



Quadruplex World

Besik Kankia^{1,2}

Received: 12 April 2021 / Accepted: 10 June 2021

© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

The RNA world hypothesis relies on the double-helix complementarity principle for both replication and catalytic activity of RNA. However, the de novo appearance of the complementarity rules, without previous evolution steps, is doubtful. Another major problem of the RNA world is its isolated nature, making it almost impossible to accommodate the genetic code and transform it into modern biochemistry. These and many other unanswered questions of the RNA world led to suggestions that some simpler molecules must have preceded RNA. Most of these alternative hypotheses proposed the double-helical polymers with different backbones but used the same complementarity principle. The current paper describes a fundamentally different idea: the de novo appearance of a nucleic acid polymer without any preexisting rules or requirements. This approach, coined as the quadruplex world hypothesis, is based on (i) the ability of guanines to form stable G-tetrads that facilitate polymerization; and (ii) the unique property of polyguanines to fold into a monomolecular tetrahelix with a strictly defined building pattern and tertiary structure. The tetrahelix is capable of high-affinity intermolecular interactions and catalytic activities. The quadruplex world hypothesis has the potential to address almost all the shortcomings of the RNA world.

Keywords Quadruplex · G-tetrad · Tetrahelix · RNA world · Origin of life · Prebiotic · Homochirality · Genetic code

The RNA world hypothesis assumes that the genetic information is encoded in an RNA template and via the complementarity principles transmitted to the replicons. The latter forms the tertiary structure, employing the same principles, responsible for molecular recognition and catalysis. Thus, the RNA world relies on the complementarity principle for both the replication and catalytic activity of RNA. However, the formation of the complementarity principle without prior evolution steps is doubtful (Koonin 2012). This and many other unresolved problems led to suggestions that RNA must be preceded by a more straightforward genetic material capable of continuous evolution into modern biochemistry

✉ Besik Kankia
kankia.1@osu.edu

¹ Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH 43210, USA

² Institute of Biophysics, Ilia State University, Tbilisi 0162, Republic of Georgia

(Anastasi et al. 2007; Joyce 2002; Orgel 2004). Several attempts have been made to design such molecules, but most of them represent double-helical polymers using the same complementarity principle (Orgel 2004). Besides, none of them is much simpler than RNA and introduces additional problems of "genetic takeover" or displacement of these genetic materials by RNA. Can a stable tertiary structure, capable of molecular recognition and catalysis, be produced without the complementarity or any other preexisting rules? To address this question, I consider here likelihood of the de novo appearance of nucleic acid quadruplexes on primitive Earth. The properties of the quadruplex world's main structural domain, G₃N (G₃NG₃NG₃NG₃; where N is any base), are discussed. Next, I consider the formation of self-reproducing tetrahelices and the role of the temperature (heating/cooling) cycles in this process. Finally, I propose that the quadruplex world addresses most of the deficiencies of the RNA world hypothesis.

Properties of G₃N Quadruplex

G₃N quadruplex is Highly Programmable

The term "programmable" in primordial evolution describes the ability of specific sequences to fold into the designed/programmed tertiary structures capable of specific interactions with other molecules or of performing some catalytic activity (DeLuca et al. 2020; Seeman 2010). The DNA duplex's programmable nature is based not only on high specificity of canonical Watson–Crick base pairs, but also on the strictly determined antiparallel alignment of strands. Thus, the tertiary structure should have strictly defined base recognition and strand alignment rules to be programmable.

G-tetrads (G•G•G•G) are square planar arrangements of guanines; each G directly interacts with adjacent Gs using both Watson–Crick and Hoogsteen bonding interfaces. In addition, each G has indirect (through cation) contact with the diagonally positioned G (Fig. 1A). It is not surprising that other nucleotides are incapable of such sophisticated interactions and other homo or mixed tetrads (i.e. A•A•A•A, G•C•G•C, or G•G•G•T) strongly destabilize quadruplexes (Mergny and Sen 2019). Despite these strictly defined

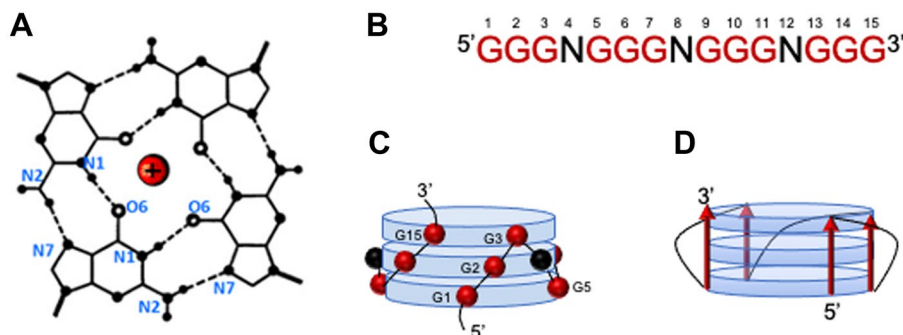


Fig. 1 **A** G-tetrad with cation (red) in the center. **B** Nucleotide sequence of G₃N (N is any base). **C** Three-dimensional (3D) representation of G₃N quadruplex with all parallel G₃ segments (red spheres) and the chain-reversal loops (black spheres); blue discs represent G-tetrads. **D** Schematic representation of G₃N quadruplex with all parallel G-tracks (red) and chain-reversal T-loops (black curved lines)

recognition/association rules, G-tetrads can adjust to any strand alignment (i.e., parallel, antiparallel and hybrid topologies) (Mergny and Sen 2019). The structural polymorphism is so strong that the topology of many monomolecular quadruplexes can depend on experimental conditions or even the history of sample preparation. One of the few monomolecular quadruplexes that demonstrates strict monomorphism is G3N, which contains all-parallel G-tracts and double chain-reversal (propeller) single nucleotide loops (Fig. 1B-D) (Kelley et al. 2011).

G3T Quadruplex is Unusually Stable

The G3T ($G_3TG_3TG_3TG_3$) quadruplex is a truncated version of a DNA aptamer (G_3T)₄, selected for binding to HIV-1 integrase (Jing and Hogan 1998). Most of the initial studies performed using the quadruplex with T-loops, G3T, revealed that G3T is the most stable known structural motif in the entire nucleic acid world. For instance, even in the presence of 0.1 mM K⁺ (the most efficient quadruplex forming cation) it demonstrates cooperative and fully reversible melting curves at 55 °C and in 50 mM K⁺ it melts near 100 °C (Kankia et al. 2016; Kelley et al. 2011). The remarkable stability of the G3T quadruplex is attributed to the all-parallel alignment of GGG-tracts and the single-nucleotide propeller loops that are responsible for its structural monomorphism (Do et al. 2011; Kelley et al. 2011). Any modification of the G3T sequence, besides nucleotide substitutions in loop positions, is accompanied by strong destabilization and structural polymorphism (Kankia 2018).

RNA Analogue of G3T has the Same Topology

While DNA and RNA form double helices with significantly different geometry (B vs A-conformations), G3T and its RNA analogue adopt precisely the same tertiary structure (Joachimi et al. 2009; Kankia 2019). The RNA quadruplex is 13° C more stable than G3T (Joachimi et al. 2009; Kankia 2019). However, not all DNA-to-RNA substitutions increase stability. Depending on the position, the substitution can either have no measurable effect, stabilize or even destabilize the structure; the effects vary between -1.5 °C and 3.5 °C per substitution (Kankia 2019).

Tetrahelical Monomolecular DNA (tmDNA)

The tmDNA is a homopolymer consisting of n number of G3N domains, poly(G3N) or (G3N) _{n} (Kankia 2014). The terminal G₃-segments of adjacent G3T domains form G₆-segments (Fig. 2A). Each G3N domain of the architecture is formed by zigzagging of G₃-segments and N-loops, while the G₆ segment sheared by adjacent domains and serves as a bridge between them (Fig. 2B-D). The G₆-segments are responsible for the structure's vertical growth, while G₃-segments and T-loops move DNA strands horizontally. The tmDNA folds rapidly and demonstrates even higher thermal stability than the G3N domain (Kankia 2014). The RNA analog of tmDNA, tmRNA, is capable of forming the same architecture, and all properties discussed below can extrapolate to RNA (Kankia 2019).

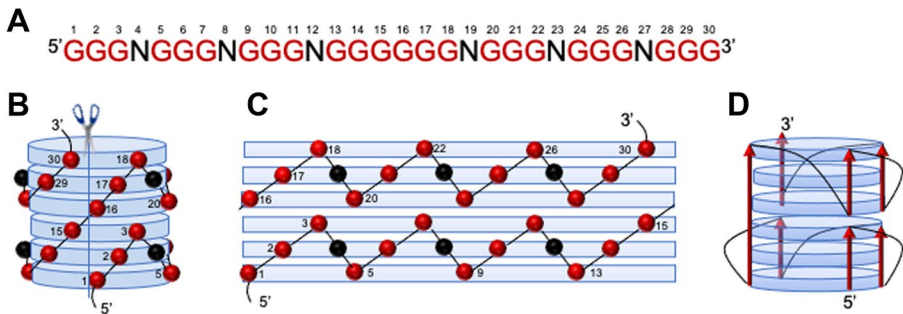


Fig. 2 **A** Nucleotide sequence of $(G3N)_2$ (N is any base). **B** 3D representation of $(G3N)_2$ quadruplex with all parallel G-tracks (red spheres) and the chain-reversal loops (black spheres); blue discs represent G-tetrads. **C** Surface of the 3D model of $(G3N)_2$ unwrapped into 2D map. **D** Schematic representation of $(G3N)_2$ quadruplex with all parallel G-tracks (red) and chain-reversal T-loops (black curved lines)

Polyguanine Folds into G3G-domain tmDNA Architecture

The tmDNA architecture can adopt only the G3N building pattern (GGG-tracks with single nucleotide loops); any other pattern (e.g., G2N ($G_2NG_2NG_2NG_2$) or G4N2 ($G_4N_2G_4N_2G_4N_2G_4$)) is incapable of forming a stable, uninterrupted tmDNA (Kankia 2018). As a result, polyguanine adopts a tmDNA structure using only the G3G building pattern (Kankia 2018). Thus, a non-specific polyguanine homopolymer can form a sophisticated tertiary structure with a strictly defined and highly programmable building pattern.

tmDNA is Highly Reproducible and Predictable

As shown earlier, G_{15} , $(G3T)_2$, or $(G3T)_7GGG$ do not show any misfolding after rapid cooling on ice (Kankia et al. 2016). The latter construct is a variant of $(G3T)_2$ with T-insertion between G3T domains, G3T-T-G3T. Since it contains eight identical G₃-segments, it is more inclined to misfold. For instance, it might form only one G3T domain in the middle of the construct with unstructured tails, i.e., $G_3TG_3T-(G3T)-TG_3TG_3$. However, it forms two perfectly folded G3T quadruplexes even upon rapid annealing on ice (Kankia 2018). Structural reversibility upon rapid annealing is uncommon for RNA, which requires careful annealing to restore initial structures.

tmDNA structures formed from sequences with A, T, or C loops can be programmed and predicted with 100% accuracy. The structure of polyguanine can also be predicted with 100% accuracy if its length equals $n \times 15$ (i.e., G_{30} , G_{45}). If this is not the case, one can still predict number of G3G domains in tmDNA, without precise positioning of the ends of the structure. For instance, G_{32} might fold into $GG-(G3G)_2$, $G-(G3G)_2-G$ or $(G3G)_2-GG$. However, G-hydrolysis at the loop positions will increase the predictability to 100% (see "Exposed Bases"). This kind of structure predictability is unusual for biopolymers. For instance, only short DNA duplexes with specifically designed Watson–Crick base pairs can be predicted with 100% accuracy, while the highest accuracy for proteins is around 80% (Heffernan et al. 2015).

Exposed Bases

tmDNA includes fully exposed nucleic acid bases similar to the Pauling model of DNA (Pauling and Corey 1953). The exposed bases create specific advantages: (i) they could be used for intermolecular base-pairing without unfolding the structure; (ii) since they are completely exposed to solvent, the loop bases can easily hydrolyze from the sugar-phosphate backbone. This would transform polyguanine into G₃ and G₆ segments separated by abasic sites. As a result, the structure predictability will increase to 100% accuracy since the abasic sites represent the most favorable chain-reversal loops (Rachwal et al. 2007); (iii) the abasic sites can be used for incorporation of other bases (e.g., cytosine) without affecting the homochirality of the polymer (see "Origin of Homochirality").

High-affinity Interactions between tmDNA Molecules

The tmDNA architecture is capable of high-affinity intermolecular binding through quadruplex-and-Mg²⁺ connections (QMCs) (Kankia 2015). QMC is based on the shape complementarity and strong stacking interaction between the quadruplexes. Figure 3A shows a pair of QMC connectors: Sub A (blue) or (G3T)₂, missing the last guanine at the 3'-end; and Sub B (red), a G3T monomer with one extra guanine at the 5'-end. Upon mixing, the substrates interact with each other and form the uninterrupted G3T trimer with one interface G-quartet interface, which serves as an interlocking mechanism and secures proper folding of the product G3T trimer. QMCs are characterized by attomolar binding affinity and dissociates only after removing excess ionic strength (i.e., removal of Mg²⁺ by EDTA) (Kankia 2015).

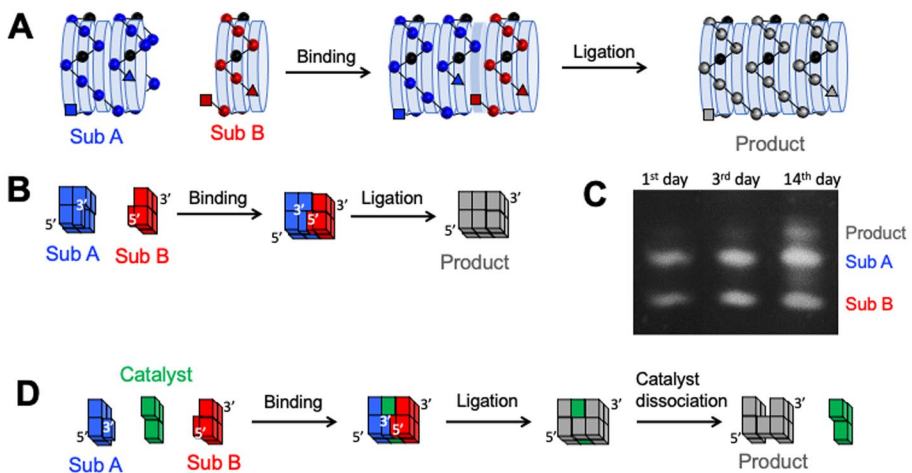


Fig. 3 **A** Design and **B** schematic of self-ligating (two-component) QMC system. **C** Ligation of the two-component system at 22 °C in 50 mM KCl, 10 mM MgCl₂, 10 mM Tris, pH 8.8 visualized on 12% polyacrylamide denaturing gel stained by SYBR Gold. **D** Model of three-component QMC system using G-triplex (green) as a quadruplex catalyst

Catalytic Activity of tmDNA

The terminal guanines of Sub A and Sub B in the formed QMC complex (Fig. 3A-B) are juxtaposed, allowing non-enzymatic ligation in the presence of a carbodiimide phosphate activating agent (e.g., EDC) (Kankia 2021). The ligation is extraordinarily robust and efficient: it can proceed in boiling solutions and complete ligation is achieved within a minute. To demonstrate the spontaneous ligation of quadruplexes the two-component QMC reaction (see Fig. 3A-B) was conducted in the aqueous solution in the absence EDC. The visible product formation is observed in 2 weeks (Fig. 3C). While the ligation of RNA in the absence of the condensing agents is a well-established phenomenon, it is uncommon for DNA. The present result is likely attributed to unusually high stability of the QMC complexes. A similarly high ligation efficiency is achieved using a quadruplex catalyst (i.e., $G_3TG_3TG_3$ or G_{11}) (see Fig. 3D). This system employs the QMC partners without a locking mechanism and is templated through the quadruplex catalyst, which demonstrates turn-over activity upon temperature cycling (Kankia 2021).

Self-polymerization of Guanines

Template-free Production of Quadruplexes

Figure 4 shows possible polymerization of polyguanines from free G-monomers. Feasibility of step 1 is based on experimental studies revealing that (i) free guanosines or GMPs form G-tetrads stacked on each other with a similar helical parameters as quadruplexes (Chen et al. 2020; Wu and Kwan 2009; Zimmerman 1976); and (ii) 3',5'-cyclic GMP demonstrates non-enzymatic self-polymerization with canonical 5'-3' phosphodiester backbone formation (Costanzo et al. 2009; Morasch et al. 2014; Pino et al. 2008). The feasibility of steps 2 and 3 is demonstrated by ligation of pG₁₀ in the presence of EDC, which results in polymers with lengths of at least a few thousand nucleotides (Kankia 2021). Feasibility of steps 5 and 6 discussed in "High-affinity Interactions between tmDNA Molecules" and "Catalytic Activity of tmDNA".

The scheme shown in Fig. 4 employs the temperature cycles at steps 2 and 4. In step 2, temperature cycling facilitates the rearrangement of short multimolecular quadruplexes into the G-wires, and in step 4, longer multimolecular quadruplexes are

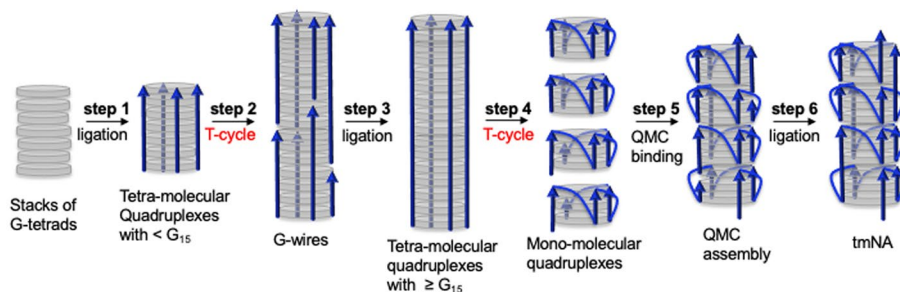


Fig. 4 Steps in possible prebiotic polymerization of quadruplexes. Only G-tetrads (gray disks) and phosphodiester backbone (arrows) are shown. Curved lines (reactions 4–6) correspond to propeller loops formed by single Gs

transformed into monomolecular tmDNA. The importance of the temperature cycles in primordial evolution is discussed in the following section.

Role of Temperature Cycles in Quadruplex Formation

In primordial Earth, the formation of biologically essential molecules was driven by fluctuations of physicochemical parameters (Salditt et al. 2020). For example, production of RNA molecules in the RNA world involves at least two steps: (i) replication of the template; and (ii) dissociation of the replicon to allow the next cycle of replication. Thus, after each replication cycle, the RNA duplex must be denatured/unfolded. This could have been achieved by the temperature cycles produced by day/night fluctuations. In the RNA world, these cycles can also have adverse effects: (i) RNA can easily misfold upon temperature cycling (Tinoco and Bustamante 1999); and (ii) if the replication session is longer than the temperature cycle, incomplete/nonfunctional RNA will be produced. In contrast, the temperature cycles fit perfectly into the quadruplex world since tmDNA does not misfold or have any length limitation. Moreover, the temperature cycles create additional opportunities in the quadruplex world (see the following sections).

The Temperature Cycles Can Overpower Detrimental Hydrolysis

One of the main barriers to abiotic nucleic acid polymerization/condensation (with release of water) is the opposite reaction, hydrolysis. In a water solution, an unstructured nucleic acid ultimately hydrolyzes to its components. To produce a polymer, the rate of condensation should be faster than the rate of hydrolysis. In other words, hydrolysis can be overpowered by increasing the rate of condensation. In the absence of a catalyst, this can be achieved by temperature elevation that decreases the activation energy barrier of the condensation. While temperature elevation would dissociate/diffuse nucleotide aggregates and unfold RNA, G-tetrads can survive high temperatures (Smith et al. 2018). Thus, the high thermal stability of G-tetrads and quadruplexes shifts the hydrolysis-condensation equilibrium towards condensation. In addition, wet-dry cycles, which can be a consequence of the heating/cooling cycles, further accelerate condensation by removing water molecules, including newly condensed ones (Deamer and Weber 2010). Thus, the quadruplex world can be easily powered by the most prominent and available energy, heat.

The Temperature Cycles Can Serve as an Enrichment Tool for Quadruplexes

Since the quadruplexes based self-recognition of guanines, G-oligomers of equal length usually can slip against each other and form multimolecular long G-wires (Marsh and Henderson 1994; Marsh et al. 1995; Protozanova and Macgregor 2001, 1996, 2000) (Fig. 4). This feature, which is a nuisance in quadruplexes' programmable programming/design, could be beneficial during abiogenesis when combined with the heating/cooling cycles. For instance, it could serve to purify quadruplexes from impurities such as other nucleotides or L-G enantiomers. This could play a critical role in breaking the chiral symmetry of nucleic acids (see "Origin of Homochirality").

Shortcomings of RNA World Hypothesis

Polymerization of the Very First Polynucleotide

There are two main barriers to polymerization of the RNA template: hydrolyses (Deamer and Weber 2010; Runnels et al. 2018) and cyclization (Cairns-Smith 1977; 2008; Ferris et al. 1996; Horowitz et al. 2010). Both barriers are removed in the quadruplex world. The hydrolysis problem is discussed in "[The Temperature Cycles Can Overpower Detrimental Hydrolysis](#)".

Specific experimental conditions have been created to overcome the problem of cyclization, albeit with limited success. For instance, polymerization on mineral surfaces (Cairns-Smith 1977, 2008; Ferris et al. 1996) or polymerization in the presence of intercalators (Horowitz et al. 2010) allows the formation of RNA strands up to 100-nt in length. As shown in Fig. 4, the quadruplex world easily overcomes the problem: pG₁₀ assembles nanowires that, upon EDC-activated ligation, form polymers of thousands of nucleotides in length (Kanki 2021). The ligation is very efficient, fast, and can be conducted even at the boiling point.

Origin of Homochirality

Another problem of the RNA world is enantiomeric cross-inhibition, whereby template-directed polymerization involving one enantiomer of RNA is inhibited by the presence of the other enantiomer (Joyce et al. 1987, 1984). While chiral selectivity of the replication was striking (poly(_D-C)-directed oligomerization of _D-G was far more efficient than of _L-G), in reactions with a racemic mixture, monomers of the opposite handedness to the template, _L-G, strongly inhibit polymerization by incorporating as chain terminators (Joyce et al. 1984). It must be noted that in addition to the chimeric products, a significant amount of short oligo(_D-G)s were found even after seven days of reaction, but it was suggested that "after extensive oligomerization, nearly all chains would have an _L-G residue at the 2'(3') terminus" (Joyce et al. 1984).

The quadruplex world can employ temperature cycles to break chiral symmetry. It was shown that G-monomers (nucleosides and GMPs) are capable of forming G-tetrad stacks with the helical properties of right-handed quadruplexes (30° rotation and 3.4 Å rise) (Chen et al. 2020; Wu and Kwan 2009; Zimmerman 1976). It must be emphasized that these helical stacks are not constrained by a covalently bonded backbone and they are free to form both left- and right-handed structures, but they favor the right-handed helical stacks. Therefore, adjacent _D-Gs in the G-tetrad stacks can ligate to each other through the 5'-3' phosphodiester backbone with thermodynamically favorable *anti*-glycosidic bonds. The _L-G monomers incorporated in the right-handed helix need to adopt *syn* conformation to have 5'-3' orientation (Joyce et al. 1984). However, this conformation is significantly less favorable for DNA quadruplexes and almost impossible for RNA quadruplexes (Fay et al. 2017). Therefore, _L-G monomers would terminate polymerization by (i) not ligating to oligo(_D-G)s or (ii) ligating and capping them at the 3' terminus (see chimeric oligos in Fig. 5). Thus, the system would contain some amount of oligo(_D-G)s with proper backbone linkages, which can be dissociated from the mixed stacks at the heating step and reassembled into the thermodynamically more favorable homogenous G-wires at the cooling step (Fig. 5).

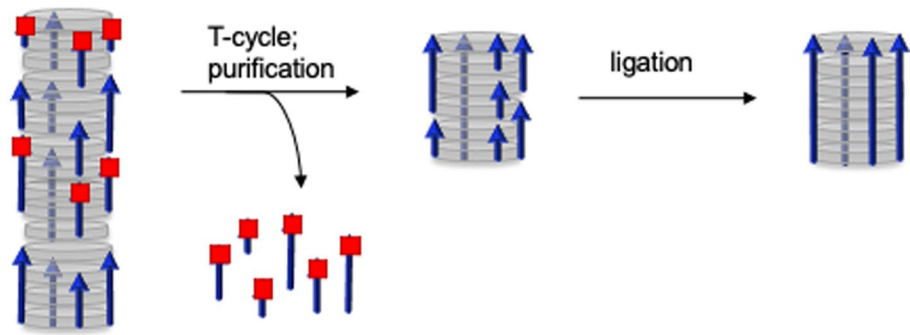


Fig. 5 Purification of quadruplexes by the heating/cooling cycles (T-cycles). Separation of oligo(D-G)s (blue arrows) from chimeric oligos (containing both D-Gs and L-Gs) (blue lines with red squares) and formation of right-handed G-wires followed by ligation

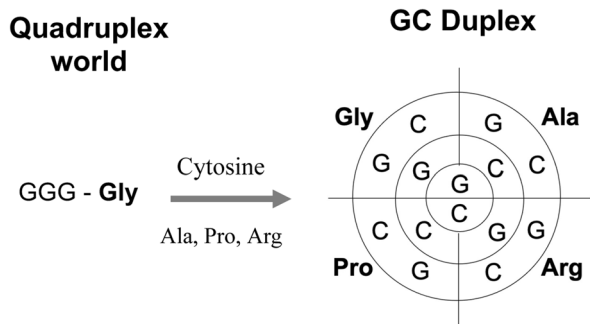
RNA-to-DNA Transformation

The RNA world hypothesis postulates that life was initiated by RNA, which later was transformed into a DNA/RNA world by step-by-step incorporation of deoxynucleotides into RNA duplexes. However, RNA and DNA double helices are significantly different (A- vs. B-conformation), which raises significant thermodynamic penalties upon RNA-to-DNA and DNA-to-RNA transformations (Gavette et al. 2016). In contrast, the DNA and RNA quadruplexes fold into exactly the same tertiary structure, making the transformations in both directions free from significant thermodynamic penalties (see "RNA Analogue of G3T has the Same Topology") (Kankia 2019; Zhou et al. 2017).

The Quadruplex World and Origin of the Genetic Code

One significant deficiency of the RNA world hypothesis is that it represents an isolated system incapable of evolving into modern day ribosome-based decoding and translation (Bowman et al. 2015; Koonin 2012; Orgel 2003). It is also unclear how the genetic code emerged and evolved in the RNA world.

Fig. 6 Transition of the quadruplex world into GC-containing duplex system



 Springer

The continuity principle assumes incremental expansion of a system while preserving earlier versions of the system and their functionality. According to this principle, earlier versions of the genetic code should be preserved in the present universal code (Hartman 1975; Jukes 1974; Kocherlakota and Acland 1982). The most prominent earlier version is the fully degenerated part of the universal code in which the type of nucleotides in the third position of base triplets does not play any role. There are eight amino acids (Gly, Ala, Pro, Arg, Ser, Val, Leu and Thr) involved in this code. Remarkably, all amino acids, except positively charged Arg, are the smallest and simplest, and the most frequent nucleotides in the meaningful first and second codon positions, are C and G bases (44% C, 31% G, 19% U and only 6% A). Based on the simplicity of the amino acids and the nucleotide frequency, the step-by-step evolution of the degenerated code was suggested earlier (Hartman 1975; Jukes 1974; Kocherlakota and Acland 1982), which is consistent with the quadruplex world hypothesis.

The code initiates as the interaction between polyglycine and polyguanine (Hartman 1975). Glycine is not just the simplest amino acids having a single hydrogen atom as its side chain, but it is the only achiral amino acid, which eliminates the possibility of enantiomeric cross-inhibition upon polymerization (see "Origin of Homochirality"). Glycine could self-polymerize, fold, and then interact with tmDNA, or it could use tmDNA as a scaffold for polymerization. The polymerization might take place at the surface of the structured tetrahelix or at exposed bases of single-stranded polyguanines. Interestingly, glycine appears to destroy stacks of G-tetrads, while alanine has almost no effect (Detellier and Laszlo 1980). This supports the specific interaction of glycine with guanine bases at the atomic groups involved in G-tetrad formation. Thus, I propose that non-hereditary polyguanine and achiral polyglycine coexisted at the initial stage of the molecular evolution and represented the first nucleic acid-protein systems.

The next step of the code expansion involved adding cytosine and three chiral amino acids (Ala, Pro, and Arg) (Hartman 1975; Kocherlakota and Acland 1982) (Fig. 6). This was accompanied by introducing the G•C complementarity principle, and only at this point, the nucleic acids become the hereditary/genetic material. The code expansion initiated a transition from the quadruplex world to double-helical nucleic acids, which requires addressing the following questions.

First, what was the mechanism of cytosine incorporation into polyguanines and what was the sequence of the first nucleic acid template? The cytosine bases could have been incorporated into polyguanines by conjugating at the abasic sites formed by spontaneous depurination of loop nucleotides (see "Exposed Bases"). The resulting strand would have

Fig. 8 Impact of frameshifting on the polypeptides produced by $(G_3CG_3CG_6C)_n$ **A** and $(G_3C)_n$ **B**

A Frameshift on tmDNA pattern

Ala Gly Gly Gly Arg Ala Gly Gly Gly Arg Ala Gly Gly Gly (AGGGR)_n
 GCGGGCGGGGGCGGGCGGGCGGGGGCGGGGGCGGGGGCGGGGG
 Arg Ala Gly Gly Gly Arg Ala Gly Gly Gly Arg Ala Gly Gly (AGGGR)_n
 GCGGGCGGGGGCGGGCGGGCGGGGGCGGGGGCGGGGGCGGGGG
 Gly Arg Gly Ala Gly Gly Arg Gly Ala Gly Gly Arg Gly Ala (AGGGR)_n
 GCGGGCGGGGGCGGGCGGGCGGGGGCGGGGGCGGGGGCGGGGG

B Frameshift in $(G_3C)_n$ pattern

Gly Arg Ala Gly Gly Arg Ala Gly Gly Arg Ala Gly (GRAG)_n
 GGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG
 Gly Gly Arg Ala Gly Gly Arg Ala Gly Gly Arg Ala Gly (GRAG)_n
 GGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG
 Ala Gly Gly Arg Ala Gly Gly Arg Ala Gly Gly Arg Ala Gly (GRAG)_n
 GGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG

a specific sequence dictated by tmDNA structure: $(G_3CG_3CG_6C)_n$ (see Fig. 7A). The replication of this sequence (probably in the presence of cations that are unfavorable for quadruplex formation, such as Li^+ or Cs^+) would produce the first double-helical polymer (Fig. 7A, bottom duplex). Another version of the first double helix could have been $(G_3C)_n$. As mentioned in "Role of Temperature Cycles in Quadruplex Formation", temperature cycles could result in untimely termination of the replication process. Since the sequence contains short repeating G_3C units, some replicons of the previous cycles could bind to the template at various places and continue elongation during subsequent cycles (Fig. 7B). As a result, $(C_3G)_n$ and $(G_3C)_n$ sequences would produce long repetitive duplexes (Fig. 7B). Interestingly, the majority of telomeric sequences represent GGG segments separated by a few nucleotides and might represent vestiges of $(G_3C)_n$ sequences. Frameshifting affects the polypeptide sequence encoded by $(G_3CG_3CG_6C)_n$, while it has no effect in the case of $(C_3G)_n$ which is always translated into $(Gly-Arg-Ala-Gly)_n$ (Fig. 8). *Second, how the chiral homogeneity of quadruplexes is maintained in the context of GC-containing duplexes?* As mentioned above, cytosine bases could have been inserted into all-G sequences by conjugating at already existing D -sugar-phosphate groups of polyguanines at abasic sites. However, the template-directed polymerization (see Fig. 7B) could proceed through association of free bases to the template followed by conjugation to D -sugar-phosphate chains produced by depurination of polyguanines.

Third, since amino acids Ala, Pro, and Arg are chiral, what is the mechanism that ensures incorporation of only L -amino acids in polypeptides? It was shown earlier that D -RNA prefers interaction with L -amino acids (Root-Bernstein 2007; Seligmann 2020).

Acknowledgements This work was supported by the Shota Rustaveli National Science Foundation (Grant FR17_140). I thank Tamari Kankia for proofreading the manuscript and Karin Musier-Forsyth for proofreading, helpful comments and discussion.

References

- Anastasi C, Buchet FF, Crowe MA, Parkes AL, Powner MW, Smith JM, Sutherland JD (2007) RNA: prebiotic product, or biotic invention? *Chem Biodivers* 4:721–739. <https://doi.org/10.1002/cbdv.200790060>
- Bowman JC, Hud NV, Williams LD (2015) The ribosome challenge to the RNA world. *J Mol Evol* 80:143–161. <https://doi.org/10.1007/s00239-015-9669-9>
- Cairns-Smith AG (1977) Takeover mechanisms and early biochemical evolution. *Biosystems* 9:105–109. [https://doi.org/10.1016/0303-2647\(77\)90019-3](https://doi.org/10.1016/0303-2647(77)90019-3)
- Cairns-Smith AG (2008) Chemistry and the missing era of evolution. *Chemistry* 14:3830–3839. <https://doi.org/10.1002/chem.200701215>
- Chen J, Liu X, Suo Z, Gao C, Xing F, Feng L, Zhao C, Hu L, Ren J, Qu X (2020) Right-/left-handed helical G-quartet nanostructures with full-color and energy transfer circularly polarized luminescence. *Chem Commun (Camb)* 56:7706–7709. <https://doi.org/10.1039/d0cc02449j>
- Costanzo G, Pino S, Ciciriello F, Di Mauro E (2009) Generation of long RNA chains in water. *J Biol Chem* 284:33206–33216. <https://doi.org/10.1074/jbc.M109.041905>
- Deamer D, Weber AL (2010) Bioenergetics and life's origins. *Cold Spring Harb Perspect Biol* 2:a004929. <https://doi.org/10.1101/cshperspect.a004929>
- DeLuca M, Shi Z, Castro CE, Arya G (2020) Dynamic DNA nanotechnology: toward functional nanoscale devices. *Nanoscale Horizons* 5:182–201. <https://doi.org/10.1073/pnas.0335459100>
- Detellier C, Laszlo P (1980) Role of Alkali Metal and Ammonium Cations in the Self-Assembly of the 5'-Guanosine Monophosphate Dianion. *J Am Chem Soc* 102:1135–1141. <https://doi.org/10.1016/j.autrev.2006.09.003>

- Do NQ, Lim KW, Teo MH, Heddi B, Phan AT (2011) Stacking of G-quadruplexes: NMR structure of a G-rich oligonucleotide with potential anti-HIV and anticancer activity. *Nucleic Acids Res* 39:9448–9457. <https://doi.org/10.1093/nar/gkr539>
- Fay MM, Lyons SM, Ivanov P (2017) RNA G-Quadruplexes in Biology: Principles and Molecular Mechanisms. *J Mol Biol* 429:2127–2147. <https://doi.org/10.1016/j.jmb.2017.05.017>
- Ferris JP, Hill AR Jr, Liu R, Orgel LE (1996) Synthesis of long prebiotic oligomers on mineral surfaces. *Nature* 381:59–61. <https://doi.org/10.1038/381059a0>
- Gavette JV, Stoop M, Hud NV, Krishnamurthy R (2016) RNA-DNA Chimeras in the Context of an RNA World Transition to an RNA/DNA World. *Angew Chem Int Ed Engl* 55:13204–13209. <https://doi.org/10.1002/anie.201607919>
- Hartman H (1975) Speculations on the evolution of the genetic code. *Orig Life* 6:423–427. <https://doi.org/10.1007/BF01130344>
- Heffernan R, Paliwal K, Lyons J, Dehzangi A, Sharma A, Wang J, Sattar A, Yang Y, Zhou Y (2015) Improving prediction of secondary structure, local backbone angles, and solvent accessible surface area of proteins by iterative deep learning. *Sci Rep* 5:11476. <https://doi.org/10.1038/srep11476>
- Horowitz ED, Engelhart AE, Chen MC, Quarles KA, Smith MW, Lynn DG, Hud NV (2010) Intercalation as a means to suppress cyclization and promote polymerization of base-pairing oligonucleotides in a prebiotic world. *Proc Natl Acad Sci U S A* 107:5288–5293. <https://doi.org/10.1073/pnas.0914172107>
- Jing N, Hogan ME (1998) Structure-activity of tetrad-forming oligonucleotides as a potent anti-HIV therapeutic drug. *J Biol Chem* 273:34992–34999
- Joachimi A, Benz A, Hartig JS (2009) A comparison of DNA and RNA quadruplex structures and stabilities. *Bioorg Med Chem* 17:6811–6815. <https://doi.org/10.1016/j.bmc.2009.08.043>
- Joyce GF (2002) The antiquity of RNA-based evolution. *Nature* 418:214–221. <https://doi.org/10.1038/418214a>
- Joyce GF, Schwartz AW, Miller SL, Orgel LE (1987) The case for an ancestral genetic system involving simple analogues of the nucleotides. *Proc Natl Acad Sci U S A* 84:4398–4402. <https://doi.org/10.1073/pnas.84.13.4398>
- Joyce GF, Visser GM, van Boeckel CA, van Boom JH, Orgel LE, van Westrenen J (1984) Chiral selection in poly(C)-directed synthesis of oligo(G). *Nature* 310:602–604. <https://doi.org/10.1038/310602a0>
- Jukes TH (1974) On the possible origin and evolution of the genetic code. *Orig Life* 5:331–350
- Kankia B (2014) Tetrahelical monomolecular architecture of DNA: a new building block for nanotechnology. *J Phys Chem B* 118:6134–6140. <https://doi.org/10.1021/jp503276q>
- Kankia B (2015) Quadruplex-and-Mg(2+) Connection (QMC) of DNA. *Sci Rep* 5:12996. <https://doi.org/10.1038/srep12996>
- Kankia B (2018) Monomolecular tetrahelix of polyguanine with a strictly defined folding pattern. *Sci Rep* 8:10115. <https://doi.org/10.1038/s41598-018-28572-x>
- Kankia B (2019) Stability Factors of the Parallel Quadruplexes: DNA Versus RNA. *J Phys Chem B* 123:1060–1067. <https://doi.org/10.1021/acs.jpcc.8b11559>
- Kankia B (2021) Quadruplex-Templated and Catalyzed Ligation of Nucleic Acids. *ChemBioChem* 22:1261–1267. <https://doi.org/10.1002/cbic.202000754>
- Kankia B, Gvarjaladze D, Rabe A, Lomidze L, Metreveli N, Musier-Forsyth K (2016) Stable Domain Assembly of a Monomolecular DNA Quadruplex: Implications for DNA-Based Nanoswitches. *Biophys J* 110:2169–2175. <https://doi.org/10.1016/j.bpj.2016.04.031>
- Kelley S, Boroda S, Musier-Forsyth K, Kankia BI (2011) HIV-integrase aptamer folds into a parallel quadruplex: a thermodynamic study. *Biophys Chem* 155:82–88. <https://doi.org/10.1016/j.bpc.2011.03.004>
- Kocherlakota RR, Acland ND (1982) Ambiguity and the evolution of the genetic code. *Orig Life* 12:71–80. <https://doi.org/10.1007/BF00926913>
- Koonin EV (2012) The logic of chance : the nature and origin of biological evolution (Pearson Education).
- Marsh TC, Henderson E (1994) G-wires: self-assembly of a telomeric oligonucleotide, d(GGGGTTGGGG), into large superstructures. *Biochemistry* 33:10718–10724
- Marsh TC, Vesenska J, Henderson E (1995) A new DNA nanostructure, the G-wire, imaged by scanning probe microscopy. *Nucleic Acids Res* 23:696–700
- Mergny JL, Sen D (2019) DNA Quadruple Helices in Nanotechnology. *Chem Rev* 119:6290–6325. <https://doi.org/10.1021/acs.chemrev.8b00629>
- Morasch M, Mast CB, Langer JK, Schilcher P, Braun D (2014) Dry polymerization of 3',5'-cyclic GMP to long strands of RNA. *ChemBioChem* 15:879–883. <https://doi.org/10.1002/cbic.201300773>
- Orgel LE (2003) Some consequences of the RNA world hypothesis. *Orig Life Evol Biosph* 33:211–218. <https://doi.org/10.1023/a:1024616317965>

- Orgel LE (2004) Prebiotic chemistry and the origin of the RNA world. *Crit Rev Biochem Mol Biol* 39:99–123. <https://doi.org/10.1080/10409230490460765>
- Pauling L, Corey RB (1953) A Proposed Structure For The Nucleic Acids. *Proc Natl Acad Sci U S A* 39:84–97
- Pino S, Ciciriello F, Costanzo G, Di Mauro E (2008) Nonenzymatic RNA ligation in water. *J Biol Chem* 283:36494–36503. <https://doi.org/10.1074/jbc.M805333200>
- Protozanova E, Macgregor R Jr (2001) Formation of DNA frayed wires is independent of the directionality of the parent strand. *Biopolymers* 58:355–358. [https://doi.org/10.1002/1097-0282\(20010405\)58:4%3c355::AID-BIP1012%3e3.0.CO;2-#](https://doi.org/10.1002/1097-0282(20010405)58:4%3c355::AID-BIP1012%3e3.0.CO;2-#)
- Protozanova E, Macgregor RB Jr (1996) Frayed wires: a thermally stable form of DNA with two distinct structural domains. *Biochemistry* 35:16638–16645. <https://doi.org/10.1021/bi960412d>
- Protozanova E, Macgregor RB Jr (2000) Thermal activation of DNA frayed wire formation. *Biophys Chem* 84:137–147
- Rachwal PA, Brown T, Fox KR (2007) Sequence effects of single base loops in intramolecular quadruplex DNA. *FEBS Lett* 581:1657–1660
- Root-Bernstein R (2007) Simultaneous origin of homochirality, the genetic code and its directionality. *BioEssays* 29:689–698. <https://doi.org/10.1002/bies.20602>
- Runnels CM, Lanier KA, Williams JK, Bowman JC, Petrov AS, Hud NV, Williams LD (2018) Folding, Assembly, and Persistence: The Essential Nature and Origins of Biopolymers. *J Mol Evol* 86:598–610. <https://doi.org/10.1007/s00239-018-9876-2>
- Salditt A, Keil LMR, Horning DP, Mast CB, Joyce GF, Braun D (2020) Thermal Habitat for RNA Amplification and Accumulation. *Phys Rev Lett* 125:048104. <https://doi.org/10.1103/PhysRevLett.125.048104>
- Seeman NC (2010) Nanomaterials based on DNA. *Annu Rev Biochem* 79:65–87. <https://doi.org/10.1146/annurev-biochem-060308-102244>
- Seligmann H (2020) First arrived, first served: competition between codons for codon-amino acid stereochemical interactions determined early genetic code assignments. *Naturwissenschaften* 107:20. <https://doi.org/10.1007/s00114-020-01676-z>
- Smith GP, Fraccia TP, Todisco M, Zanchetta G, Zhu C, Hayden E, Bellini T, Clark NA (2018) Backbone-free duplex-stacked monomer nucleic acids exhibiting Watson-Crick selectivity. *Proc Natl Acad Sci U S A* 115:E7658–E7664. <https://doi.org/10.1073/pnas.1721369115>
- Tinoco I Jr, Bustamante C (1999) How RNA Folds *J Mol Biol* 293:271–281. <https://doi.org/10.1006/jmbi.1999.3001>
- Wu G, Kwan IC (2009) Helical structure of disodium 5'-guanosine monophosphate self-assembly in neutral solution. *J Am Chem Soc* 131:3180–3182. <https://doi.org/10.1021/ja809258y>
- Zhou J, Amrane S, Rosu F, Salgado GF, Bian Y, Tateishi-Karimata H, Largy E, Korkut DN, Bourdoncle A, Miyoshi D et al (2017) Unexpected Position-Dependent Effects of Ribose G-Quartets in G-Quadruplexes. *J Am Chem Soc* 139:7768–7779. <https://doi.org/10.1021/jacs.7b00648>
- Zimmerman SB (1976) X-ray study by fiber diffraction methods of a self-aggregate of guanosine-5'-phosphate with the same helical parameters as poly(rG). *J Mol Biol* 106:663–672