Biochimie 146 (2018) 156–165

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Research paper

Effect of dC → d(m^5C) substitutions on the folding of intramolecular triplexes with mixed TAT and C^+GC base triplets

Carolyn E. Carr, Rajkumar Ganugula [1], Ronald Shikiya [2], Ana Maria Soto [3], Luis A. Marky*

Department of Pharmaceutical Sciences, University of Nebraska Medical Center, 986025 Nebraska Medical Center, Omaha, NE, 68198-6025, USA

ARTICLE INFO

Article history:
Received 19 October 2017
Accepted 19 December 2017
Available online 24 December 2017

Keywords:
Triplex
Duplex
Methylation
Epigenetics
Thermodynamics

ABSTRACT

Oligonucleotide-directed triple helix formation has been recognized as a potential tool for targeting genes with high specificity. Cystosine methylation in the 5′ position is both ubiquitous and a stable regulatory modification, which could potentially stabilize triple helix formation. In this work, we have used a combination of calorimetric and spectroscopic techniques to study the intramolecular unfolding of four triplexes and two duplexes. We used the following triplex control sequence, named *Control Tri*, d(AGAGAC_5TCTCTC_5TCTCT), where C_5 are loops of five cytosines. From this sequence, we studied three other sequences with dC → d(m^5C) substitutions on the Hoogsteen strand (*2MeH*), Crick strand (*2MeC*) and both strands (*4MeHC*). Calorimetric studies determined that methylation does increase the thermal and enthalpic stability, leading to an overall favorable free energy, and that this increased stability is cumulative, i.e. methylation on both the Hoogsteen and Crick strands yields the largest favorable free energy. The differential uptake of protons, counterions and water was determined. It was found that methylation increases cytosine protonation by shifting the apparent pK_a value to a higher pH; this increase in proton uptake coincides with a release of counterions during folding of the triplex, likely due to repulsion from the increased positive charge from the protonated cytosines. The immobilization of water was not affected for triplexes with methylated cytosines on their Hoogsteen or Crick strands, but was seen for the triplex where both strands are methylated. This may be due to the alignment in the major groove of the methyl groups on the cytosines with the methyl groups on the thymines which causes an increase in structural water along the spine of the triplex.

1. Introduction

Triple stranded DNA, or triplexes, are a regulatory structure found within chromosomal DNA [1,2]. Because triplexes are inherently unstable, a feature which implicates the involvement of triplexes in cancer and neurodegenerative disorders [3–9], they can be hard to detect and verify. Because of triplex instability, they have a high rate of generating repeat sequences such as in the case of Freidrich's ataxia, which becomes progressively worse in successive generations [8,10–12]. Despite their perceived instability, triplexes have been associated with a variety of roles related to genomic regulation; this role is supported by the fact that most identified intramolecular DNA triplex forming sequences occur only once in the genome, which imparts specificity to triplexes as a regulatory binding element [3,13]. Triplexes have been shown to be involved in genomic regulation by acting as pause sites during replication [6,14–18], posttranslational processing through mRNA splicing [19,20], and chromatin organization by triplex formation through distant sequences [21–23]. In addition, it has been shown that triplexes promote methyltransferase recruitment and methylation of downstream cystosines [11,24–26], which alters gene expression and is important for gene imprinting and cell differentiation [27–33]. Recent work has suggested that the solution conditions of the cell, not often replicated *in vitro*, significantly stabilize Hoogsteen hydrogen bonds, leading to stable formation of triplexes [34]. This research, coupled with the identification of over a thousand sites for triplex formation by long noncoding RNAs (lncRNAs) [35,36], suggests that triplexes may be a more

* Corresponding author.
E-mail address: lmarky@unmc.edu (L.A. Marky).
[1] Current address: Aisin Cosmos R&D Co. Ltd, Hyderabad, Telangana, India.
[2] Current address: Department of Med. Microbiology & Immunology, Creighton University, Omaha, Nebraska, USA.
[3] Current address: Department of Chemistry, Towson University, Towson, Maryland, USA.

https://doi.org/10.1016/j.biochi.2017.12.008
0300-9084/

widespread regulatory mechanism than originally hypothesized. In addition, significant synthetic work has been done to generate a triplex forming third strand with greater stability and pH independence for use as an antigene therapy [37–40].

Epigenetic affects due to cytosine methylation have long been known to be a critical factor in vertebrates [28,33,41,42], and are associated with key biological processes such as genomic imprinting [32,33,43,44], X-chromosome inactivation [45–48], silencing of transposons [33,49–51] and cell differentiation [45,52,53]. These methylation-dependent processes typically occur during early development and exclusively occur at CpG dinucleotides. However, current research has shown that dynamic methylation occurs during other times and in other cell types, with a preference for CpA dinucleotides [54,55]. This has shown to be especially prevalent in repeat sequences and transposons of fungi, which display little CpG methylation but an abundance of CpA and CpT methylation [56,57]. The methylation of the repeat sequences is thought to silence transposons and prevent the repeat sequences from being extended. CpA methylation is also seen in methylation studies of *Drosophilia,* which was previously believed to not utilize methylation for gene regulation [58]. Specifically, significant CpH methylation has been found in mature human and mouse brains cells, but not in infant brain cells [54,59,60]. These studies revealed that non CpG methylation accumulates appreciably through human brain development [54]. Embryonic stem cells, pluripotent stem cells and oocytes have shown significant levels of non-CpG methylation which is then lost after cell differentiation, suggesting a dynamic role in gene regulation for non-CpG methylation rather than the static role of CpG methylation [54,55,61–64]. Because triplexes have been linked with chromatin organization and enhanced methylation of downstream cystosines [4,24,65,66], and studies have shown dynamic methylation at non-CpG sites, we decided to investigate the thermodynamic stability of methylated pyrimidine triplexes, with specific focus on determining the differential binding of protons, counterions, and water. The immobilization of these molecules is important for determining DNA stability and for their ability to interfere with polymerase binding; they also have implications in antigene strategies that seek to generate a methylated triplex that could hinder polymerase binding, or disrupt a triplex by forming two duplexes.

Triplexes consist of three strands, with two of these strands forming canonical Watson-Crick base-pairs with each other, and a third strand binding to the Watson strand through Hoogsteen or reverse-Hoogsteen hydrogen bonds. There are two types of intramolecular DNA triplexes, a pyrimidine or parallel triplex, and a purine or anti-parallel triplex. The pyrimidine triplex has a pyrimidine-rich Hoogsteen strand, which runs parallel to the Watson strand and binds to it through Hoogsteen base-pairing. If this third strand contains cytosines, then only one hydrogen bond is possible without protonation at its N3 position, and thus these triplexes are stabilized by low pH to facilitate protonation [38,67]. A pyrimidine triplex may also contain only thymines and these triplexes are stabilized by mono and divalent cations. If the third strand is rich in purines then it will run anti-parallel to the Watson strand and bind using reverse-Hoogsteen hydrogen bonds; purine rich triplexes are stabilized by mono and divalent cations [39,68]. The triplexes studied below and shown in Fig. 1 are examples of pyrimidine, or parallel, triplexes, with the third strand binding to the Watson strand with Hoogsteen hydrogen bonds.

Ions and water molecules are critical for the native function of DNA, regardless of structure, but are often overlooked. As a biopolymer, DNA is negatively charged and the repulsive forces from these charges play a role in dictating the structure of DNA. Structures such as intramolecular junctions, which have a high charge density around the junction, have a strong structural

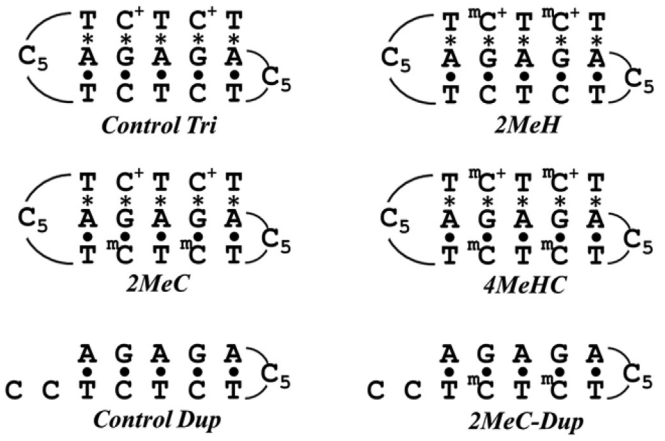

Fig. 1. Hypothesized structures of triplexes and duplexes. • indicates Watson-Crick base-pairing, * indicates Hoogsteen base-pairing, and $^{m}C^{+}$ represents methylated cytosines.

requirement for cations to neutralize the negative charge repulsion [69,70]. The same is true for triplexes, which have a third strand inserted into the major groove and bound to the Watson strand [71–73]. While triplexes may have a higher counterion uptake, the cytosine protonation necessary for pyrimidine triplex formation would exclude counterions and may have an impact on the overall stability. In addition to influencing triplex stability, any ions bound to the surface of the DNA would need to be displaced, an entropically favorable process, in order to interact with a complementary strand, such as in an antigene therapy, or for a protein to bind.

Water also plays a fundamental role in determining the secondary and tertiary structure of oligonucleotides [74], as previous research has indicated that nucleic acids are heavily hydrated [75,76]. The overall hydration of an oligonucleotide is dependent on its conformation, nucleic acid composition, and sequence [77–80]. The precise details and determinants of hydration have yet to be fully elucidated, especially when taking into account the different types of water bound to the surface of nucleic acids; hydrophobic or structural water, associated with polar and non-polar groups, and electrostricted water, which are immobilized by charged groups [81,82]. These two types of water are difficult to differentiate and the results are further confused when considering the hydration sphere of the ions bound to oligonucleotides. In a similar manner to ions, water must be displaced in order for another molecule to bind to the surface of DNA. Cation and water release have a major effect on the binding of proteins to nucleic acids due to the contribution to a favorable binding entropy.

Our lab is focused on understanding the unfolding thermodynamics of DNA structures, the physical factors that control which structures they form and the free energy of these forms. In addition, we seek to complete these thermodynamic profiles by measuring the uptake of water, protons and counterions, which are sequence and structure dependent. In this study, we used a combination of spectroscopic and calorimetric techniques to investigate the unfolding behaviour of a pyrimidine triplex (*Control Tri*) at pH 5.2 and 6.2, and compared its behaviour with that of a triplex containing 5-methylcytosines on the Hoogsteen strand (*2MeH*), the Crick strand (*2MeC*), and both strands (*4MeHC*). Our results suggest that methylation increases the thermal stability and free energy of formation of the duplex domain, but not the triplex. Methylation increases the pH range at which the cytosines become protonated; a higher amount of protonation causes a greater exclusion of counterions due to the positively charged cystosines. Methylation did not appear to increase water immobilization to any significant