Systems Chemistry in the Chemical Origins of Life: The 18th Camel Paradigm

Ramanarayanan Krishnamurthy[1*]

ABSTRACT

Developing an understanding of the prebiotic chemical *Origins of Life* has long been defined and dominated by various biology-driven hypotheses, such as the RNA-first world or metabolism-first world or lipid-first world, or protein-first world, each of which then gave rise to the other classes of compounds and, then, set the stage for the appearance of life. These '*one-after-the-other*' concepts are a result of a 'top-down' view that extrapolates extant biological paradigms and processes, linearly backward in time, relying largely on molecular phylogenetic analysis for clues to seek evolutionary historical relationships in biochemical pathways. As a result, the focus has been *only* on the chemical origins of the biological building blocks of RNA, proteins, metabolites *etc*., at the expense of ignoring the roles of other prebiotically relevant molecules. However, in recent times, this one-dimensional reductionist thinking has been slowly, but surely, challenged by the influence of *Systems Chemistry* in prebiotic chemistry – leading to a 'bottom-up' co-existence and co-evolution of various molecules that can give rise to a (dynamic) network of interacting entities capable of chemical evolution. This personal review describes how our research program, which was once driven by the reductionist-linear approaches, has been increasingly influenced by the principles and paradigms of Systems Chemistry. And how it, in turn, has led to fundamental changes in our approaches to investigating the chemical Origins of Life by considering alternative prebiotic molecules and chemistries that may have played a role in getting the prebiotic chemistry started, but have not been retained in their original forms in extant biology. This has led to a "*18th camel paradigm*" in our research – one that is providing unconventional venues and alternative perspectives to comprehending some of the long-standing issues in this field.

Keywords: origin of life, depsipeptides, DNA, peptides, prebiotic Chemistry, RNA, Systems Chemistry.

Introduction

The field of *Origins of Life* – particularly referring to how inanimate chemicals can be transformed to a collection of supramolecular assemblies that begin to exhibit properties that can be associated with life – has been primarily driven by our understanding of the chemical origins and behavior of biopolymers and bio-assemblies of extant biology [1 - 7]. How biological molecules are synthesized, transformed and function has been used as a guide, for providing clues as to what may have happened billions of years ago on the early Earth. Since biology relies on RNA(DNA), proteins,

*Correspondence: rkrishna@scripps.edu
[1]Department of Chemistry, The Scripps Research Institute, La Jolla, CA 92037, USA

lipids, and metabolites, each of these classes of molecules has been used as candidates to understand the *Origins of Life* as we know it (*Fig. 1*). This has been a reasonable hypothesis-driven reductionist strategy, and has given rise to many different approaches in developing an understanding of how the molecular processes of life's chemistry could have started on the early Earth. One early conjecture was the "protein world" that posited the formation of long-chain peptides and aggregates (formed from amino acids) as the first entity, which led to the advent of other biomolecules [8 - 10]. The RNA-world hypothesis, developed based on the ideas of RNA encoding for protein synthesis [11 - 14], postulated that RNA molecules were first synthesized, self-replicated and then gave rise to the proteins, DNA and so on [15 - 18]. However, there were some modifications to this postulate in the subsequent years which called for ancestors of RNA, called pre-RNA or proto-RNA, which later gave rise to RNA [19 - 21]. In the interim, there was a metabolism-first approach [22 - 27], wherein a collection of molecules and their conversions in a net-work like setting, laid the groundwork for forming the building blocks of RNA, proteins, lipids *etc.* [28 - 31]. Later, it was the lipid world wherein the original lipid- dominated supramolecular assemblies gave rise later to other biopolymers [32, 33].

Thus, each class of biomolecules of extant life was given the central and primal role, with the others following naturally – relying on existing biological chemistries as historical pathways (supported by molecular phylogenetic analysis) to explain how they would proceed to give rise to life. For example, in the RNA-only world approach, it was enough that one found a prebiotically plausible pathway to the nucleosides and nucleotides of RNA – then one could claim that the "RNA world" would take over, and *Darwinian* evolution and selection would give rise to life, because RNA would give rise to proteins, which would then give rise to DNA and so forth (*Fig. 1*). This "one-after-the-other" paradigm

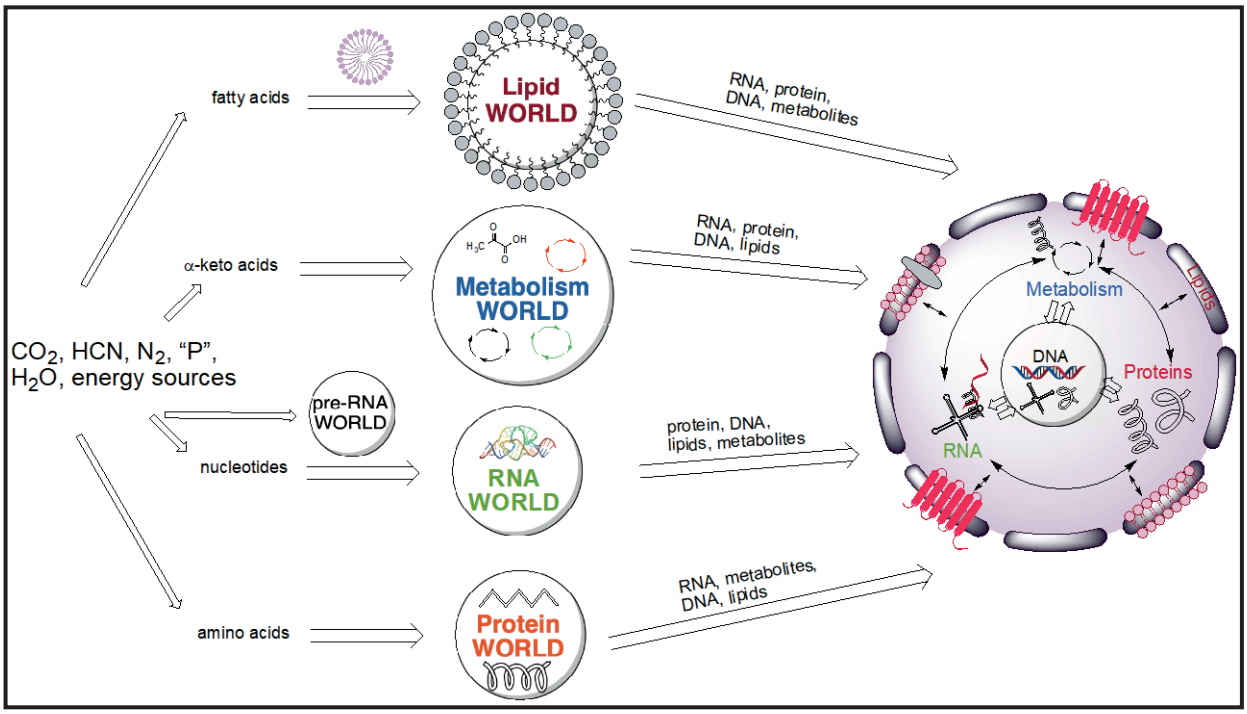

Figure 1. The Conventional "Me-First" Hypotheses in Origins of Life. Historically, the *Origins of Life* field has been fragmented based on the primary biopolymers and biomolecules of extant life, wherein each of these classes of compounds were proposed to have appeared first and then invented the other classes of biomolecules at a later stage.

was the hall mark of each of these "me-first" camps, and had their compelling reasons rooted in how biology synthesized the building blocks of the biopolymers [34]. It was the RNA world, which pretty much came to dominate the scene and soon became the 'the team to beat' [35, 36]. This *status quo* would face challenges now and then either from an experimental demonstration of a spectacular peptide replication [37], or from autocatalyzed self-growth of fatty acid vesicles [33], or from formation of compounds important for metabolism [38], and hypothesis supported mostly by hydrothermal observations [39 - 42] – however, the RNA world eminence was never seriously challenged. The reason was the dominant role of RNA in biology (giving rise to proteins and DNA) coupled with the conviction that ribosome is a 'ribozyme' [21, 43], and the intriguing demonstrations of the ever expanding repertoire of RNA-catalyzed reactions [44 - 46], of which the most important is the potential for the almost unfettered self-replication [47].

But, with the advent of *Systems Chemistry* [48] in 2005 – a concept that was an outcome of the work of *Günter von Kiedrowski* dealing with replication of nucleic acids [49, 50] – all of the research paradigms mentioned in the *Introduction* have been impacted in one way or the other and have become more sensitized to consider the interactions between various classes of molecules, and not in isolation as they were originally envisioned. The goal of this personal review is to document how *Prebiotic Systems Chemistry* (for the purpose of this review, we use the working definition: "*the emerging covalent and non-covalent interactions between molecules /supramolecules, leading to a network of reaction pathways and transformations with potential for feedback and chemical evolution*") has altered our thinking and our research program in developing an understanding of the chemical Origins of Life, by highlighting couple of examples from our work. For how the overall field has been impacted, I would point the readers to some comprehensive reviews [51 - 57] and selected individual works [58 - 62], to judge how the concept of Systems Chemistry is taking hold [63] in the Origins of Life field, and to the importance of bringing the separate parts together right from the start ("bottom-up") and not at some later stages of chemical or *Darwinian* evolution.

My Entry into Origins-of-Life Research

When I got introduced (or, should I say, hooked?) to the field of Origins of Life by a lecture from *Albert Eschenmoser,* I was a graduate student at The Ohio State University in early 1990s and unaware of many of these dueling views. *Eschenmoser*'s lecture dealt with the *Chemical Etiology of Nucleic Acid Structure* [64] (*Fig. 2*) and left such an indelible impression on me that it induced me to join the *Eschenmoser* group at ETH, Zurich in 1994, to work on this topic (and have been to this day). Even in the early 1980s, while RNA world was the model to follow, *Eschenmoser* took the unconventional approach of asking "*Why RNA*?" and not "*How RNA?*" [65]. That, in itself, was clear statement that RNA is not the only molecular structure that we should consider, but rather a library of plausible structures derived from the neighborhood of RNA – structures that could be formed by the same chemistries that led to the formation of RNA [66]. This way of reasoning, in my opinion, would be a forerunner to the "*Systems Chemistry*" in a different way, since the preference of RNA from a library of other nucleic acids implied a selection that is based on an emergent property of functions that are manifested only at the level of a polymer (such as base-pairing structures and the catalytic activity). Thus, even though it appeared that the *Eschenmoser* approach could be classified within the RNA world, it differed enough in its character (of asking why RNA) that it stood apart from the usual RNA-only-world approach. As stated elegantly "*Asking the central question regarding the criteria for RNA's natural selection and extending the inquiry to whether its emergence was dominated by combinatorial generation and functional selection or by synthetic contingency could mean to embark on a program of much more comprehensive chemical screening of potentially natural nucleic acid alternatives*" [64].

In this "*why RNA*" approach, the important twist

was also to ask "*why not XNA?*", where X was other sugars in place of ribofuranose, which led *Eschenmoser* to conduct a systematic investigation of the properties of the hexose-sugar backbone-nucleic acids [67], the pentopyranosyl nucleic acids (*e.g.*, p-RNA) [68, 69], and later of the threofuranosyl nucleic acids (TNA) [70, 71] (*Fig. 2*). From these studies, the answer to the question "*why RNA*?" was that the selection of RNA was based on the "*optimal (and not maximizing the) base-pairing properties*" [72]. This reference to *an optimum*, in turn, implies an end result of a natural selection from a library of candidates for a particular environmental condition (aqueous medium, near neutral pH, ambient temperature *etc.*) – and has all the connotations of *Systems Chemistry* [73].

The exploration of the properties of XNAs, naturally connected with the idea of a pre-RNA world, which was put forward to overcome the "*nightmare scenario*" of not being able to (at that time) prebiotically synthesize the nucleos(t)ides of RNA [74]. A simpler pre-RNA candidate that would be compatible with prebiotic synthesis would form first and then give rise to RNA. Thus, the threofuranosyl-NA (TNA) was considered to be one such candidate [75]; however, *Eschenmoser* himself had reservations against such a scenario, expressing the difficulties if one were to take this pre-RNA to RNA world transition seriously [76]. Around 2004, the research program at *Scripps*, in its search for the prebiotic contemporaries to RNA, had started investigating dipeptide backbones (in place of sugar-phosphates), that were tagged with

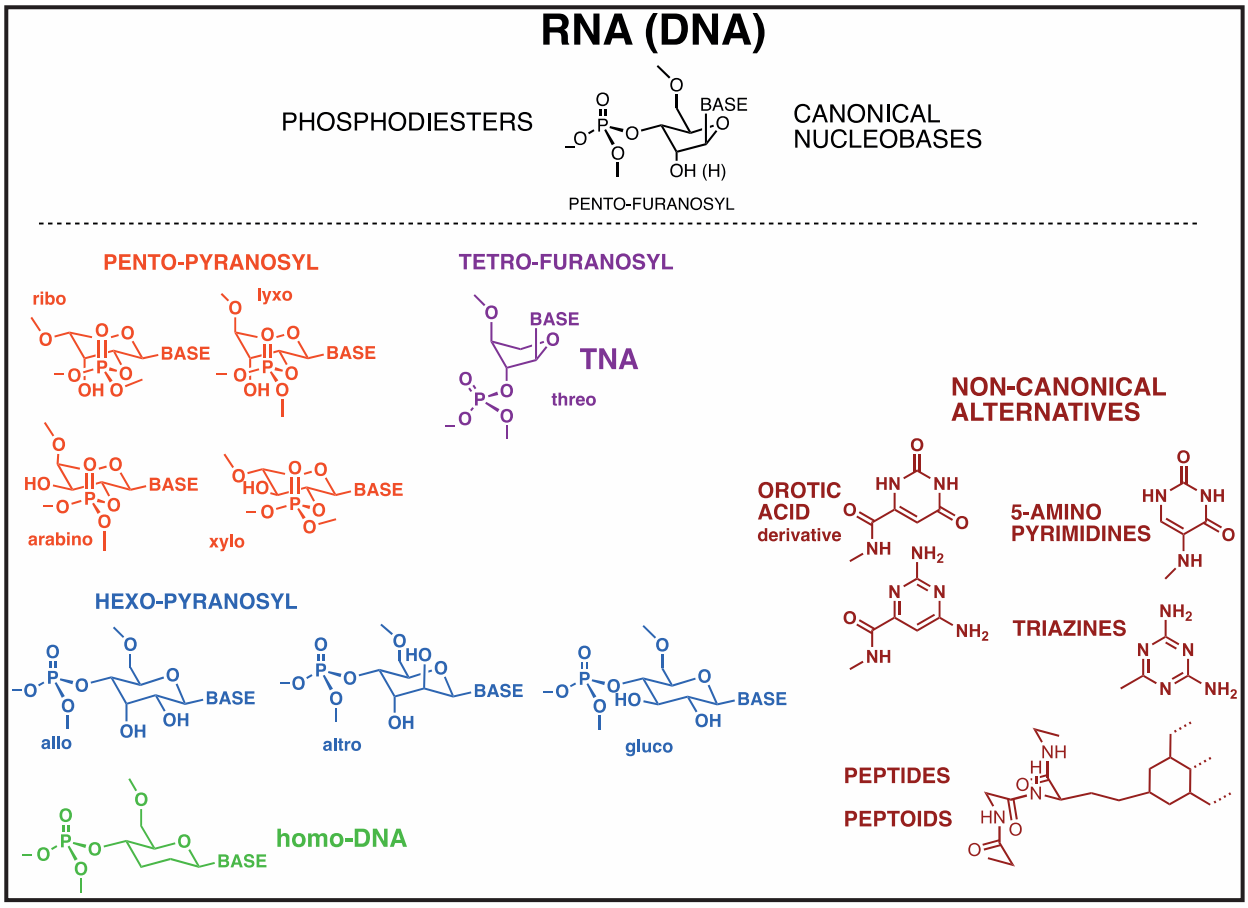

Figure 2. Why RNA and Why not XNA? The chemical etiology of nucleic acid structure pioneered by *Eschenmoser* is a systematic investigation into the base-pairing properties of nucleic-acid alternatives drawn from the structural neighborhood of RNA and was later expanded to include the non-canonical nucleobases and dipeptide backbones.

alternative nucleobases (and not the canonical nucleobases) [77, 78] (*Fig. 2*). The study of these alternative nucleobases led us to consider orotic acid as a plausible prebiotically available nucleobase [79, 80] (*Fig. 2*). The results from the study of the orotic acid-containing oligomers gave us the first glimpse into the problems associated with both an RNA-first or pre-RNA-first approach in trying to understand the Origins of Life's biomolecules [80].

Orotic Acid – Our Entry into Systems Chemistry

Orotic acid is a unique nucleobase amongst all of the nucleobases in biology, in that it is the only nucleobase to be formed *de novo*, in its native heterocyclic form, starting from aspartic acid [81, 82] (*Fig. 3*). While adenine, uracil, guanine, and cytosine are the most famous 'gang of four' in RNA, none of these canonical nucleobases are synthesized in their native form by biological pathways. Once the orotic acid has been synthesized, it is then coupled ('ribosylated') with the 5-phospho-ribosyl-1-pyrophosphate to form the orotidine nucleotide that is decarboxylated to give the uridine derivative which is then converted to the cytidine derivative. The purine nucleobases are built part-by-part on the 5-phosphoribosyl framework [81, 82]. The biosynthetic pathway of the nucleotides of RNA (and not the nucleobases as such) has never been reproduced in a prebiotic setting, and all of the work to date relies on prebiotically plausible chemistries that bear little resemblance to what biology is using today [83, 84]. Even there, the investigations based on the RNA world (or the pre-RNA world) have paid little attention to orotic acid, just for the arbitrary reason that orotidine is not present in the final functional palette of nucleotides in RNA. A majority of the pre-RNA studies changed only the backbone of the sugar-phosphate, while still maintaining the 'gang of four' canonical nucleobases [85, 86]. This omission is strange, since orotic acid is the only 'canonical' nucleobase that is available both by prebiotic (starting from HCN [87-89]) and biotic pathways [81, 82, 87, 88] (*Fig. 3*).

Why Not Orotic Acid in RNA?

We investigated the base-pairing properties of orotidine-containing RNA and found orotidine to be an inferior base-pairing partner, so much so, that even a single incorporation of orotidine in a dodecamer duplex completely destabilized the

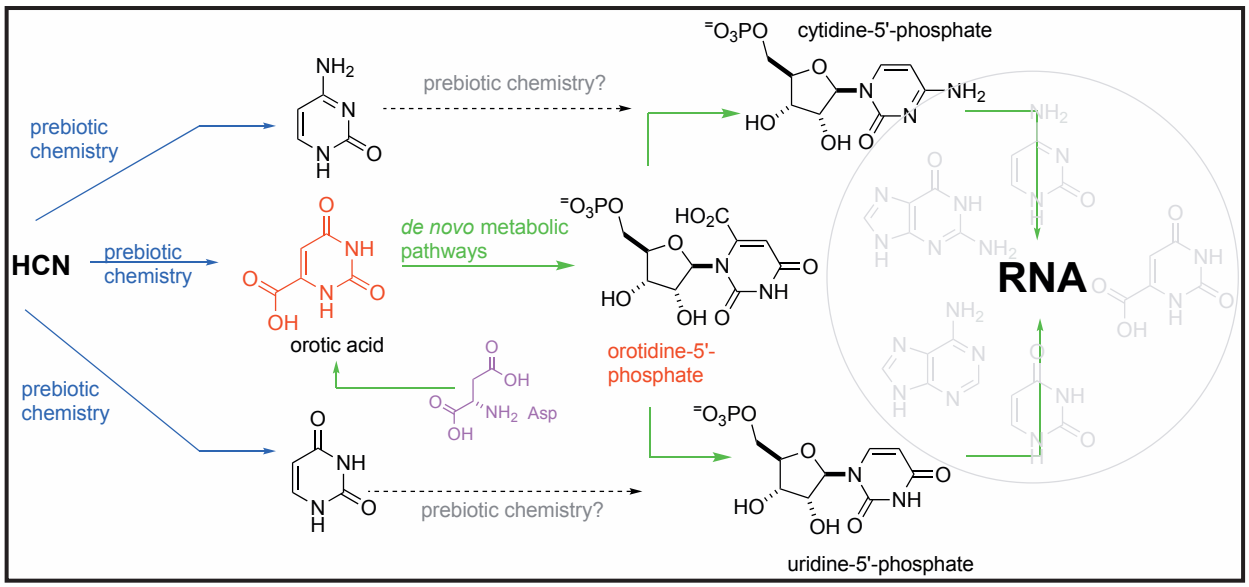

Figure 3. Why not orotic Acid? The presence of orotic acid in *de novo* biological pathways and the formation of orotic acid from HCN seems to suggest a link from the past to the present, a feature that is absent for the other four canonical nucleobases of RNA, *i. e.*, adenine, uracil, guanine, and cytosine.

duplex formation [80]. While this could be rationally interpreted as nature selecting a more functional uracil or cytosine over the non-functional orotic acid, such an interpretation only highlights the issue of ignoring orotic acid in the first place. If one were to accept the premise that uracil and cytosine were selected based on the base-paring capacity, and that orotic acid was discarded on its inability to form base pairs, then the question that naturally arises is: at what level did this selection of nucleobases takes place? If the functional base-pairing property was the criterion, then selection had to have happened at the level of an oligomer in the presence of complementary base-pairing partner. If that is the case, then one cannot make the claim that selection of the four nucleobases could have happened at the level of their prebiotic formation either as the nucleobase or the nucleoside or the nucleotide, because there was no base-pairing property to select for at that monomeric level. This points out to the untenability of using the mere presence of these nucleobases or ribose sugars in meteorites or in a prebiotic synthesis pathway to argue for their presence in RNA. While the formation of parts of RNA is a critical part of the process, it is not the only criteria to state "*… therefore RNA world!*". Even if RNA nucleosides were the only possible outcome in a prebiotic scenario (by any stretch of imagination), even then RNA would have to be selected for, based on its functional capability – and before we say "*Why not*? *Isn't that the case?",* we have to take into consideration that is not a given. If one were to ruminate about what other environmental requirements (solvent, salt concentrations, pH of the medium, temperature *etc*.) are needed for RNA to be functional, one can envision different environments under which RNA would not function, let alone survive. Such a thought process led us realize that selection (and existence of) RNA in biology is not the result of a mere synthesis of its building blocks ('*it could be made*'), but whose selection is based on its functional role and the environment where it can express its function. This realization fits with what *Eschenmoser* had stated in his *Chemical Etiology of Nucleic Acid Structure*, namely, "*Biological reasoning would emphasize that moderate base-pairing strength, as encountered in RNA and resulting from the high conformational flexibility of the ribofuranose backbone, was essential for the evolution of a rich diversity of nucleic acid-related biological functions*" [64]. This realization has to be balanced with the real possibility that the selected product(s) could be very well different if the conditions were different (for, *e.g*., in an exoplanet with alternative environments, solvents, pHs, prebiotic chemistries leading to a different biology) [90, 91].

This experience with orotic acid in RNA provided us with the first taste of the '*Systems Chemistry*' flavor in an *Origins-of-Life* perspective – that the selection of nucleobases, sugars, and the phosphate linker is made at the level of an oligonucleotide system and its emergent functional property, and not at the level of a monomeric building block (nucleobase or nucleoside or nucleotide) that is not capable of expressing that property [34, 91]. Moreover, it also made us aware that there are molecules – that are also available by the same prebiotic chemical pathways and are also used in extant biological pathways – however, do not end up in their original form in the final product [90]. In this case, the orotic acid had been transformed to become the functionally useful uracil and cytosine – and in that sense, orotic acid and orotidine can be considered as a 'pre-RNA'. At the same time, it is important to note that extant biology still has not devised a method to synthesize *de novo* free uracil or cytosine and couple it with 5-phospho-ribosyl-1-pyrophosphate (discounting the salvage pathway) [81, 82].

The "17 Camel Problem"

Around this time, I came across an intriguing anecdotal puzzle called "*the 17 camels and 3 sons*". The problem is depicted in *Fig. 4*, where a father leaves 17 camels to his three sons, and to be divided among them according to his will. It was instructive to read this puzzle and realize that it was not possible to solve the problem according to the 'will of the father' within the parameters of the problem; the mathematics, if followed strictly, would result in an unhappy camel. The difficulty in

solving the problem only with the parameters defined by the puzzle, pointed to the parallels of the conundrum faced in the chemical Origins of Life, wherein, we also have the equivalent of these 17 camels, called the "*pillars of prebiotic chemistry*" [72, 91, 92], such as the formose reaction leading to ribose, HCN chemistry leading to the canonical nucleobases, and the *Strecker* reaction (*Urey-Miller* spark discharge experiments) giving rise to amino acids. Such remarkable and striking connections of products of prebiotic chemistry of simple molecules with molecular building blocks of life is riveting and leads one to conclude that the chemical Origins of Life must be straightforward. To make matters more complicated, the findings of the same sort of building blocks of ribose, amino acids, and nucleobases in meteorites leads to an even more grandiose claim, "*not only life on earth, but in the universe*". Such 'irrational exuberance', while understandable, has not led to any concrete connection between such prebiotic building blocks and the chemical Origins of Life. For example, the mere presence of nucleobases and ribose does not guarantee the formation of nucleosides. Rather, using them as the sole guides has created more complications and led to search for alternative solutions for the formation of RNA nucleosides [82]. Nor has the prebiotic presence of amino acids has led straightforwardly to peptides. The desire to connect one prebiotic data point (ribose, amino acids, and nucleobases in prebiotic chemistry and in meteorites) with the second biological data point (the same molecules in extant life), separated by millions of years by simple extrapolations, has led to an apparent straight line that has misled (and

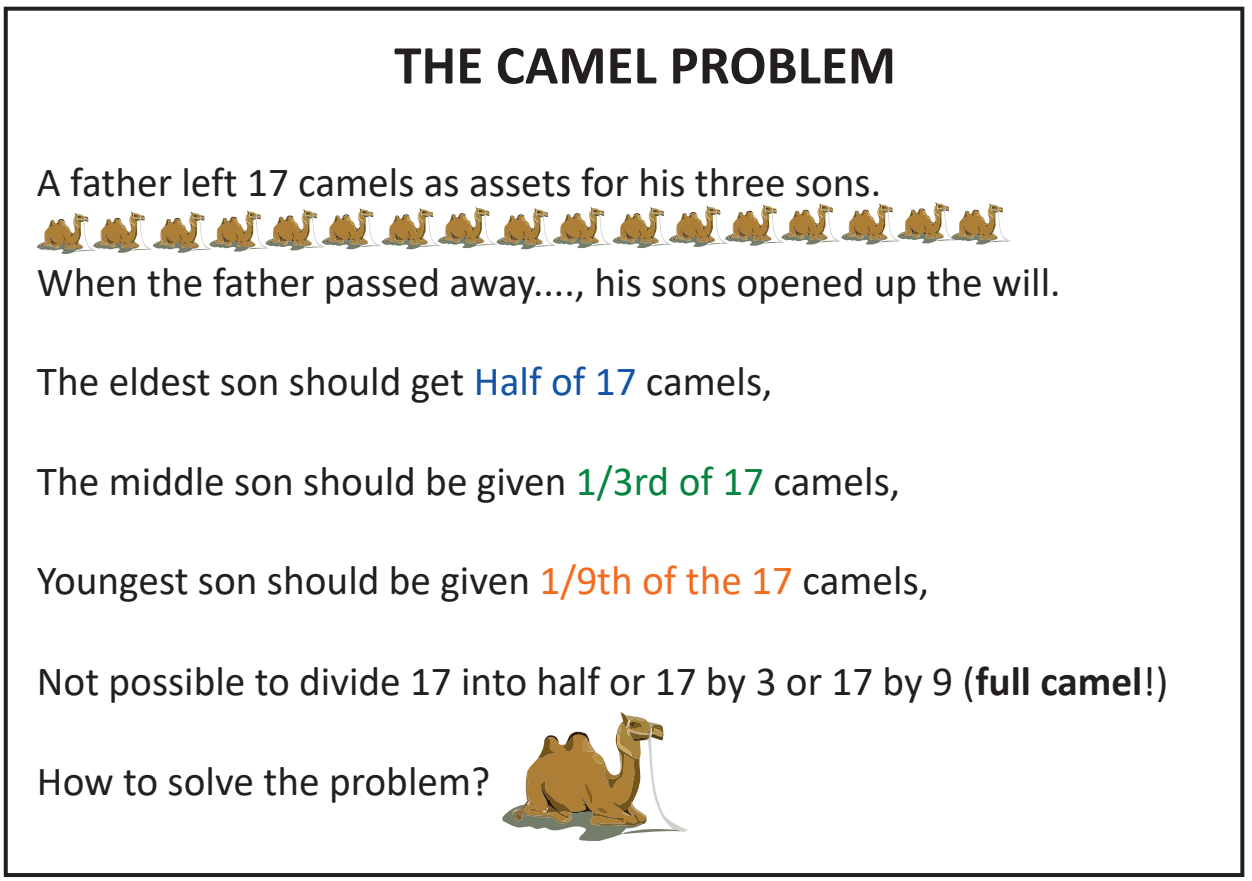

Figure 4. The '17 Camel Problem'. The puzzle shows how difficult it is to solve the problem with only the parameters as defined by the will of the father. This is similar to looking only for the biologically relevant molecules in prebiotic chemistry and trying to forge a (direct and shortest) solution to the chemical Origins of Life problem.

still is misleading) this field. This is where the parallels with the "*the 17 camels and 3 sons*" puzzle becomes apparent – we are trying to solve the Origins of Life problem with what is observed in biology and then limiting the search to only those molecules in prebiotic chemistry, while ignoring the rest of the prebiotic inventory [90].

Before we jump into solving the "*17-camel problem*", I want to introduce the next problem of emergence of polypeptides that showed us how the approach of Systems Chemistry in a prebiotic context could provide solutions that are prebiotically realistic. Understanding the prebiotic pathways of formation of polypeptides starting from only amino acids was a natural consequence of trying to connect the two points: amino acids in prebiotic chemistry with proteins in biology. However, in spite of many attempts there are still problems. First, it is prebiotically unrealistic to start only with pure amino acids as a starting point [94]. Second, many of the types of prebiotic chemical activations used to make peptide bonds would not be the ones that can be ported over to biology. Therefore, on the one hand there is need to focus on biological proteins and investigate how they can be made in prebiotic chemistry, but then the prebiotic chemistry that would make the peptides could not be made compatible with biological processes of how peptides are synthesized. And this is where, the concepts of Systems Chemistry provided an alternative solution that is not only prebiotically plausible, but one that is also portable to biological scenarios. And there, is the logical connection to the "*17 camels and 3 sons*" problem, and how that puzzle is solved.

As seen from the puzzle depicted in *Fig. 4*, it is not possible to solve the problem with the given parameters and have a full camel at the end. As the anecdotal story goes, the three sons unable to solve the problem visit a wise old man who patiently listens to them and declares that he can solve the problem. The wise old man adds one of his own camels to the pack of 17 to make it 18. Then he reads the will of the father and gives ½ of 18 (9 camels) to the first son, 1/3 of 18 (6 camels) to the second son, and the 1/9 of 18 (2 camels) to the third son. If we add up these numbers 9+6+2, it comes out to 17 camels. The remaining camel (which the wise old man brought in the first place) was taken back by the wise old man, thus solving the seemingly unsolvable problem. Some may object that one of the sons has been shortchanged by this solution; but that is better than distributing parts of *a* camel to all three of them! This approach showed how "*outside the box*" thinking can suggest solutions – by adding something else to the problem, make it solvable and then removing what was added. In chemical terms, one would call that a '*catalyst*'! As to how this *18th-camel* solution parallels a solution to the prebiotic peptide synthesis is illustrative of how "*Systems Chemistry*" approach can be very useful in prebiotic chemistry. In other words, are there prebiotically relevant molecules that could have been added at the initial stages that enable a chemical transformation (which otherwise would be difficult) and are later removed (or are converted)?

The α-Hydroxy Acids and the 18th Camel

Around 2011, as part of the collaborative effort in the *Center for Chemical Evolution* (CCE), there was an effort underway, by *Irena Mamajanov* (*Nick Hud*'s group at Georgia Tech), to study the formation of polyesters by a wet-dry cycle starting from hydroxy acids. It was motivated by *Hud*'s idea that polyesters would be able to form secondary structures (for, *e.g.*, by coordinating to metals), develop catalytic capabilities and could act as forerunners of the modern polypeptides. It was around this point that I became involved in the discussion and asked them, "*but how does the polyester give rise to a polypeptide in an evolutionary context*?" [95]. Understandably, this question did not gain much traction, as there were good reasons to pursue the polyester project given the promise of finding catalytic activity. I kept pressing the issue and proposed the following idea to the members of the CCE (*Fig. 5*): why don't we mix amino acids along with the α-hydroxy acids in the wet-dry cycle experiments? It should lead to peptide-bond formation, since the initial esters formed would be attacked by the amino group of

the amino acid to give a thermodynamically stable amide bond. This new dipeptide would still have a free carboxylic acid group which can be esterified again, and a newer amino acid could attack this ester bond. This process, repeated multiple times, would naturally lead to increased peptide bonds in the growing chain, driven by the thermodynamic stability of the amide bond *via* an ester-peptide exchange (*Fig. 5*) [95]. This suggestion, at that time, was based on the following considerations: *a*) α-hydroxy acids are known to easily oligomerize to give polyesters by a simple, prebiotically plausible, wet-dry cycle – a result of the kinetically facile and slightly thermodynamically favored ester-bond formation– as shown by *Weber* [96] and (at that time) the ongoing work of *Irena Mamajanov*, *Nick Hud*, and co-workers. It is also common knowledge from organic chemistry that esters can be converted to amides by reacting with amines (from the pioneering works of *Murray Goodman* [97] and others [98, 99]) driven by the formation of the thermodynamically stable amide bond. Therefore, it was only natural to surmise that if the polyesters could be attacked by amino acids, then it would naturally lead to a peptide bond. *b*) The choice of α-hydroxy acids was based on the fact that these are also formed in copious amounts in the *Urey-Miller* spark discharge experiments [100, 101] and are also found in meteorites [102] (but have been largely ignored because extant biology does not use the oligomers of hydroxy acids [18]).

But, this fact was noticed by *Alex Rich* who stated, "*The large amounts of α-hydroxy acids produced in these experiments means that it is distinctly possible that early polymerizing mechanisms may have resulted in polymers containing both amides and esters*" [103]. It is important to note that this suggestion of a mixed amide-ester polymer by *Rich* was not based on the ester-bond forming first and then converting to an amide bond, but rather based on the presumption of the formation of amide bonds as well as ester bonds during the random co-polymerization of amino acids and hydroxy acids – based on the demonstration that ribosomes are also able to oligomerize α-hydroxy acids into oligoesters [103, 104] and the presumed 'similar reactivity'. As *Rich* stated, "*Thus it is likely that α-hydroxy and α-amino acids were present in the period during which abiogenic polymerization took place. It is possible that many of the polymerization mechanisms may not have differentiated between them since amides and esters are somewhat similar in their reactivity*" [103].

Finally, after a Skype conversation and e-mail exchanges with *Irena* on Dec. 5, 2011, she did the first experiment (later in December 2011) of mixing malic acid with aspartic acid (at 100°C for 4 days), and recorded preliminary ^{1}H- and ^{13}C-NMR data suggestive of a co-polymerization – but this was not pursued further until September 2013 when I asked *Irena* again about the mixture of amino acids and hydroxy acids. But, by then *Irena* was working on the malic acid-oligomerization manuscript [105] and left for her new position. Later, *Sheng-Sheng Yu* (*Martha Grover*'s group) took over and started a systematic investigation and was soon joined by *Jay Forsythe* (*Facundo Fernandez*'s group). This dynamic duo took the project to its logical completion by demonstrating that, indeed, a mixture of α-amino acids and α-hydroxy acids in a simple wet-dry cycling experiment at moderate temperatures formed ester bonds that were converted to depsipeptides with increasing incorporation of amino acids in the growing polymer – forming peptide bonds at the expense of ester bonds, as shown by various analytical techniques [106, 107]. The MS-MS sequence analysis of the depsipeptides clearly showed the continuous enrichment of the sequences with amino-acid residues by replacement of the hydroxy acids through an ester-amide exchange, as the wet-dry cycling was continued. Overall, the polyester was slowly converted to a poly-depsipeptide, with the hydroxy acids being regenerated, as the amino acids were being incorporated [106, 107].

Thus, from a Systems Chemistry point of view, the interaction of mixture of α-hydroxy acids with α-amino acids, by the process of removal of water by drying, naturally led to the appearance of the peptide bond as consequence of *a*) the kinetics of ester bond formation, which then *b*) allowed the attack of the amino group of the amino acid to

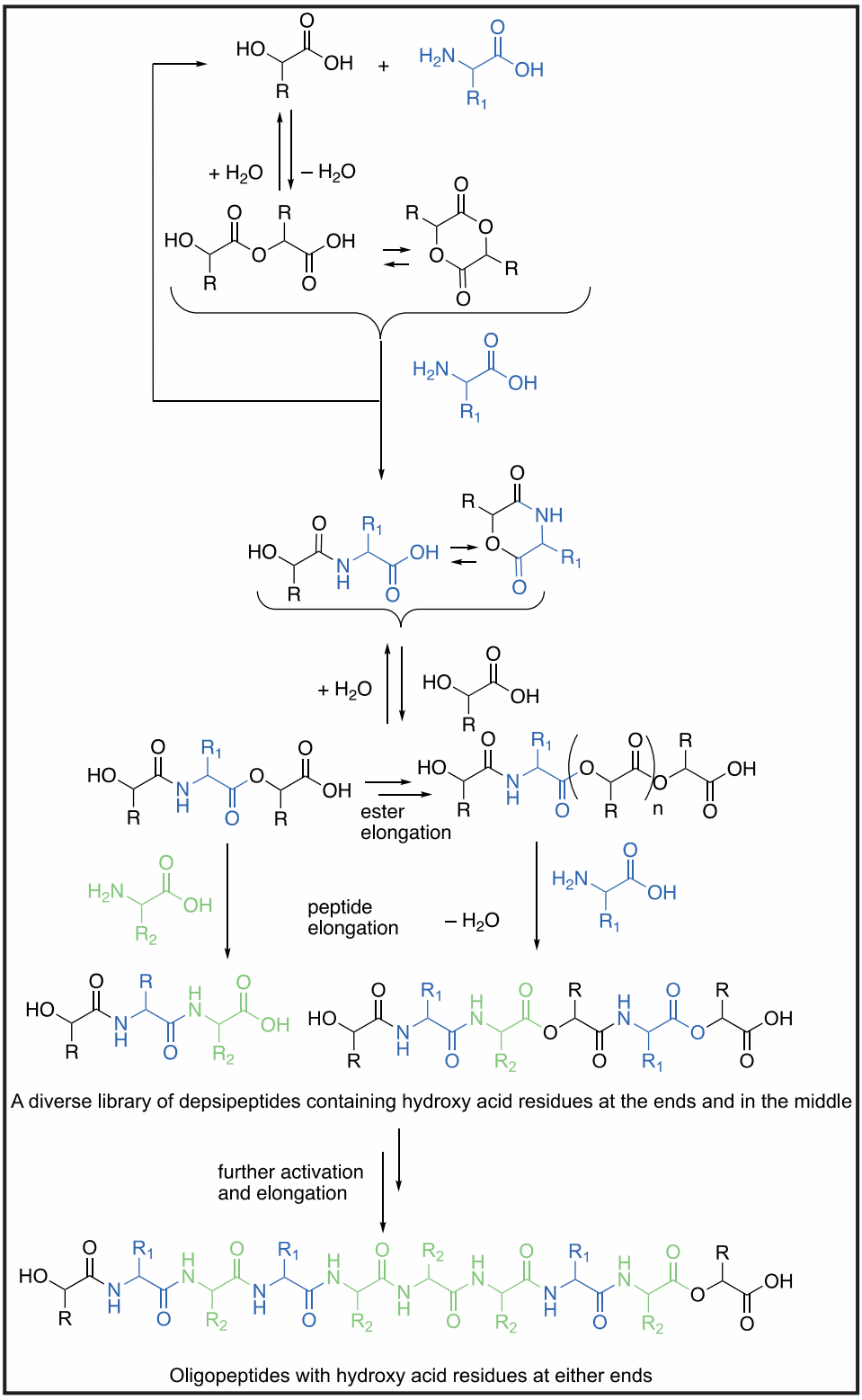

Figure 5. The α-Hydroxy acid acting as an initiator and a catalyst. The concept of α-hydroxy acid (ester-bond) mediated peptide-bond synthesis that illustrates a System Chemistry approach in providing a pathway to peptides under plausible prebiotic scenarios [90, 95].

form the thermodynamically stable amide bond (*Fig. 6*, top). In this process, the α-hydroxy acid was first consumed to form an ester and then was regenerated when the amide bond was formed. In other words, the α-hydroxy acid was catalyzing the peptide-bond formation, and therein lies the connection the "18th camel story"! The *a*-hydroxy acid was the "18th camel" that was added to solve the "17 camel problem" of trying to make peptide bond starting from only amino acids based on the "will of the father". What is more intriguing is that this principle of ester formation, followed by amine attack to make the peptide bond, is exactly the same chemistry that takes place within the ribosome [81, 90], where the amino acids are esterified on the 3'(2')-hydroxy groups of t-RNA, and the free amino moiety of the esterified amino acid on one t-RNA attacks the ester bond on the neighboring t-RNA to form the amide bond (*Fig. 6*, bottom). The same is true for the non-ribosomal peptide synthesis, where a thioester is the central activated species (*Fig. 6*). Such a resemblance of the chemistries both in the abiotic and biotic pathway indicates that it is important (and illuminating) to search for such chemical principles *via* a Systems Chemistry approach, rather than trying to focus only on prebiotic amino acids and biological proteins, and try to understand how the latter would emerge from the former *via* the shortest possible route (*Fig. 6*, top).

The 18th Camel Solution and the RNA World

When I first presented ester-amide exchange work,

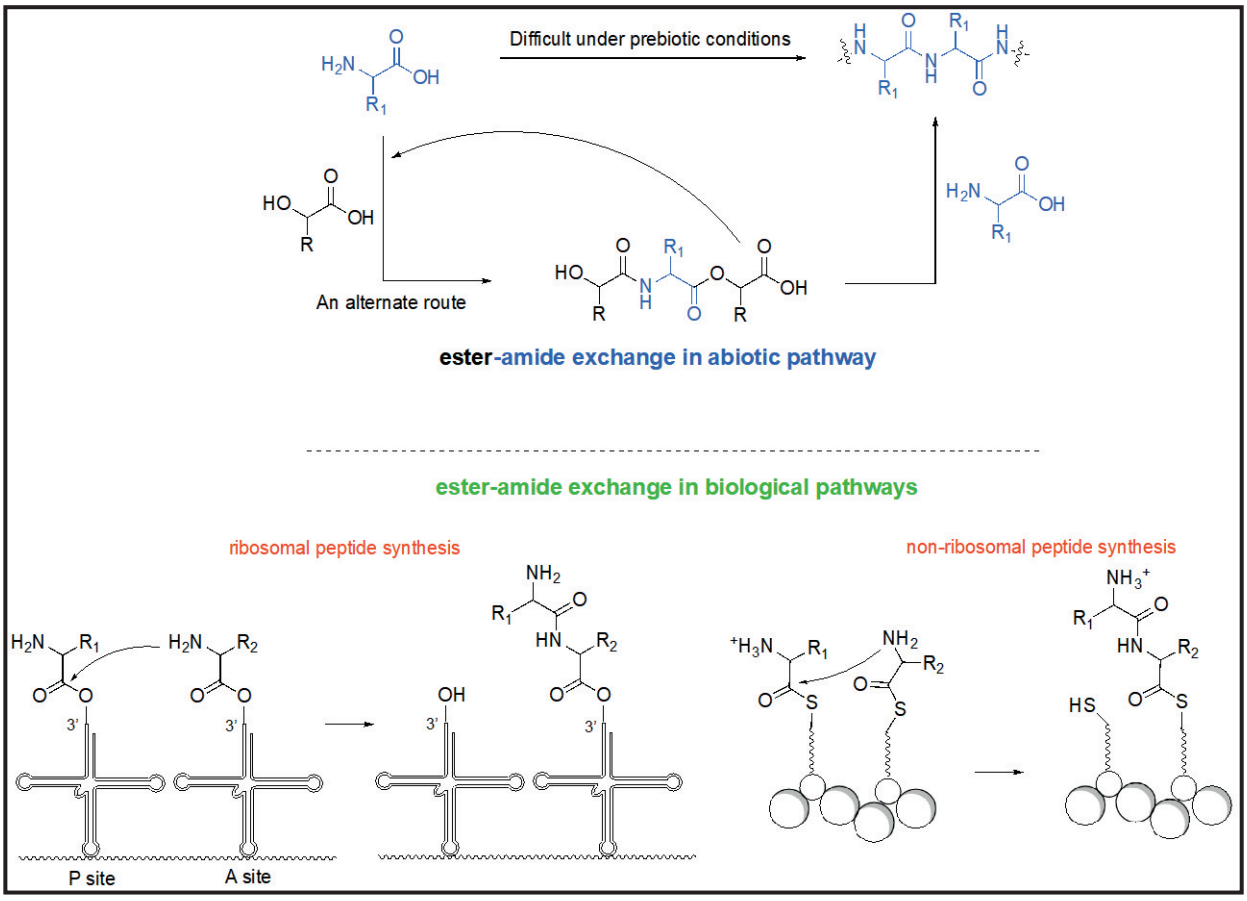

Figure 6. A 'coincidental' preservation of the primordial exchange mechanism? The mechanism of ester-peptide exchange operating under plausible prebiotic condition seems to have been, coincidentally, "preserved" in modern biological pathways, both in the ribosomal and non-ribosomal peptide synthesis.

on behalf of the CCE collaboration, at *The Stanley Miller Memorial Lecture* in May 2015, I was approached by a chemist after the lecture who said (and I am paraphrasing), "*but this same principle cannot be shown for the nucleic acids and the RNA world*". At that time, having no experimental results, I had no answer. It was a time when I was getting disillusioned with pursuing a pre-RNA-world hypothesis, according to which prebiotically plausible pre-RNA systems were the forerunners giving rise to the more difficult-to-assemble RNA. In a proposal, submitted to NASA in July 2013, I expressed my misgivings (which are similar to what had been expressed by *Eschenmoser* for TNA [76]): "*The introduction of an 'ancestor of RNA' brings additional complications. For example, there are questions as to how a pre-RNA world transitioned to an RNA world: (a) Did the 'pre-RNA world' have its own proto-prebiotic chemistry with its own genotype and phenotype? (b) How much of these processes could be (and were) ported over to the 'RNA world'? (c) Or did the 'RNA world' have to develop all of its, genotype, phenotype and its associated processes, 'de novo' independent of the previously developed systems' chemistry? More importantly, a clean transition from a homogeneous-backbone pre-RNA world to a homogeneous-backbone RNA world (akin to an RNA world to a DNA-RNA-protein world transition) would have been highly unlikely unless there were sophisticated mechanisms to separate their prebiotic chemistries. More often than not, it may have been more of a mixture of (heterogeneous-backbone) systems, their chemistries and the interaction between them that would have dictated the path of chemical evolution. Further selection pressures, such as stability (towards hydrolysis and decomposition) and the ability to form complex structures (conferring the ability to replicate, act as a catalyst and selectively bind to small molecules) would have fine-tuned the mixture towards homogeneity*". It was apparent that the Systems Chemistry strategy was taking hold in our thinking naturally, as a consequence of the problems that were created by the reductionist and linear, RNA-first or pre-RNA-first, approaches. And that takes us to the next part of the Systems Chemistry approach in our laboratory – that suggested how RNA and DNA can appear together, challenging the very notion of an RNA-first or an RNA-only world.

By the end of 2015, we had a project that was underway in our laboratory to understand the base-pairing behavior of chimeric nucleic acids in the context of chemical evolution. In the 2013 NASA proposal, we had the following hypothesis: "*This proposal aims to tackle the question of a 'pre-RNA world' from a gradual chemical evolution view-point by postulating that there need not have been a clean homogeneous system to another clean homogeneous system transition, but rather a combinatorial mixing-in and "cleaning-out" of individual chemical-structural elements resulting in a progressive evolution that led to the emergence of RNA. In other words, RNA (and RNA world) would have been a product of incremental constitutional replacement of the previous system rather than a replacement of the system as a whole* (*Fig. 7*). *This, interestingly, a) implies that the previous system(s) need not be homogeneous; rather, this thought process allows for the ancestors of RNA to be a heterogeneous mixture of backbones, recognition elements and linkers, or an entirely different paradigm of transition; b) avoids the reinvention of chemistries associated with each transition of one homogeneous system to the other; and c) implies that self-sorting selection via interaction between the various partners (e.g., base-pairing/template mediated replication) could allow for chemical evolution of a homogeneous backbone. The underlying assumption for this hypothesis is that the same type of (abiotic) chemistry that gave rise to RNA building blocks from ribose, would have also operated on potential alternatives (available by the same chemistries), and produce alternative building blocks (backbone and recognition elements) that would have been available for further processing by combinatorial chemical evolution*." In this proposal, however, we were not considering that two systems could appear simultaneously, only that the emergence of RNA is possible from a chimeric backbone system [91]. As

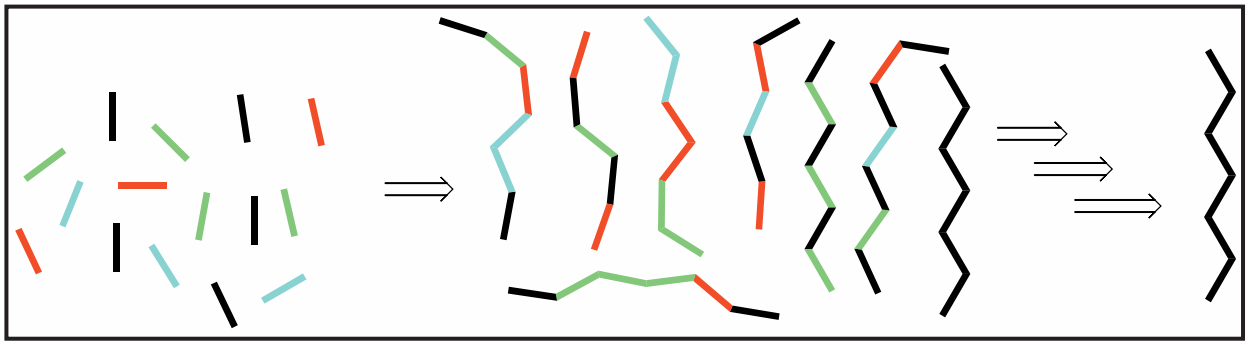

Figure 7. The original Figure in the 2013 NASA proposal. A homogeneous backbone system can emerge from a mixture of heterogeneous-backbone oligomeric systems.

it turned out, we had a lot more to learn from applying the concepts of Systems Chemistry to the emergence of RNA *via* '*heterogeneous mixture of backbones*" or chimeric nucleic acids.

Chimeric Nucleic Acids – "The 18th Camel"

The incentive for exploring chimeric nucleic acids arose from our studies of pentulose nucleic acids [108, 109] (*Fig. 8*). Inspired by the search for alternative nucleic acids starting from the structural neighborhood of RNA, we reasoned that another isomer of ribose to consider would be the corresponding pentuloses, ribulose, and xylulose, which are formed in much greater amounts than ribose in the formose reaction [108]. We synthesized the corresponding xylulose-NA and ribulose-NA, and observed that they were devoid of base-pairing capacity, indicating that they would not be able to compete with RNA in terms of base-pair- mediated functions [108]. Insertion of one or two units of pentulose-NA in RNA also drastically weakened the base pairing of RNA [108]; however, much to our surprise, a strictly alternating (chimeric) xylulose-ribose-NA or ribulose-ribose-NA with 50% incorporation showed (sequence-dependent) strong base pairing, sometimes exceeding even the duplexes from comparable RNA sequences (*Fig. 8*) [109]. The fact that nucleic-acid systems that have no base-pairing properties and have no cross-pairing with RNA, can become 'functional' when mixed with RNA (in an alternating sequence) led us to consider initiating a study with chimeric sequences of RNA and DNA – two systems that are known to strongly undergo base pairing and cross-pairing with each other [110]. Our idea was to test the RNA-world to RNA-DNA-world transition and to investigate the base-pairing properties of the chimeric sequences of mixed RNA-DNA ('RDNA') that is expected to result in such prebiotic transitions – with the (naïve) expectation that there would be a smooth base-pairing landscape of transition from a pure RNA sequence to a pure DNA sequence *via* the chimeric sequences, because RNA and DNA have strong affinity for each other. However, unexpectedly, all of the duplexes from RDNA chimeric sequences (without exception) had much lower thermal stability when compared to the corresponding RNA or DNA duplexes (*Fig. 9, a*) [110]. This meant that the energy landscape in transitioning from RNA sequences (in an RNA world) to an RNA-DNA world would have to reckon with an 'energy barrier' in the form of RDNA sequences that are inferior in terms of base-pairing strength (*Fig. 9, 10*) [110]. While this poses a problem for starting from a ribonucleos(t)ides/RNA-first-only world, it opened up a Systems Chemistry approach when we considered the possibility of starting from a mixture of ribose and deoxyribose nucleos(t)ides. Since, as implied in *Fig. 10,* one can envision that this mixture can lead to RDNA sequences with varying degrees of RNA and DNA incorporation. And based on the thermodynamic preference for duplex formation, the system may eventually self-select only those that are capable of base pairing.

This preference is expected to lead to systems that have more homogeneous-sugar-backbone sequences. Thus, began our journey into the Systems Chemistry of chimeric and homogeneous nucleic acids, where we operated on the (erroneous) assumption that chimeric sequences would be 'dead ends' because of their inability to form duplexes with their complementary counterparts. It turned out to be quite the opposite and taught us a good lesson of how we can convince ourselves to be right with our ideas (when we are wrong)!

The first chimeric system we investigated in this context comprised of TNA–RNA residues (TRNA), based on the assumption that the similar chemical pathways for the prebiotic synthesis of TNA and RNA would allow for their co-formation, co-existence, co-polymerization, and co-evolution [111]. Based on previous observations of TNA base-pairing properties [71], we synthesized two self-complementary TRNA chimeric sequences: one with TNA units alternating with RNA units and the other a block sequence with stretches of TNA residues and RNA residues (*Fig. 11, a*). The duplexes from these sequences exhibited weaker affinities with peculiar pairing behavior as observed by their temperature-dependent (UV-T_m) melting curves. A self-complementary chimera with an alternating arrangement [TNA(T)-RNA(A)]$_n$ formed a stronger duplex compared to the reverse [TNA(A)-RNA(T)]$_n$, while the exact opposite behavior [TNA(A)$_n$-RNA(T)$_n$ > TNA(T)$_n$-RNA(A)$_n$] was observed in a chimeric block arrangement. Guided by this observation, we designed two TRNA chimeric non-self-complementary strands where the block portion incorporated RNA(T) and TNA(A), while the alternate portion contained TNA(T) and RNA(A), anticipating a strong duplex. Unexpectedly, not only did these two complementary chimeric heterogeneous sequences have weak affinity for each other, but even more surprisingly, these chimeric TNA-RNA sequences formed stronger duplexes with the corresponding complementary homogeneous-backbone RNA (or TNA) sequences (*Fig. 11, a*). Further studies confirmed this

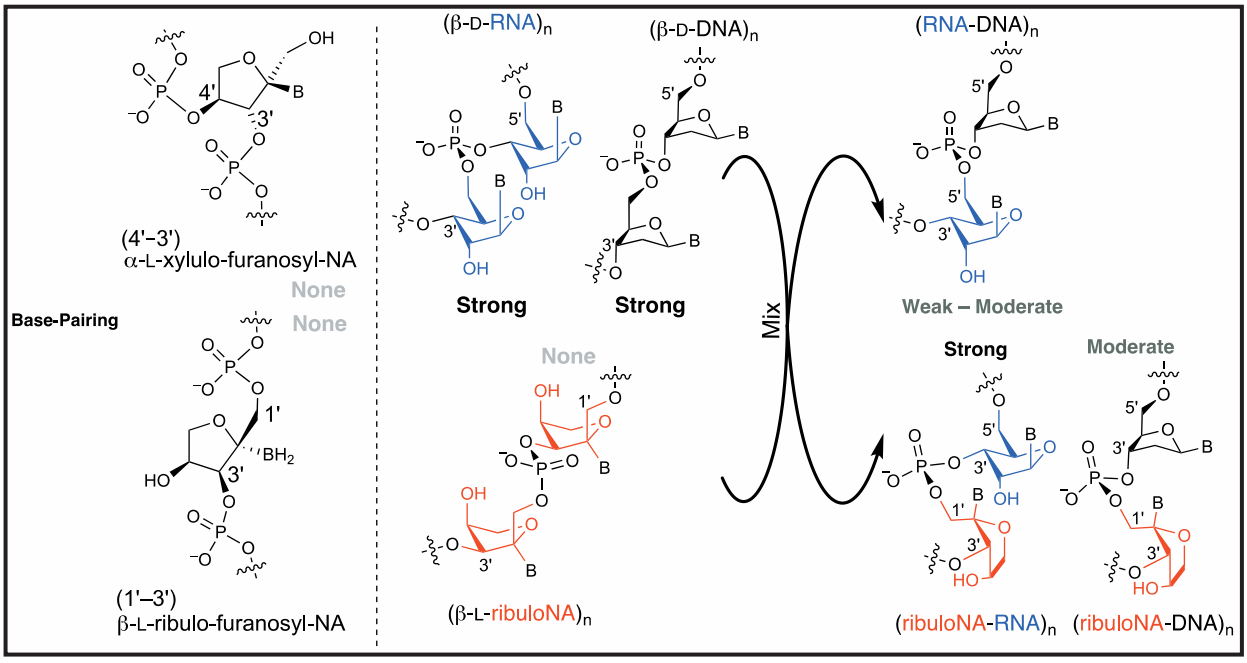

Figure 8. The chimeric pentulose-pentose nucleic acids. The pentulose nucleic acid systems, which have no base-pairing properties, surprisingly give rise to moderate-to-strong (or in some cases stronger than RNA or DNA) duplexes when interspersed in RNA or DNA sequences in a strictly alternating manner.

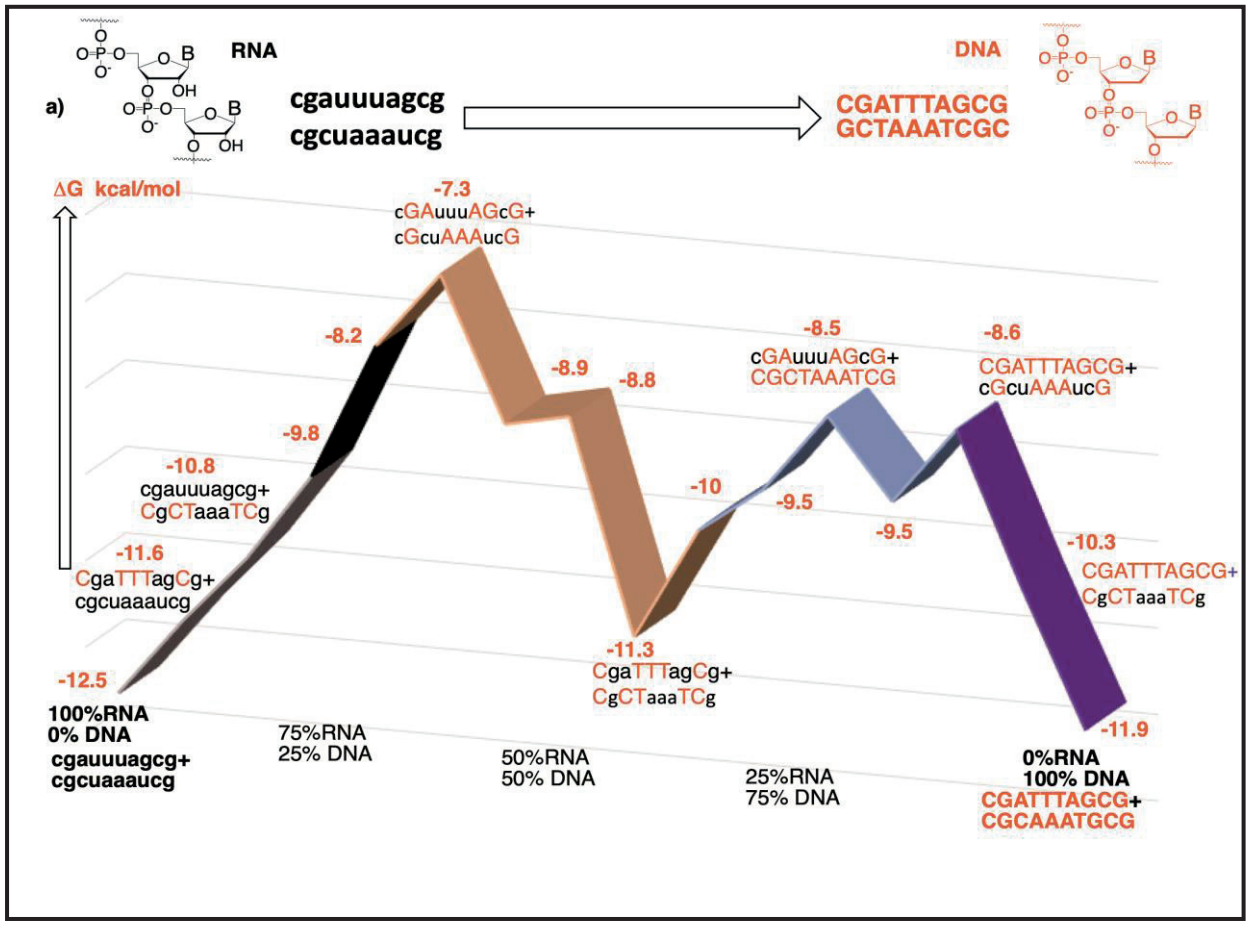

Figure 9. Thermodynamic Energies of Chimeric Sequences in the Progression of RNA to DNA. The energy landscape of transition from an RNA-duplex to a DNA-duplex *via* chimeric-heterogeneous-RDNA duplexes encounters a "thermodynamic energy barrier".

behavior to be true and general for TRNA chimeric sequences containing all four nucleobases (A, T/U, G and C). *This was the observation* that led us to consider the possibility that these chimeric and heterogeneous TRNA sequences could preferentially act as templates for non-enzymatic ligations of the complementary homogeneous RNA and TNA strands, in a mixture of sequences [111]. When we tested this idea of non-enzymatic ligation with chimeric TRNA templates and a mixture of complementary TRNA ligands and RNA ligands (*Fig. 11, b*), we did observe the preferential formation of the corresponding RNA-ligation product accompanied by very little of the chimeric TRNA product [111]. Thus, was born the answer to the question that was posed by the chemist to me in 2015 – yes, the same principle of a Systems Chemistry approach that was used for the emergence of peptides from depsipeptides is also applicable to understanding of the appearance of homogeneous sugar-backbone-containing nucleic acids starting from a mixture of chimeric and unmixed nucleic acid sequences (*Fig. 11, c*).

Since the synthetic availability of TNA was resource- and time-limiting, we quickly switched to investigating the chimeric RDNA system based on our previous studies (*Fig. 9*) and the ease of their commercial availability. This also allowed us an opportunity to test whether the paradigm of chimeric templates preferentially binding the corresponding complementary homogeneous sequences is a general phenomenon and not limited to the TRNA, TNA, RNA combination. Indeed, it turned out that the chimeric RDNA

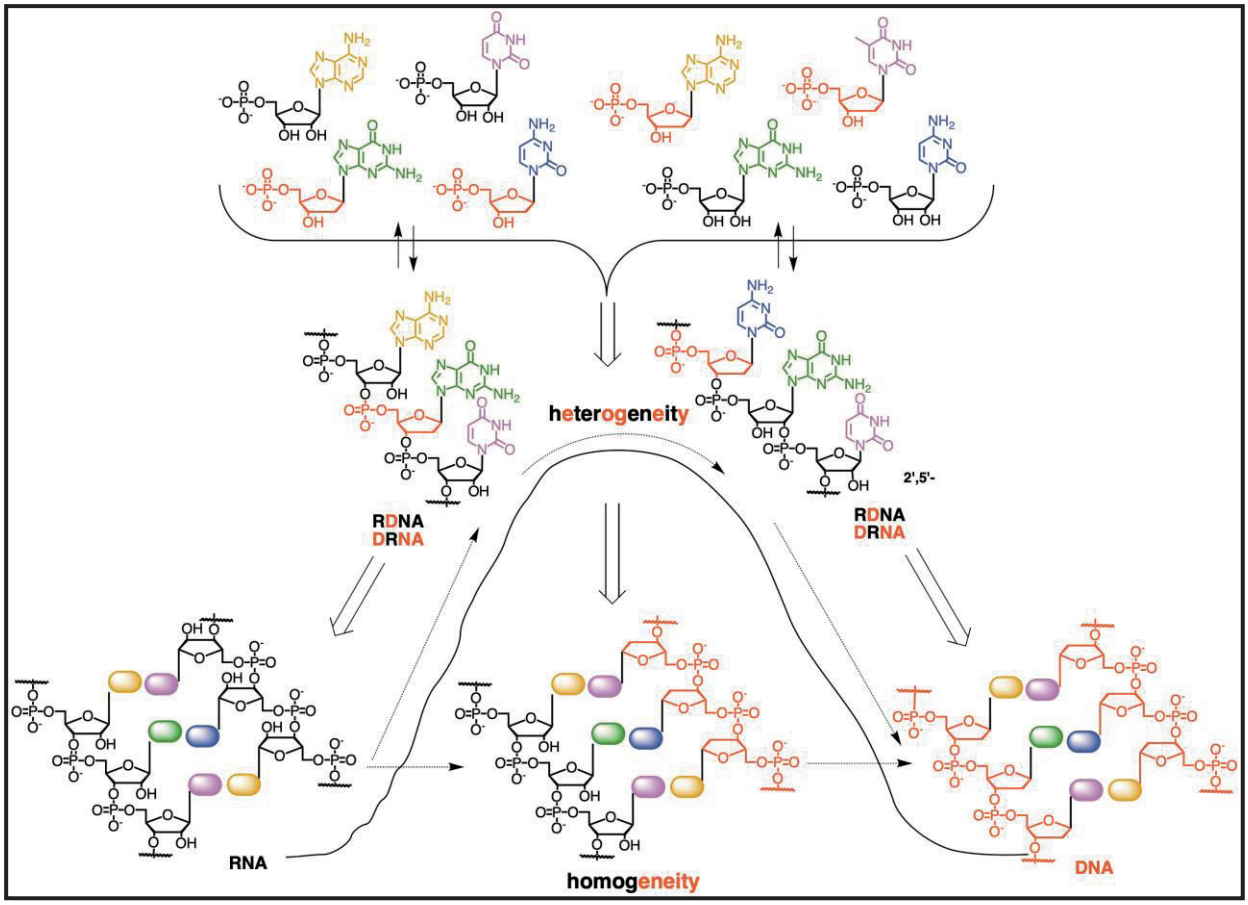

Figure 10. Chimeric RNA-DNA in the transition of RNA world to an RNA-DNA world. If one were to start from a mixture of RNA and DNA nucleotides and produce the chimeric- and unmixed-backbone strands, the thermodynamic sink of homogeneous-backbone duplexes may drive the emergence of simultaneous RNA and DNA sequences. This paradigm argues against the generally accepted RNA-world hypothesis of RNA-first DNA-second scenario.

templates did preferentially ligate the homogeneous RNA and DNA ligands (over the RDNA chimeric ligands) in the presence of a mixture of these ligands. This observation set the stage for the next logical stage of the System Chemistry scenario as to whether such a thermodynamic driven preference of duplexes, which are in equilibrium, can be pushed towards one side by the presence of another sequence that is complementary, and can bind, to the newly formed ligation product of RNA (or DNA) and releasing the chimeric template to engage in another round of ligation to form the homogeneous RNA (or DNA) product (*Fig. 12, a*). That such a template-product inhibition, characteristic of the RNA world, could be overcome by the use of chimeric RDNA templates was demonstrated by comparing the efficiency of RDNA templates *vs.* the corresponding RNA templates [111]. In every case, the RDNA chimeric template outperformed the RNA template in terms of efficiency of forming the final ligation (RNA) product. This observation was also true when the system was subject to a selection pressure by diluting the system, such that the chimeric RDNA or the homogeneous RNA templates had to compete for the ligands. Such observations again reinforced the benefits of a System Chemistry approach of using the mixtures of chimeras along with homogeneous-backbone sequences in overcoming the limitations of the classic template-product inhibition encountered when using only a (prebiotically implausible) pure

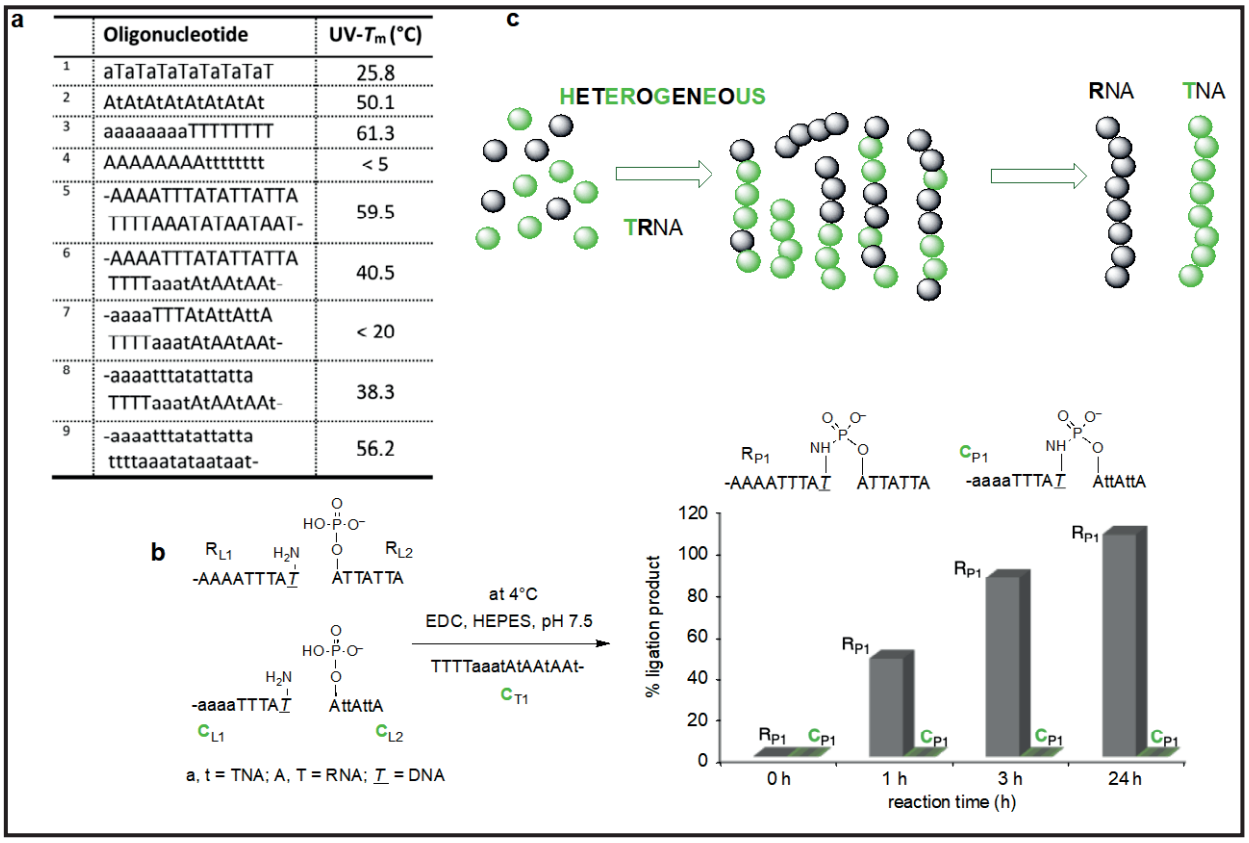

Figure 11. The beneficial role of chimeric TRNA sequences. *a*) The peculiar base-pairing behavior of chimeric TRNA sequences led to the possibility that *b*) the chimeric TRNA sequences can act as templates for ligating homogeneous-RNA (or TNA) ligands preferentially over chimeric TRNA ligands, suggesting that *c*) homogeneous-backbone-RNA and -TNA strands could emerge starting from a mixture of sequences. For conditions of measurements, see [111].

RNA system. There were further benefits of implementing the System Chemistry strategy in the above mixture of nucleic acid sequences. For example, if the system is able to reproduce the chimeric RDNA template starting from chimeric ligands, then there is a possibility of inducing cross-catalytic synthesis of both the homogeneous and chimeric sequences leading to possibly amplified production of the homogeneous RNA product (*Fig. 12*, *b*). We were able to achieve a proof-of-principle demonstration of this, wherein a brew of the RNA ligands along with the chimeric RDNA ligands in the presence of the RDNA template produced much more (160%) of the final RNA product when compared to the mixture that lacked the chimeric RDNA ligands (108%) [111] (*Fig. 12*, *b*). This result, combined with the others described above, clearly shows the advantages of the Systems Chemistry *modus operandi*, and what outcomes can be obtained – the very same ones that have been very difficult to attain when pursuing experiments based on an RNA-only scenario. The "18th camel", in this example, would be the chimeric RDNA sequences, which (when added to the mixture) enable the emergence of the homogeneous RNA and DNA backbones, but they themselves (as chimeric sequences) are not present in the final biologically functionally relevant systems.

Systems Chemistry and the RNA World (an Overlooked Story)

It was only when we were writing up this work for publication that I became aware of the hypotheses of *Woese* [13] (thanks to *Antonio Lazcano* who pointed this reference to me) and *Orgel* [112],

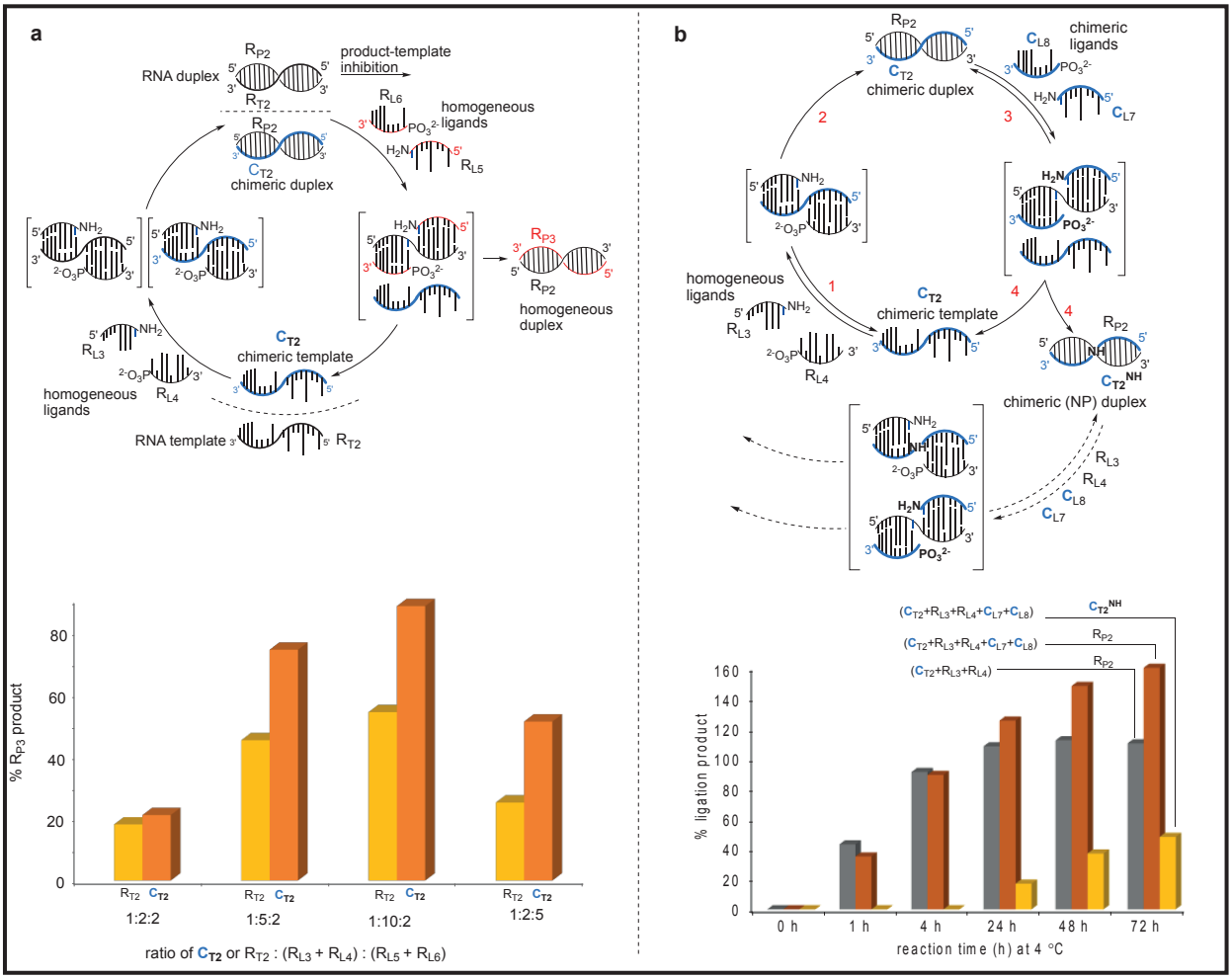

Figure 12. Chimeric RDNA templates are better than RNA (or DNA) sequences in replication. *a)* The chimeric RDNA templates are more efficient in overcoming template-product inhibition (when compared to RNA sequences) and thus enable the emergence of homogeneous-backbone-RNA (and -DNA) sequences starting from a combination of RDNA, RNA, and DNA ligands. *b*) By the same principle, the chimeric RDNA templates also enable their self-replication *via* cross-replication of the corresponding RNA product, thus creating the potential for an auto-amplification scenario. For conditions of measurements, see [111].

where they had considered the possibility that a mixture of nucleic acids would have existed in prebiotic chemistry. For example, *Woese* in 1967 wrote, "*Similarly, any polynucleotides made by unguided polymerization could contain many kinds of bases as well as a variety of sugars and a variety of linkages among the different components*" [13]. *Orgel* in 1974 was even more explicit [112], writing "*We believe that nucleosides, deoxynucleosides, and derivatives of related sugars must have coexisted in the prebiotic soup, and copolymerized to form the first prebiotic nucleic acid like molecules; we do not think, therefore, that the question – which came first, DNA or RNA? – is a good one. We have used ribonucleoside-containing models extensively because ribonucleoside derivatives are readily available and undergo efficient template-directed condensations. Further work on derivatives of other sugars would be desirable*" (unfortunately, *Orgel* seems to have had a change of heart on this sentiment, telling *Gerald Joyce* that he did not believe in it anymore – as conveyed by *Gerald Joyce* to me *via* personal communication in 2019). Despite this, it is marvelous to see how much foresight these pioneers had, one that is coming

true in the age of Systems Chemistry. Another pioneer, *John Oró*, in 1974, expressed a similar idea stating that "*It appears probable that DNA and RNA oligomers could have arisen simultaneously on the primitive Earth*", based on the consideration of the physicochemical properties of these nucleic acids, though he did not provide much details [113]. Later, *Sutherland* [114], *Miller* and *Lazcano* [115], *Follman* [116], *Szostak* and *Powner* [117], and *Wächtershäuser* [118] considered (in various forms) that RNA and DNA could have co-existed at various stages of chemical evolution. *Szostak* and *Powner* went so far as to state "*Thus, chimeric RNA/DNA polymers may have been sufficient for the emergence of life. The primordial biochemical exploitation of a mixed RNA and DNA genetic system could eliminate the requirement for a genetic takeover (of RNA by DNA), and would arguably result in a simplification of the transition from chemistry to biology*" [117]. Thus, it is evident that the concept of Systems Chemistry was present in various forms (even in the RNA world), though it may not have been explicitly recognized as such. It is becoming clear that the "RNA world" has to take into account the Systems Chemistry approach and the (slow but sure) change of heart in many of the RNA-world practitioners [119] is beautifully illustrated by the title of a recent perspective "*The difficult case of an RNA-only origin of life*" [120]!

Is Systems Chemistry the Panacea?

While the two systems described above –*a*) the ester-amide exchange chemistry in the context of enriching a depsipeptide backbone with increasing amounts of peptide bonds and *b*) the chimeric RDNA systems leading to homogeneous RNA and DNA sequences – are promising examples, they are only a start. The paradigms of Systems Chemistry needs to be applicable also at the more fundamental level of prebiotic chemistry that gives rise not only to the building blocks of amino acids and nucleic acids, but also lipids [121] and sugars [122]. The examples highlighted in this article from our work are what I would consider to be relatively clean starting points which are *a*) amenable to further manipulations and *b*) give comparatively clean outcomes that can be analyzed in a manner that give clear-cut answers. In that sense, these are only proof-of-principle experiments which demonstrate the need and importance of Systems Chemistry in the larger picture, but in itself are not the final word. For the wider acceptance of Systems Chemistry approaches in Origins of Life studies, we need to continue to drill down to understand how these building blocks would have formed in the first place and continue to interact – and we may come up on the realization that there may be a limit to how and where the Systems Chemistry approaches can be applied [123]. For example, the current approaches in our laboratory are largely driven by applying the principles of Systems Chemistry to explore new venues in investigating the chemistries that could lead to generation of complexities, both in terms of the chemical entities and their interactions.

Therein, we come face to face with the harsh realities of dealing with prebiotic mixtures that have been called "*prebiotic clutter*", "*tar*", and "*messy chemistry*" [124, 125]. The term, *Systems Chemistry*, when applied to prebiotic chemistry conjures up all of these descriptions and with good reasons. Starting from the formose reaction to the *Urey-Miller* spark discharge experiment to HCN-polymerization chemistry, we have seen how difficult it has been to move beyond the complex mixtures. There have been serious attempts [107, 126] by simplifying the systems and reducing the parameters involved; however, it has still been challenging to identify chemistries that would emerge from these simple building blocks and give rise to chemical entities that begin to interact, leading to the emergence of self-sustaining Systems Chemistry. If the System Chemistry approach is to be taken seriously in the context of the chemical Origins of Life [123], experimental verifications must be provided at all levels of complexities and not just at the level that deals with clean starting materials. While that is a real challenge [34] (especially given the historical difficulties), I think it is also the strength of Systems Chemistry – leading us to discover the *18th camel* by prompting us to look beyond the molecules that are used by extant biology and

including the prebiotically relevant molecules that would have co-existed along with biogenic ones. And that may lead us to recognize and realize the unexpected benefits of bringing such co-existing molecules together – that can interact in a covalent and non-covalent fashion – giving rise to emergent properties and behavior that may provide unanticipated solutions to problems that were once '*unsolvable*'. It may also allow us to identify '*sweet spots*' of complex mixtures which are prebiotically relevant but do not cause runaway messiness, enabling us to spot the "*habitable zone*" equivalents of prebiotic clutter [82]. In that process, it may reinforce the most important lesson of the difference between "*can*" and "*could*" in our pursuit of an understanding of the chemical Origins of Life. As stated by *Eschenmoser* and *Kisakürek*, "*The natural genesis of life on Earth is a hypothesis of evolutionary science; it is the task of synthetic organic chemistry to test this hypothesis experimentally. The aim of an experimental aetiological chemistry is not primarily to delineate the pathways along which our ('natural') life on Earth could have originated, but to provide decisive experimental evidence, through the realization of model systems ('artificial chemical life'), that life can arise as a result of the organization of organic matter*" [127]. In this context, understanding the import of '*can*' in that sentence is also the equivalent of the *18th camel*, and, that Systems Chemistry is not the panacea but a possibility!

Acknowledgements

Funding jointly from the NSF and the NASA Exobiology program under the NSF-Center for Chemical Evolution (CH-1504217) and an award from the Simons Foundation (327124) to RK is gratefully acknowledged. I am grateful to Professor Albert Eschenmoser for having introduced me to, and mentored me in, this fascinating field. I am indebted to Nick Hud for the invitation (and camaraderie) to be part of the Center for Chemical Evolution starting in 2010 that led to the ester-amide project, and the subsequent collaborative efforts that has resulted in many of the depsipeptide projects currently underway in the Center. And I thank C.B.S. Mani for bringing the '17 camel problem' to my attention, Luke Leman and my group colleagues for reading the manuscript and providing feedback. Finally, I would like to thank Professor *Stuart Schreiber* who first encouraged me and insisted that I publish the "*18th Camel Paradigm*" in the context of the Chemical Evolution, during his visit to *Scripps* on April 4, 2016.

Author details

[1]Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

Competing interests

The author declares that he has no competing interests.

References

1 Kitadai, N., Maruyama, S.: Origins of Building Blocks of Life: A Review. Geosci Front 9, (2018) 1117-1153.

2 Pross, A.: What is Life?: How Chemistry Becomes Biology. Oxford University Press, Oxford (2012).

3 De Duve, C.: A Research Proposal on the Origin of Life. Orig Life Evol Biosph 33, (2003) 559-574.

4 Orgel, L. E.: The Origin of Life - A Review of Facts and Speculations. Trends in Biochem Sci 23, (1998) 491-495.

5 Forterre, P., Gribaldo, S.: The Origin of Modern Terrestrial Life. Human Front Sci Prog 1, (2007) 156-168.

6 Miller, S. L.: The Prebiotic Synthesis of Organic Molecules and Polymers. Adv Chem Phys 55, (1984) 85-107.

7 Fox, S. W.: A Theory of Macromolecular and Cellular Origins. Nature 205, (1965) 328-340.

8 Fox, S. W.: Self-ordered Polymers and Propagative Cell-like Systems. Naturwissenschaften 56, (1969) 1-9

9 Fox, S. W.: Molecular Evolution to the First Cells. Pure Appl Chem 34, (1973) 641-670.

10 Andras, P., Andras, C.: The Origins of Life – The 'Protein Interaction World' Hypothesis: Protein Interactions were the First Form of Self-Reproducing Life and Nucleic Acids Evolved Later as Memory Molecules. Med Hypotheses 64, (2005) 678-688.

11 Rich, A.: On the Problems of Evolution and Biochemical Information Transfer. 103-126 Academic Press, (1962).

12 Crick, F. H.: The Origin of the Genetic Code. J Mol Biol 38, 367-379 (1968).

13 Woese, C. R.: The Genetic Code: The Molecular Basis for Genetic Expression. Harper and Row, New York (1967).

14 Orgel, L. E.: Evolution of the Genetic Apparatus. J Mol Biol 38, (1968) 381-393.

15 Gilbert, W.: Origin of Life - The RNA World. Nature 319, (1986) 618-618.

16 Schwartz, A. W.: The RNA World and its Origins. Planet Space Sci 43, (1995) 161-165.

17 Yarus, M.: Boundaries for an RNA World. Curr Opin Chem Biol 3, (1999) 260-267.

18 Orgel, L. E.: Some Consequences of the RNA World Hypothesis. Orig Life Evol Biosph 33, (2003) 211-218.

19 Joyce, G. F., Schwartz, A. W., Miller, S. L., Orgel, L. E.: The Case for an Ancestral Genetic System Involving Simple Analogs of the Nucleotides. Proc Natl Acad Sci USA 84, (1987) 4398-4402.

20 Gesteland, R. F., Cech, T. R., Atkins, J. F.: The RNA world: The Nature of Modern RNA Suggests a Prebiotic RNA. Cold Spring Harbor Laboratory Press, New York (1999)

21 Joyce, G. F., Orgel, L. E.: Progress Towards Understanding the Origin of the RNA world. Cold Spring Harbor Monograph Series 43, (2006) 23-56

22 Cairns-Smith, A. G., Walker, G. L.: Primitive Metabolism.

BioSystems 5, (1974) 173-186.

23 Hartman, H.: Speculations on the Origin and Evolution of Metabolism. J Mol Evol 4, (1975) 359-370.

24 Wächtershäuser, G.: Before Enzymes and Templates: Theory of Surface Metabolism. Microbiol Rev 52, (1988) 452-484.

25 Morowitz, H. J., Kostelnik, J. D., Yang, J., Cody, G. D.: The Origin of Intermediary Metabolism. Proc Natl Acad Sci U S A 97, (2000) 7704-7708.

26 Morowitz, H. J.: A Theory of Biochemical Organization, Metabolic Pathways, and Evolution. Complexity 4, (1999) 39-53.

27 Kauffman, S. A.: Investigations. Oxford University Press, Oxford (2000).

28 Goldford, J. E., Segrè, D.: Modern Views of Ancient Metabolic Networks. Cur Opin Syst Biol 8, (2018) 117-124.

29 Shapiro, R.: A Simpler Origin for Life. Sci Amer 296, (2007) 46-53.

30 Kauffman, S. A.: Approaches to the Origin of Life on Earth. Life 1, (2011) 34-48.

31 Copley, S. D., Smith, E., Morowitz, H. J.: The Origin of the RNA World: Co-evolution of Genes and Metabolism. Bioorg Chem 35, (2007) 430-443.

32 Segré, D., Ben-Eli, D., Deamer, D. W., Lancet, D.: The Lipid World. Orig Life Evol Biosph 31, (2001) 119-145.

33 Walde, P., Wick, R., Fresta, M., Mangone, A., Luisi, P. L.: Autopoietic Self-Reproduction of Fatty Acid Vesicles. J Am Chem Soc 116, (1994) 11649-11654.

34 Krishnamurthy, R.: Giving Rise to Life: Transition from Prebiotic Chemistry to Protobiology. Acc Chem Res 50, (2017) 455-459.

35 Lazcano, A.: The RNA World and the Origin of Life: A Short History of a Tidy Evolutionary Narrative. BIO Web of Conf 4,(2015) 00013, 1-8.

36 Bernhardt, H. S.: The RNA World Hypothesis: The Worst Theory of the Early Evolution of Life (Except for all the others). Biol Direct 7, (2012) 23.

37 Lee, D. H., Granja, J. R., Martinez, J. A., Severin, K., Ghadri M, R.: A Self-Replicating Peptide. Nature 382, (1996) 525 – 528.

38 Cody, G. D., Boctor, N. Z., Filley, T. R., Hazen, R. M., Scott, J. H., Sharma, A., Yoder Jr., H. S.: Primordial Carbonylated Iron-Sulfur Compounds and the Synthesis of Pyruvate. Science 289, (2000) 1337-1340.

39 Cody, G. D., Boctor, N. Z., Hazen, R. M., Morowitz, H. J., Yoder, Jr. H. S.: Geochemical Roots of Autotrophic Carbon Fixation: Hydrothermal Experiments in the System Citric Acid, H_2O-(±FeS)–(±NiS). Geochim Cosmochim Acta 65, (2001) 3557-3576.

40 Martin, W., Russell, M. J.: On the Origin of Biochemistry at an Alkaline Hydrothermal Vent. Philos Trans R Soc B Biol Sci 362, (2007) 1887-1926.

41 Novikov, Y., Copley, S. D.: Reactivity Landscape of Pyruvate under Simulated Hydrothermal Vent Conditions. Proc Natl Acad Sci USA 110, (2013) 13283-13288.

42 Sojo, V., Herschy, B., Whicher, A., Camprubí, E., Lane, N.: The Origin of Life in Alkaline Hydrothermal Vents. Astrobiol 16, (2016) 181-197.

43 Cech, T. R.: The Ribosome is a Ribozyme. Science 289, (2000) 878-879.

44 Breaker, R. R., Joyce, G. F.: The Expanding View of RNA and DNA Function. Chem Biol 21, (2014) 1059-1065.

45 Sczepanski, J. T., Joyce, G. F.: A Cross-chiral RNA Polymerase Ribozyme. Nature 515, (2014) 440-442.

46 Lincoln, T. A., Joyce, G. F.: Self-sustained Replication of an RNA Enzyme. Science 323, (2009) 1229-1232.

47 Wochner, A., Attwater, J., Coulson, A., Holliger, P.: Ribozyme-catalyzed Transcription of an Active Ribozyme. Science 332, (2011) 209-212.

48 Stankiewicz, J., Eckardt, L. H.: Chembiogenesis 2005 and Systems Chemistry Workshop. Angew Chem Int Ed 45, (2006) 342-344.

49 Kindermann, M., Stahl, I., Reimold, M., Pankau, W., von Kiedrowski, G.: Systems Chemistry: Kinetic and Computational Analysis of a Nearly Exponential Organic Replicator. Angew Chem Int Ed 44, (2005) 6750-6755.

50 Hayden, E., von Kiedrowski, G., Lehman, N.: Systems Chemistry on Ribozyme Self-construction: Evidence for Anabolic Autocatalysis in a Recombination Network. Angew Chem Int Ed 47, (2008) 8424-8428.

51 Ludlow, R., Otto, S.: Systems Chemistry. Chem Soc Rev 37, (2008) 101-108.

52 Ruiz-Mirazo, K., Briones, C., de la Escosura, A.: Prebiotic Systems Chemistry: New Perspectives for the Origins of Life. Chem Rev 114, (2014) 285-366.

53 Peyralans, J., Otto, S.: Recent Highlights in Systems Chemistry. Curr Opin Chem Biol 13, (2009) 705-713.

54 Ashkenasy, G., Hermans, T. M., Otto, S., Taylor, A. F.: Systems Chemistry. Chem Soc Rev 46, (2017) 2543-2554.

55 Kee, T. P., Monnard, P.-A.: Chemical Systems, Chemical Contiguity and the Emergence of Life. Beil J Org Chem 13, (2017) 1551-1563.

56 Bai, Y. Chotera, A., Taran, O., Liang, C, Ashkenasy, G., Lynn, D.: Achieving Biopolymer Synergy in Systems Chemistry. Chem Soc Rev 47, (2018) 5444-5456.

57 Szostak, J. W. An Optimal Degree of Physical and Chemical Heterogeneity for the Origin of Life? Philos Trans R Soc B Biol Sci 366, (2011) 2894-2901.

58 Szostak, J.: Systems Chemistry on Early Earth. Nature 459, (2009) 171-172.

59 Islam, S., Powner, M. W.: Prebiotic Systems Chemistry: Complexity Overcoming Clutter. Chem 2, (2017) 470-501.

60 Lancet, D., Zidovetzki, R., Markovitch, O.: Systems Protobiology: Origin of Life in Lipid Catalytic Networks. J R Soc Interface 15, (2018) 20180159, 1-36.

61 Ashkenasy, G., Jagasia, R., Yadav, M. Ghadiri, R. M.: Design of a Directed Molecular Network. Proc Natl Acad Sci USA 101, (2004) 10872 – 10877.

62 Kauffman, S.: Systems Chemistry Sketches. Chem Evol Space Time 981, (2008) 310 – 324.

63 Tor, Y.: Catalyst: On Organic Systems Chemistry and Origins Research. Chem 2, (2017) 448-450.

64 Eschenmoser, A.: Chemical Etiology of Nucleic Acid Structure. Science 284, (1999) 2118-2124.

65 Eschenmoser, A.: Why Pentose-nucleic Acids and not Hexose-nucleic Acids? Nachrichten aus Chemie, Technik und Laboratorium 39, (1991) 795-798.

66 Eschenmoser, A. Loewenthal, E.: Chemistry of Potentially Prebiological Natural Products. Chem Soc Rev 21, (1992) 1-16.

67 Eschenmoser, A. Dobler, M. Warum Pentose- und nicht Hexose-Nucleinsauren?? Teil I. Einleitung und

Problemstellung, Konformationsanalyse fur Oligonucleotid-Ketten aus 2',3'-Dideoxyglucopyranosyl-Bausteinen ('Homo-DNS') sowie Betrachtungen zur Konformation von A- und B-DNS. Helv Chim Acta 75, (1992) 218-259.

68 Beier, M., Reck, F., Wagner, T., Krishnamurthy, R. Eschenmoser, A.: Chemical Etiology of Nucleic Acid Structure: Comparing Pentopyranosyl-(2'->4') Oligonucleotides with RNA. Science 283, (1999) 699-703.

69 Pitsch, S., Wendeborn, S., Krishnamurthy, R., Holzner, A., Minton, M., Bolli, M., Miculka, C., Windhab, N., Micura, R., Stanek, M., Jaun, B., Eschenmoser, A.: Pentopyranosyl Oligonucleotide Systems. 9th Communication. Helv Chim Acta 86, (2003) 4270-4363.

70 Schöning, K. U., Scholz, P., Guntha, S., Wu, X., Krishnamurthy, R., Eschenmoser, A.: Chemical Etiology of Nucleic Acid Structure: the alpha-Threofuranosyl-(3'-2') Oligonucleotide System. Science 290, (2000) 1347.

71 Schöning, K. U., Scholz, P., Wu, X., Guntha, S., Delgado, G., Krishnamurthy, R., Eschenmoser, A.: The α-L-Threofuranosyl-(3'□2')-oligonucleotide System ('TNA'): Synthesis and Pairing Properties. Helv Chim Acta 85, (2002) 4111-4153.

72 Eschenmoser, A.: Etiology of Potentially Primordial Biomolecular Structures: From Vitamin B12 to the Nucleic Acids and an Inquiry into the Chemistry of Life's Origin: A Retrospective. Angew Chemie Int Ed 50, (2011) 12412-12472.

73 Krishnamurthy, R.: RNA as an Emergent Entity: An Understanding Gained Through Studying its Nonfunctional Alternatives. Synlett 25, (2014) 1511-1517.

74 Joyce, G. F., Orgel, L. E.: Prospects for Understanding the Origin of the RNA World. Cold Spring Harbor Monograph Series 37, (1999) 49–77.

75 Orgel, L.: Perspectives: Origin of life: A Simpler Nucleic Acid. Science 290, (2000)1306-1307.

76 Eschenmoser, A.: The TNA-Family of Nucleic Acid Systems: Properties and Prospects. Orig Life Evol Biosph 34, (2004) 277-306.

77 Mittapalli, G. K., Reddy, K. R., Xiong, H., Munoz, O., Han, B., De Riccardis, F., Krishnamurthy, R., Eschenmoser, A.: Mapping the Landscape of Potentially Primordial Informational Oligomers: Oligodipeptides and Oligodipeptoids Tagged with Triazines as Recognition Elements. Angew Chem Int Ed 46, (2007) 2470-2477.

78 Mittapalli, G. K., Osorinio, Y. M., Guerrero, M. A., Reddy, K. R., Krishnamurthy, R., Eschenmoser, A.: Mapping the Landscape of Potentially Primordial Informational Oligomers: Oligodipeptides Tagged with 2,4-Disubstituted 5-Aminopyrimidines as Recognition Elements. Angew Chem Int Ed 46, (2007) 2478-2484.

79 Zhang, X. , Krishnamurthy, R.: Mapping the Landscape of Potentially Primordial Informational Oligomers: Oligo-dipeptides Tagged with Orotic Acid Derivatives as Recognition Elements. Angew Chem Int Ed 48, (2009) 8124-8128.

80 Kim, E.-K., Martin, V. , Krishnamurthy, R.: Orotidine-Containing RNA: Implications for the Hierarchical Selection (Systems Chemistry Emergence) of RNA. Chem Euro J 23, (2017) 12668-12675.

81 McMurry, J. , Begley, T.: The Organic Chemistry of Biological Pathways, Roberts and Company Publishers, Englewood (2005).

82 Yadav, M.; Kumar, R.; Krishnamurthy, R.: Chemistry of Abiotic Nucleotide Synthesis. Chem Rev. 120, (2020) 4766-4805.

83 Sutherland, J. D.: The Origin of Life—Out of the Blue. Angew Chem Int Ed 55, (2016) 104-121.

84 Becker, S., Feldmann, J., Wiedemann, S., Okamura, H., Schneider, C., Iwan, K., Crisp, A., Rossa, M., Amatov, T., Carell, T.: Unified Prebiotically Plausible Synthesis of Pyrimidine and Purine RNA Ribonucleotides. Science 366, (2019) 76-82.

85 Karri, P., Punna, V., Kim, K. , Krishnamurthy, R.: Base-Pairing Properties of a Structural Isomer of Glycerol Nucleic Acid. Angew Chem Int Ed 52, (2013) 5840-5844.

86 Nelson, K. E., Lazcano, A., Miller, S. L.: Peptide Nucleic Acids as a Possible Primordial Genetic Polymer. First Steps in the Origin of Life in the Universe, Proceedings of the Trieste Conference on Chemical Evolution, 6th, Trieste, Italy, Sept. 18-22, 2000, (2001) 3-10.

87 Miyakawa, S., Cleaves, H. J., Miller, S. L.: The Cold Origin of Life: B. Implications Based on Pyrimidines and Purines Produced From Frozen Ammonium Cyanide Solutions. Orig Life Evol Biosph 32, (2002) 209-218.

88 Ferris, J. P., Joshi, P. C., Edelson, E. H., Lawless, J. G.: Chemical Evolution. XXX. Hydrogen Cyanide: A Plausible Source of Purines, Pyrimidines and Amino Acids on the Primitive Earth. J Mol Evol 11, (1978) 293-311.

89 Ferris, J. P., Joshi, P. C., Lawless, J. G.: Chemical Evolution. XXIX. Pyrimidines from Hydrogen Cyanide. BioSystems 9, (1977) 81-86.

90 Krishnamurthy, R.: Life's Biological Chemistry: A Destiny or Destination Starting from Prebiotic Chemistry? Chem Eur J 24, (2018) 16708-16715.

91 Krishnamurthy, R.: On the Emergence of RNA. Israel J Chem 55, (2015) 837-850.

92 Miller, S. L., Orgel, L. E.: The Origins of Life on the Earth, Prentice Hall Inc, New Jersey (1973).

93 Wills, C., Bada, J.: The Spark of Life: Darwin and the Primeval Soup. Oxford University Press, Oxford (2001).

94 Raggi, L., Bada, J. L., Lazcano, A.: On the Lack of Evolutionary Continuity Between Prebiotic Peptides and Extant Enzymes. Phys Chem Chem Phys 18, (2016) 20028-20032.

95 Forsythe, J. G., Krishnamurthy, R., Fernandez, F.: The Story Behind the Depsipeptide Project in the CCE, <http://www.scripps.edu/krishnamurthy/The%20Prebiotic%20Depsipeptide%20Story.html> (2017).

96 Weber, A. L.: Thermal Synthesis and Hydrolysis of Polyglyceric Acid. Origins Life Evol. Biosphere 19, (1989) 7-19.

97 Goodman, M., Stueben, K. C.: Peptide Syntheses via Amino Acid Active Esters. J Am Chem Soc 81, (1959) 3980-3983.

98 Pillon, L. Z., Utracki, L. A.: Compatibilization of Polyester/Polyamide Blends via Catalytic Ester-Amide Interchange Reaction. Polym Eng Sci 24, (1984) 1300-1305.

99 Hoegberg, T., Stroem, P., Ebner, M., Raemsby, S.: Cyanide as an Efficient and Mild Catalyst in the Sminolysis of Esters. J Org Chem 52, (1987) 2033-2036.

100 Miller, S. L., Urey, H. C.: Organic Compound Synthesis on the Primitive Earth. Several Questions about the Origin of Life have been Answered, but Much Remains to be Studied. Science 130, (1959) 245-251.

101 Huber, C., Wächtershäuser, G.: α-Hydroxy and α-Amino

Acids Under Possible Hadean, Volcanic Origin-of-Life Conditions. Science 314, (2006) 630-632.

102 Peltzer, E. T., Bada, J. L.: [alpha]-Hydroxycarboxylic Acids in the Murchison Meteorite. Nature 272, (1978) 443-444.

103 Rich, A.: The Possible Participation of Esters as well as Amides in Prebiotic Polymers. In: Buvet, R. Ponnamperuma, C. (eds.) Chemical Evolution and the Origin of Life: Proceedings of the International Conference on the Origin of Life 180-196 (North-Holland, Amsterdam (1971).

104 Fahnestock, S., Rich, A.: Ribosome-Catalyzed Polyester Formation. Science 173, (1971) 340-343.

105 Mamajanov, I., MacDonald, P. J., Ying, J., Duncanson, D. M., Dowdy, G. R., Walker, C. A., Engelhart, A. E., Fernández, F. M., Grover, M. A., Hud, N. V., Schork, F. J.: Ester Formation and Hydrolysis during Wet–Dry Cycles: Generation of Far-from-Equilibrium Polymers in a Model Prebiotic Reaction. Macromol 47, (2014) 1334-1343.

106 Forsythe, J. G., Yu, S. S., Mamajanov, I., Grover, M. A., Krishnamurthy, R., Fernández, F. M., Hud, N. V.: Ester-Mediated Amide Bond Formation Driven by Wet–Dry Cycles: A Possible Path to Polypeptides on the Prebiotic Earth. Angew Chem Int Ed 127, (2015) 10009-10013.

107 Forsythe, J. G., Petrov, A. S., Millar, C. A., Yu, S. S., Krishnamurthy, R., Grover, M. A., Hud, N. V., Fernández, F. M.: Surveying the Sequence Diversity of Model Prebiotic Peptides by Mass Spectrometry. Proc Natl Acad Sci USA 114, (2017) E7652-E7659.

108 Stoop, M., Meher, G., Karri, P., Krishnamurthy, R.: Chemical Etiology of Nucleic Acid Structure: The Pentulofuranosyl Oligonucleotide Systems: The (1'□3')-β-L-Ribulo, (4'□3')-α-L-Xylulo, and (1'□3')-α-L-Xylulo Nucleic Acids. Chem Eur J 19, (2013) 15336-15345.

109 Efthymiou, T., Gavette, J., Stoop, M., Froeyen, M., Herdejin, P., Krishnamurthy, R.: Chimeric XNA: An Unconventional Design for Orthogonal Informational Systems. Chem Eur J 24, (2018) 12811-12819.

110 Gavette, J. V., Stoop, M., Hud, N. V., Krishnamurthy, R.: RNA–DNA Chimeras in the Context of an RNA World Transition to an RNA/DNA World. Angew Chemie Int Ed 55, (2016) 13204-13209.

111 Bhowmik, S., Krishnamurthy, R.: The Role of Sugar-backbone Heterogeneity and Chimeras in the Simultaneous Emergence of RNA and DNA. Nat Chem 11, (2019) 1009-1018.

112 Orgel, L. E., Lohrmann, R.: Prebiotic Chemistry and Nucleic Acid Replication. Acc Chem Res 7, (1974) 368-377.

113 Stephen-Sherwood, E., Odom, D. G., Oro, J.: Prebiotic Synthesis of Deoxythymidine Oligonucleotides. J Mol Evol 3, (1974) 323-330.

114 Sutherland, J. D., Whitfield, J. N.: Prebiotic Chemistry: a Bioorganic Perspective. Tetrahedron 53, (1997) 11493-11527.

115 Dworkin J. P., Lazcano, A., Miller Stanley, L. The Roads to and from the RNA World. J Theor Biol 222, (2003) 127-134.

116 Follmann, H.: Deoxyribonucleotides: the Unusual Chemistry and Biochemistry of DNA Precursors. Chem Soc Rev 33, (2004) 225-233.

117 Powner, M. W., Zheng, S.-L., Szostak, J. W.: Multicomponent Assembly of Proposed DNA Precursors in Water. J Am Chem Soc 134, (2012) 13889-13895.

118 Wächtershäuser, G.: The Place of RNA in the Origin and Early Evolution of the Genetic Machinery. Life 4, (2014) 1050-1091.

119 Xu, J., Chmela, V., Green, N. J., Russel, D. A., Janicki, M. J., Góra, R. W., Szabla, R., Bond, A. D., Sutherland, J. D.: Selective Prebiotic Formation of RNA Pyrimidine and DNA Purine Nucleosides. Nature 582, (2020) 60-66.

120 Le Vay, K., Mutschler, H.: The Difficult Case of an RNA-only Origin of Life. Emerg Topics Life Sci 3, (2019) 469-475.

121 Lopez, A., Fiore, M.: Investigating Prebiotic Protocells for A Comprehensive Understanding of the Origins of Life: A Prebiotic Systems Chemistry Perspective. Life 9, (2019) 49.

122 Ritson, D., Sutherland, J. D.: Prebiotic Synthesis of Simple Sugars by Photoredox Systems Chemistry. J Mol Evol 4, (2012) 895-899.

123 Strazewski, P.: The Beginning of Systems Chemistry. Life 9, (2019) 11.

124 Guttenberg, N., Virgo, N., Chandru, K., Scharf, C., Mamajanov, I.: Bulk Measurements of Messy Chemistries are Needed for a Theory of the Origins of Life. Phil Trans R Soc A 375, (2017), 20160347, 1-6.

125 Walker, S. I., Packard, N., Cody, G. D.: Re-conceptualizing the Origins of Life. Phil Trans R Soc A 375, (2017) 20160337, 1-11.

126 Suárez-Marina, I., Abdul-Haja, Y. M., Turk-MacLeod, R., Gromski, P. S., Cooper, G. J. T., Olivé, A. O., Colón-Santos, S., Cronin, L.: Integrated Synthesis of Nucleotide and Nucleosides Influenced by Amino Acids. Commun Chem 2, (2019) 28.

127 Eschenmoser, A., Kisakürek, V. M.: Chemistry and the Origin of Life. Helv Chim Acta 79, (1996) 1249-1259.