ISSN 2075-1729
www.mdpi.com/journal/life

Peer-Review Record:

What RNA World? Why a Peptide/RNA Partnership Merits Renewed Experimental Attention

Charles W. Carter, Jr.

***Life* 2015, *5*, 294-320, doi:10.3390/life5010294**

Reviewer 1: Anonymous
Reviewer 2: Anonymous
Editor: Niles Lehman (Guest Editor of Special Issue "The Origins and Early Evolution of RNA")

Received: 24 November 2014 / Accepted: 12 January 2015 / Published: 23 January 2015

First Round of Evaluation

Round 1: Reviewer 1 Report and Author Response

This is a review and therefore can not, as the form asks, actually be original. However, in the context of a review it is quite adequate. The case is made that coded peptide synthesis came about by ongoing interactions between peptides and oligonucleotides without the necessity of an RNA World. The emphasis on what can be learned by attempting to unravel the evolutionary history of the synthetases. Refreshingly, no attempt is made to deduce the origin of specific codons. The emphasis on investigating the catalytic potential of small peptides in various origin of life contexts is especially important. Overall, this is a good review of the relevant literature in this area. Minor comments as follows:

(1) It is unfair to male readers to refer to colors on figures as wheat, firebrick, etc as they are typically only aware of red, green, blue, yellow and shades thereof.

Response: These colors are those used in PYMOL, and hence should be familiar to a wide user base. However, as a male with marginal color discrimination myself, but who can distinguish the colors used, I have nonetheless tried to accommodate the reviewer in most cases.

(2) In multiple/most cases the figure captions do not adequately explain the figures. In Figure 5, for example, it is not clear what blue and green are and there is no CP1 label as the caption suggests. All the figure captions should be reviewed.

Response: Reviewer 2 also raised numerous concerns about the figure legends. All have been revised as requested. Figure 5 (now Figure 4) has been revised to indicate CP1.

(3) The text on Line 146 is garbled-probably should read "led to greatly".

Response: Corrected

(4) In the vicinity of Line 154, the diminished capabilities of the aaRS Urzymes is discussed. It should be noted that this may be the result of evolutionary adaptation to inclusion in the context of the modern enzyme. An interesting experiment (I am not requesting author to do it) might be to attempt to evolve the extracted urzyme to determine if a few sequences changes can restore some of the lost activity. It is unfortunate that such experiments are much harder with peptides than they are with RNA.

Response: The paragraph at the bottom of p. 7 in Section 2 is extensively revised to address all aspects of this concern, including reference to the mutation of the TrpRS Urzyme active-site residue D146A and its effect to increase catalytic activity and the possibility, which we have recently validated, that Urzymes likely differ functionally in some respects, by virtue of the additional masses in the contemporary enzymes.

(5) In the vicinity of line 184 it would seem appropriate to reference and perhaps comment on the work of Turk and Yarus relating to amino acid activation using very small RNAs.

Response: See above response to Reviewer 1 on the inclusion of Section 3.7.1, which addresses this concern as well.

(6) Lines 205–206 make the author's point very well!

Response: Many thanks! Each time I read this compliment, I have to locate what was said in the original.

(7) On Line 254, reference 40 does not seem correct. I see no mention of L2 or L3 in this reference.

Response: Excellent point. Reference 40 discusses the resistance of the PTase to protein extraction procedures. As indicated by a new citation of personal communication, I had a recent communication with Harry Noller in which he amplified what was not explicit in that paper: L2 and L3 were always present whenever activity was observed. If necessary, I can provide that email. However, it is stored at a remote location to which I have no access at the moment.

(8) What is status of reference 15? Nature likely decided by now.

Response: This manuscript has been problematic for various reasons related to the quite unusual nature of both the approaches and the results. Three Nature journals declined to review it; one explicitly replied that it was largely already published in Carter, et al., Biology Direct, (2014) 9:1. I am currently revising a submission to JBC, which has been reviewed and rejected, but which the managing Associate Editor has agreed to return to the original reviewers along with my extensive responses to their concerns.

Round 1: Reviewer 2 Report and Author Response

In this manuscript the author reviews the evidence for a model of life's origins that invokes assemblies of peptide-RNA complexes rather than a purely RNA World. Most of the work reviewed emerges from the author's laboratory and centers on his groups' discovery that fairly small domains of tRNA synthetases from both structural classes are capable of amino acid activation and tRNA

aminoacylation. In addition, the author's work has elaborated upon and tested the Rodin-Ohno hypothesis that two families of enzymes are encoded by sense and antisense strands of the same gene, and his group has demonstrated a new bioinformatic metric for detecting the occurrence of examples. I am personally convinced that this work represents a very important contribution to our understanding of life's origins and am happy to see this review. I think the manuscript is generally solid but it should benefit from consideration of the following points:

- I do not think the author does a good job in promoting his ideas by drawing such a sharp distinction between his RNA-peptide world and the "RNA World". Contrary to his statement (p. 3) there are many RNA World proponents who do not go into fits of apoplexy at the mention of the word "peptide". [At p. 9, bottom, the phrase "blind devotion" is also over the top]. A very good formulation of the RNA World was given by Joyce & Orgel, in The RNA World (2nd ed, 1999). These authors suggest that the only required aspects of the RNA World are that genetic continuity was assured by RNA replication grounded in WC base-pairing, and that genetically encoded proteins were not involved as catalysts. This leaves plenty of room for considering how non-coded peptides may have played an important role.

Response: I agree with the reviewer's point that hyperbole tends to weaken the manuscript; examples including those mentioned here have been deleted or reworded. I am most grateful to the reviewer for suggesting the earlier defense of the RNA World hypothesis by Joyce and Orgel, because it provides a framework missing in the original draft of this manuscript. For that reason, I have emphasized at various points the importance of rudimentary versions of coding implicit in the Carter & Kraut model, and expanded discussion in the introduction (p. 4) and a new section of results (Section 3.7, p. 20) of how that model incorporates origins of replication, catalysis and coding.

I disagree, however, that the Joyce and Orgel formulation is in any way consistent with what is presented in this manuscript. Specifically, I argue (hopefully more coherently in the revision) that replication, specificity, catalysis, and coding all developed gradually and in concert from the stereochemical complementarity of the two classes of polymers, and hence that the RNA World formulation is based on the wrong assumptions. Thus, although the Joyce and Orgel argument may accommodate the marginal participation of peptides, it fails completely to incorporate peptides in an equal partnership in the emergence of the three properties necessary for molecular biology to evolve rapidly.

- The new distinction insisted on by the authors here, of course, is that the protein and RNA may have templated each other (reciprocal prebiotic autocatalysis, p. 2). This is a genuinely distinct notion, and it should be contrasted explicitly with the key tenets of the RNA World hypothesis—either as formulated by Joyce and Orgel or by others. Perhaps the author's proposal really does do a better job of providing a testable hypothesis for the origin of contemporary coding than anything an RNA World proponent has come up with. If so, that is the probably the most important reason why we should take the Carter-Kraut model seriously. If the authors could lay out these ideas more clearly in the beginning it would provide a stronger foundation for what follows.

Response: The introduction has been revised to establish the comparison with the Joyce-Orgel articulation of the RNA World hypothesis and a new section, Section 3.7, has been included to make the detailed comparisons requested by the reviewer.

- The author has a tendency toward hyperbole that detracts from the message. For example, at the top of p. 4 he states that Urzymes retain "impressive fractions" of native aaRS activities, rather than simply indicating what the fractions are and letting the reader judge whether they are impressive or not. On p. 8 at the end of the penultimate paragraph, the word "stunning" referring to the author's own observations may cause eyes to roll (it did for me). On p. 12, last paragraph, I do not think acceptor stem base coding should be described as "unambiguous", At p. 15, bottom, again, please let the readers decide if Urzymes are a watershed or not.

Response: Examples of hyperbole have been deleted and each example cited here has been revised to tone down self-promotion. Description of the acceptor stem operational code remains as "unambiguous" for lack of a better word. The paper describing this work is only being submitted to PNAS by Dr. Wolfenden as a member, because he wished to finish work on a new paper describing the new data on amino acid phase transfer free energies before submitting that paper. This creates a somewhat sticky situation, as that paper is not yet available, as I thought it would be when I wrote this paper, having completed it eight months ago. Similar issues apply to the paper about the catalytic activity of the 46mer, as a revision of that paper is being prepared for re-submission to JBC.

- At p. 8, bottom, the comments about allostery come out of the blue and then go back into the blue. Can this be expanded to provide some more background and context? This is a really important point that will be lost in the present form of the writing.

Response: I sympathize with the reviewer here. I now refer at the top of what is now p. 11 to the paragraph in question in Section 3.6, now on p. 19–20, which has been amplified and should clear up the motivation for introducing the idea of allostery as an essential component of the specificity necessary for a fully developed genetic code.

- On p 8, the sentence beginning "The parallel catalytic proficiencies..." lacks punctuation. More importantly, the distinction between the catalytic domain size and the Urzyme size is exaggerated and misleading. Class I Rossmann folds, for example, only consist of 5–6 parallel beta strands with connecting helices, and they are often no more than 150 residues in size (sometimes less), not 250. Urzymes are 125 or 130 amino acids in size (Figure 3 legend), so the Urzymes are not significantly smaller than the catalytic domains.

Response: I'm grateful to the referee for pointing out the stylistic problems in this paragraph and have worked the text, now in the middle of p. 9, to resolve some confusion evident in the rest of this critique. With regard to the relative sizes of Urzymes and other enzymes, both Class I and II aaRS have long insertion (sub)domains within their catalytic domains. In particular, catalytic domains in Class I are all substantially larger (by 76–160 residues) than the smallest Rossmannoid proteins (e.g., response receivers) which do indeed have ~130 residues. With this increased size comes significant functional sophistication. We have now demonstrated that the Class I insertion participates in allosteric behavior, which is discussed more fully in response to a previous concern. New citations are given to put this size difference into proper perspective in the discussion in the first paragraph of p. 11, showing that Urzymes indeed are smaller than the catalytic domains by many multiples of the standard error of the mean size. See also response to a similar concern below.

- p. 7 top, it is stated that Urzymes are 5 logs down in activity as compared to contemporary aaRS, but inspection of Fig 4b seems to suggest that 6–7 logs down is more accurate. Can the author provide a Table citing actual representative catalytic parameters for wt and a few of the Urzymes from each class?

Response: The reviewer is correct. The constructs we have called "Urzymes" exhibit a range of catalytic activities. The most active are ~100-fold more active than the least active. We have clarified this in a revised paragraph, now at the bottom of p. 7 in Section 2. Table 1 has been included on p. 9 with experimental apparent second order rate constants.

- On p. 7–8 the authors mention a 46 amino acid peptide and then ~50 amino acid peptide ("...and now to the ATP binding sites of ~50 residues...") However, it is not made clear until p. 10 what the actual activity of the "~50 amino acid" ATP binding site is—and it is never specified what sections of the class I and class II Urzymes or full length enzymes these polypeptides represent. It seems implausible that such a small polypeptide would stably fold given that it is not wrapping around a metal ion and has not been selected for stability. The author should provide a drawing of the contemporary class I and class II catalytic domains with color indicating where this polypeptide lies. Since there are no stable 50 aa subdomains in either class of contemporary aaRS, this figure would necessarily show that the 50 aa peptide depends on surrounding regions of the protein for structure and stability. How, then, can it in an isolated state bind not just ATP but also amino acid, and then accelerate rate enhancement by 40–60,000 fold (p. 10)? Presumably this enhancement is with respect to the very slow uncatalyzed rate? How are such slow rates measured? There is nothing in the Experimental Section about this. Even though this is a review paper, the claim made for two-substrate binding and then catalysis by these small peptides is pretty exceptional, and it needs to be validated here, not just by reference to another experimental paper. Further, the authors cite the Mildvan papers (refs. 44–48) as precedent. I am not familiar with the details of this work, but reading through the titles, it looks like Mildvan showed ATP binding by 50-mer peptides by various methods, but maybe not ATPase activity or some other activity? The sentence citing Mildvan does not specify.

Response: Again, the reviewer has made salient points. The activities of the 46-residue peptides are exceptionally surprising, and the discussion about this work is hampered by the fact that the experimental paper is under revision and hence not available. Nevertheless, I have done what I could to clarify the various concerns voiced in this long paragraph: (i) estimation of the uncatalyzed rate of amino acid activation is clarified in the first paragraph of Section 3, p. 9. (ii) Figure 1 has been enhanced by inclusion of Figure 1a, which is the requested drawing showing the location of the 46-residue ATP binding sites of both aaRS classes. (iii) Section 3.4 has been extensively revised to place the work of Mildvan in proper context as background for the work described here and in reference [18], which is provided for the referees benefit. (iv) Activities of the two 46-residue peptides, with both wild type sequences and the peptides from the sense antisense gene are, indeed, a very big deal. The relevant data from reference [18] are summarized paragraph in Section 3.4 at the top of p. 13.

- There is a small black line on the left side of Figure 4b that seems to have something to do with a 46 aa ATP binding peptide. What is this representing? Figure 4 in general also does not say whether it is depicting amino acid activation or tRNA aminoacylation.

Response: The legend for Figure 4 now explicitly states that the activities represented are for amino acid activation, as measured by 32PPi exchange.

- On p. 8 the authors state that they have demonstrated "irrefutably" that peptide-based catalysis and specificity are striking attributes of peptides far shorter than contemporary enzymes. However, there are of course examples of contemporary enzymes that are 125–130 amino acids in size, the same size as the Urzymes. Therefore, this claim rests on the capacity of the 46-mers to catalyze aminoacyl adenylate synthesis, which should be substantiated as mentioned above.

Response: The reviewer correctly points out, first, that there are enzymes as small as the aaRS Urzymes and, hence, that our claim for catalytic activity in much shorter peptides rests heavily on the activity shown for the 46mers. Several points should be made. (i) We believe that by showing that active site mutations eliminate the activity of the 46mers we have established the authenticity of their activities. (ii) the enzymes that are only ~125–130 amino acids long are hydrolases and isomerases, which are relatively unsophisticated activities requiring but a single substrate. Synthetases represent more complex enzymes that bind more than one substrate. (iii) The Urzymes are surprising not only in their catalytic proficiencies, but also because they acylate tRNA and especially because they retain the activated amino acids with high affinity. This material has been included in Section 3.1, p. 10, with new references to archival data on enzyme masses.

- P. 8 bottom— what is meant by 20% of the specificities? If a typical nonediting aaRS has 5 logs of specificity against a noncognate aa for the activation step, does this mean that the Urzymes have over 4 logs of specificity? That does not appear to be the case from Figure 6. Such an interpretation is also not consistent with the production of statistical proteins (Conclusion). Please also rewrite the Figure 6 legend so that it is more descriptive and intelligible.

Response: Section 3.2 has been revised to clarify that the percentages discussed here derive, as they should, from the relative Gibbs energies of specific recognition. The caption to Figure 6 has also been revised to state explicitly that specificities plotted are differences in Gibbs energies for cognate and all non-cognate amino acid k_{cat}/K_M values.

- P. 11 bottom, the paragraph about the binary coding information and regression models is very very condensed and difficult to penetrate. Can this be rewritten and preferably expanded so we see a little more of what is going on here.

Response: Section 3.4 has been sub-divided into separate sub-sections, in which responses to this and other concerns are addressed in detail. In particular Section 3.4.2 now expands the description of our (submitted) investigation of tRNA coding properties.

- P. 10, bottom, please define MBP

Response: MBP (middle (codon) base-pairing) is adequately defined in the last paragraph of Section 2. The abbreviation has been avoided elsewhere, except in the last paragraph of Section 3.4.1, where it is again spelled out.

- Figure 5 legend, the N and C-termini of CP1 cannot be separated by exactly the distance of a peptide bond, 1.5 Angstroms or so.

Response: The text in the legend to Figure 4 (note change in number) has been revised to specify that it is the alpha carbons left at the C- and N-termini after removing CP1 that are separated by this distance. As surprising as this may be, it is true and can be readily verified using PYMOL.

- In Figure 9 the legend needs to make clear precisely what all the differently shaped and colored icons in the bottom part of the figure represent.

Response: The legend has been revised as requested.

- P. 4, top—"The catalytic power of peptides related by phylogeny to contemporary enzymes is thus far greater than was anticipated..."—please provide citations for what was anticipated.

Response: The first par. at the top of p. 5 has been rewritten to explain the meaning of "anticipated".

- P. 15, third paragraph from the end, cite to Figure 9 should be to Figure 10.

Response: Corrected.

- P. 15, first paragraph "...most secluded nook", the author should indicate "as discovered so far" or some similar limiting statement. Many biological nooks remain unpenetrated by human scientists. Also, there should be some citation, somewhere, to the Bartel ribozyme replicases, which show that in vitro RNA can replicate itself. That point should be forthrightly acknowledged as favoring a pure RNA World hypothesis.

Response: As part of the extended response to the initial concerns voiced by reviewer 1, I have included in Section 3.7.1, a summary of the work on ribozymal replicases, as well as the work of M. Yarus on other sorts of aptamers.

I disagree strongly, however, that the Bartel and other ribozymal replicases are evidence favoring a pure RNA world. In order for that to be true, it has to be shown that there is a path to their evolution that is much faster than first evolving humans and developing 21st century technology, and that doesn't involve peptide polymerases.