approach is simple and still powerful as it can characterize physical aspects of biological problems unambiguously and quantitatively.

2. Methods and materials

2.1. Samples

We purchased all the oligonucleotide samples from Integrated DNA Technologies, Inc (Coralville, Iowa, USA). For the intramolecular pyrimidine-motif triplex (intraY), the following strands were used: 5' AAG AAG AAG AAG AAG (Cy5) TGG CGA CGG CAG CGA (Bio) 3' and 5' TCG CTG CCG TCG CCA CTT CTT CTT CTT CTT TTT TCT TCT TCT TCT TCT TCT TCT TCT (Cy3) 3', where Bio is a biotin label for surface immobilization, and Cy3 and Cy5 are a donor and an acceptor dye for FRET measurement, respectively. For the intermolecular pyrimidine-motif triplex (interY), the following strands were used: 5' AAG AAG AAG AAG AAG (Cy5) TGG CGA CGG CAG CGA 3' and 5' (Bio) TCG CTG CCG TCG CCA CTT CTT CTT CTT CTT TTT TCT TCT TCT TCT TCT TCT TCT (Cy3) 3'. The first two oligonucleotides were hybridized first and then the third strand bound as a TFO. The sequence involved in triplex formation was identical to the one in the intramolecular triplex. For the intramolecular purine-motif triplex (intraR), the following strands were used: 5' (Cy3) GAA GAA GAA GAA GAA CTT TTA AGA AGA AGA AGA AGT GGC GAC GGC AGC GA 3' and 5' (Bio) TCG CTG CCG TCG CCA CTT CTT CTT CTT CTT TTT TCT TCT TCT TCT TCT TCT TCT (Cy3) 3'. For the intermolecular purine-motif triplex (interR), the first two strands for interY were used again and the third strand was replaced with 5' (Cy3) GAA GAA GAA GAA GAA (Cy3) 3'. We also tested the same third oligonucleotide with a Cy3 dye labeled in the opposite position: 5' (Cy3) GAA GAA GAA GAA GAA (Cy3) 3'. Again, the sequence involved in triplex formation was identical to the one in the intramolecular triplex.

Fig. 1. Base arrangements in DNA triplexes. In each base triad, the bases in the middle and in the right are from the duplex and paired by Watson–Crick base pairing (dotted) and the base in the left is from the third strand (TFO). A pyrimidine-motif triplex uses the triads shown in the upper panels (a) C–GT and (b) T–AT, and the bases on the third strand bind to the purine-rich sequence (target strand) by Hoogsteen base pairing (dashed). Here C indicates a protonated cytosine. A purine-motif triplex uses the triads shown in the lower panels ((c) A–AT and (d) G–GC) and the bases on the third strand bind to the purine-rich sequence by reverse Hoogsteen base pairing (dashed). Asterisks and dots in the triad notations indicate (reverse) Hoogsteen and Watson–Crick pairing, respectively. R represents the deoxyribose residue connected to the base.
To assemble the molecular modules for FRET measurements, we first dissolved the sample in T50 buffer (10 mM Tris–HCl, 50 mM NaCl, pH = 7.4) and mixed all the oligonucleotides in the same buffer and incubated them above the melting temperature (>90 °C) and slowly cooled down to make sure that the molecules were properly hybridized to form the final constructs. Schematic diagrams of the sample molecules are shown in Fig. 2.

We purchased buffering reagents (MES, HEPES, Tris) from Sigma–Aldrich. A MES buffer with pH = 6.5 contained 50 mM MES, 2 mM Trolox, and ~30 mM NaOH. A HEPES buffer with pH = 7.5 contained 50 mM HEPES, 2 mM Trolox, and ~30 mM NaOH. A Tris buffer with pH = 8.5 contained 50 mM Tris and 2 mM Trolox. For single-molecule FRET measurements, we added “gloxy” (10 μg/ml glucose oxidase (Sigma–Aldrich) and 0.4 μg/ml catalase (Sigma–Aldrich) in T50) and glucose [14] right before injecting the mixture to the sample chamber.

2.2. Single-molecule FRET measurement

Fluorescence resonance energy transfer developed by Förster [15] is an optical technique that provides the nanometer scale spatial information by simply measuring spectral change of fluorescence from two fluorescent dyes. When a (donor) dye whose absorption spectrum lies in the range of shorter wavelength is excited, it will emit fluorescence in its own spectral range. If another dye (acceptor), whose absorption spectrum is overlapped with the donor’s emission spectrum, is within 10 nm from the donor, the energy given to the donor is transferred to the acceptor and the acceptor will emit fluorescence instead. Consequently, the fluorescence is red-shifted. This transfer is more efficient when the two dyes are closer: 50% transfer efficiency is achieved when the two dyes are separated by the so-called Förster distance (typically 4–5 nm).

A single-molecule version of this technique was developed by Ha et al. [16] and has proven to be very useful in studying biological problems. Details of the technique can be found in various review articles [17,18]. Here we briefly outline the experimental scheme [19]. A 532 nm laser (Crystalaer 10 mW) beam was sent to the sample chamber through the edge of a high-NA objective lens (Nikon APO TIRF, 60x, NA = 1.49) for total-internal reflection. The fluorescence from donor and acceptor dyes were collected by the same objective lens and imaged on the EM-CCD (Andor Ixon DV887) after spectral division by a dichroic mirror (Chroma 630DCXR).

To investigate the molecular arrangement of DNA triplex, DNA constructs in T50 buffer were injected in a streptavidin-coated chamber. Then the chamber was washed out by T50 buffer three times (200 μl × 3). Immediately before fluorescence measurements, a solution containing gloxy and glucose was injected to avoid the drift of pH due to the oxygen scavenging system [20] and the sample chamber was mounted on the setup for fluorescence measurements.

For each sample, we measured ~50 frames and each frame has approximately 100 molecules on average, which is sufficient to produce a statistically reliable FRET efficiency (FE) histogram. To obtain time-lapse FRET signal traces, we acquired the data for a longer period (>200 s). The FRET data were then analyzed according to the standard procedure [14].

3. Results

3.1. A pyrimidine-motif triplex was formed in a pH-dependent manner

To demonstrate the formation of DNA triplex, we tested the intraY module at various pH conditions as shown in Fig. 3a–c and Fig. 4. In a weak basic condition (50 mM Tris, pH = 8.5), the molecule exhibited a low FRET efficiency (~0.1) as shown in Fig. 3c. This indicates that the single-stranded tail was not folded (see the left figure of Fig. 2(a)) and the donor dye was far from the acceptor dye located in the middle of the duplex stem. The peak at FE = 0 was from donor-only molecules and shall be ignored. When the buffer was changed to 50 mM MES at pH = 6.5, the FE histogram was changed dramatically (see Fig. 3(a)). The peak at the low FE disappeared and a new peak at FE = 0.85 emerged, which implies that the two dyes were close to each other as depicted in the right

Fig. 2. Molecular constructs used in this study. (a) In the intraY module, a donor (Cy3: hatched, green ball)-labeled 3'(TTC)5 tail serves as a triplex-forming oligonucleotide. Once it is folded to form an H-DNA-like triplex by protonation, a high FRET efficiency would be observed due to the shorter distance between the donor and the acceptor (Cy5: solid, red ball) on the duplex stem. Relative sizes of the dyes show which dye emits more fluorescence. Hereafter dots and asterisks indicate Watson–Crick base pairing and Hoogsteen pairing and the bases in the third strand are shown in red, bold face. (b) In the interY module, a separate Cy3-labeled third strand forms a triplex. (c) In the intraR module, a donor (Cy3)-labeled 5'(GAA)3 tail is linked to the purine-rich strand of the duplex and folded back to form a H’-DNA-like triplex, Mg2+ facilitates the folding transition. (d) Different from the interY module, the interR module contains a separate strand of (GAA)5 sequence. Depending on the dye location, the parallel alignment of the third strand yields different FRET efficiencies. Inverted triangles at the bottom of the constructs represent biotin labeling. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
This result is consistent with the well-known geometry of the parallel pyrimidine-motif triplex: the mirror-symmetric pyrimidine sequence was folded back in the middle to bind to the purine sequence and thus form a parallel triplex. ‘Parallel’ means that the directionalities of the third strand and the target (purine) strand are the same. When we tested the effect of a more acidic condition (pH < 6.5), the histogram basically remained unchanged (data not shown). The folded structure for the third pyrimidine strand is the characteristic feature of the H-DNA structure.

In order to see whether the direct linkage of the third strand has any role in the formation of the folded structure and to demonstrate the formation of DNA triplex in more general circumstance, we examined the formation of triplex using a DNA duplex and a separate TFO. As shown in Fig. 5 and Fig. 2 (b), the third strand was bound to the duplex in the same orientation as in the intraY module. One notable difference is that in order to induce the association of the third strand, a small physiological concentration (1 mM) of Mg$^{2+}$ was necessary presumably because Mg$^{2+}$ helped the TFO overcome the repulsive electrostatic barrier imposed by negatively charged DNA. In the absence of Mg$^{2+}$ or at neutral pH, we could not observe the formation of triplex (data not shown). On the contrary, the TFO in the intraY module was attached to the duplex stem and allowed to search an only limited volume around the pivot loop, which expedited the formation of the triplex.

A similar Mg$^{2+}$-dependent enhancement of triplex formation was also found when the intraY module was in a weak basic condition. As described in Fig. 3 (d), the addition of a small concentration (0.3 mM) of Mg$^{2+}$ helped the intraY module form
a triplex as seen by the large peak at high FE. Certainly, divalent
cation promoted the formation of triplex.

3.2. A pyrimidine-motif triplex can be formed at neutral pH,
displaying dynamic inter-conversions between folded and unfolded
states

Next, we checked how the molecule behaved at neutral pH.
When the molecule was at pH ~ 7.5, we observed two peaks at
high and low FE as shown in Fig. 3 (b) while molecules at pH = 6.5
and 8.5 showed a FE peak at either ~0.9 or ~0.1. The sensitivity of
the pyrimidine-motif triplex to the solution pH is already well
established because the stability of the structure depends upon the
protonation of cytosine. Considering that the $pK_a$ of cytosine is in
fact quite low (~4.5 or less [21]), the stable folding at neutral pH is
rather surprising and this will be discussed in the section of
discussions and conclusions.

By analyzing the FRET signals of individual molecules in the
course of time, we captured the dynamic nature of the transition.
As shown in Fig. 4, the FRET from the intraR module alternated in
two different states, indicating that the molecules underwent
frequent inter-conversions between folded and unfolded confor-
mations in the time-scale of seconds.

3.3. A purine-motif triplex was stabilized by divalent cation

If there is a mirror-symmetric pyrimidine-rich (or homopyr-
imidine) sequence, its complementary sequence should be also
mirror-symmetric and purine-rich (or homopurine). It is also
known that a mirror-symmetric purine-rich sequence can form
da DNA triplex [4]. Thus, we then checked whether we can see
a third homopurine strand bind to a duplex to form a DNA triplex.
Since it is easier to study such a transition with the intramolecular
configuration, we first monitored the FRET change from the intraR
module. As the formation of purine-motif triplex does not require

![Fig. 6](image_url)

**Fig. 6.** The presence of Mg$^{2+}$ ions is critical for the formation of an H*-DNA-like intramolecular triplex. FRET histograms from the intraR module at (a) [Mg$^{2+}$] = 0 mM, (b) 1 mM, and (c) 10 mM. Here the solution pH was 7.5.

the acidic condition, we only tested the effect of divalent cation.
**Fig. 6** (a) shows that in the absence of Mg$^{2+}$, the molecule exhibited
a low FRET efficiency, indicating that the single-stranded homo-

purine tail was not folded and the donor was far away from the
acceptor. In the presence of ~10 mM Mg$^{2+}$, the low FRET peak
almost diminished and a new peak at high FE increased remarkably
(see Fig. 6 (c)). The folded structure by the third purine strand is the
characteristic feature of the H*-DNA structure [22]. The physio-

logical level of Mg$^{2+}$ was sufficient to permit the stable formation
of DNA triplex in the neutral pH condition, which implies that the
purine-motif triplex may be more relevant to biological
phenomena. One thing worth mentioning is that when the folding
was induced by divalent cation such as Mg$^{2+}$, there were almost
always a small number of molecules showing mid FRET values. This
is presumably due to some scrunching of the single-stranded tail by
divalent cations so that the distance between the dyes became
intermediate between the fully folded and fully unfolded states.

3.4. The parallel purine-motif triplex was dominant over the
antiparallel triplex

Next, we would check the molecular orientation observed with
the intraR module using the intermolecular arrangement. Thus,
we introduced a third homopurine TFO (3’ Cy3-labeled) to the
sample chamber in which duplex molecules bearing the same
homopurine/homopyrimidine tract were surface-immobilized. To
our surprise, the FE was high (~0.9) (Fig. 7 (a)), indicating that the
third strand bound to the duplex was in the opposite orientation
as shown in Fig. 2 (d). The homopurine TFO strand appeared to
align parallel to the target homopurine sequence. It was quite
unexpected because from literature the antiparallel purine-motif
triplx is believed to be more stable than the parallel counter-
part [10]. In order to double-check it and to rule out the possibility
that this unexpectedly high FRET efficiency is a merely artifact
due to strong sequence-nonspecific scrunching of the whole molecule
by the large Mg$^{2+}$ concentration (~10 mM), we tested the same
TFO except for the dye location (Cy3 label at 5’ end). If the high
FRET is due to strong ion-induced condensation, we will see
similar FRET signal with this molecule. Or if the high FRET ef-

iciency was due to the parallel alignment this molecule would also
align in the same manner, the FRET efficiency will decrease
because the donor would be 15 bp away from the acceptor. As
shown in Fig. 7 (b), the FE was reduced to ~0.33, which is consistent with the parallel attachment of the third strand. If the
third strand is antiparallel, we expect to observe high FE for the 5'-labeled TFO. From Fig. 7 (b), there is no sign of the antiparallel alignment. The strand arrangements are depicted in Fig. 2 (d).

4. Discussions and conclusions

Single-molecule FRET technique allowed us to obtain straightforward, intuitive information about the configuration of the strands involved in triplex formation. Molecular arrangements including the orientation of the third strand can be clearly distinguished according to distinct FE distributions. For each molecule, we can immediately tell which orientation the third strand adopts. By virtue of the simplicity and clarity of the method, we can address several issues in this well-known problem of DNA triplex with certainty and reliability.

First, we were able to confirm the essential features of various triplex configurations. We can clearly see the pH-driven formation of pyrimidine-motif triplexes and the stabilization of triplexes by divalent cations. The molecular arrangement, mainly relative orientations of the strands, was verified by the direct assessment of the distance between two points in the molecule.

Interestingly, the pyrimidine-motif triplex remained reasonably stable even at neutral pH which was at first thought quite puzzling because the protonation of cytosine is critical for the stability of triplex and the pK<sub>a</sub> of free cytosine is quite low (~4.5 or less). Recent experimental and theoretical work claimed that the pK<sub>a</sub> of cytosine buried in the major groove is as high as 7 or 8 [23,24]. Our finding is well suited to these previous reports and strongly supports their findings.

Another more intriguing finding of our study is that in the intermolecular triplex configuration, the third homopurine strand was bound to the duplex in the parallel orientation. Since the target sequence is an identical homopurine sequence, the target and the TFO have the same sequence and orientation, resembling the synaptic complex found in homologous recombination. In fact, such a triplex is called the R-form triplex (recombination-form) and has been known for many years as an intermediate form of homologous recombination. The R-form triplex has been mainly implicated in synaptic complex found in homologous recombination. In fact, such a triplex in the absence of auxiliary proteins [10,13]. Our finding demonstrates the existence of such R-form triplex. It, however, calls for further studies to understand what factors determine the orientation of the third strand and why we observed the parallel alignment dominantly although a majority of earlier work supported the opposite situation.

In this report, we clearly determined the molecular arrangements of a variety of triplex conformations by single-molecule FRET measurements and investigated the effects of various factors on the formation of triplex. From the results, we demonstrated that our approach is powerful and versatile in studying nano-scale conformational diversity of biomolecules such as nucleic acids and in examining the physical backdrop of various intriguing structures such as triplex.

Acknowledgment

This work was supported by Mid-career Research Program through NRF grant funded by the MEST (NRF 2009 0084933) (N.-K. Lee and S.-C. Hong).

References