

Elements of Loren's Style:  
Manuscript/Thesis/Figure Preparation in the Williams Group  
Loren Dean Williams, Georgia Tech  
revised 9/30/2023

**Table of Contents**

<b>General Principles</b>	<b>3</b>
<b>The Outline</b>	<b>3</b>
<b>The Abstract</b>	<b>3</b>
<b>The Introduction</b>	<b>4</b>
<b>The Results</b>	<b>5</b>
<b>The Discussion</b>	<b>5</b>
<b>The Figures</b>	<b>6</b>
<b>The Central Figure.</b>	<b>6</b>
<b>Figure Legends</b>	<b>8</b>
<b>The References.</b>	<b>8</b>
<b>Internal review before submission.</b>	<b>8</b>
<b>Peer Review.</b>	<b>9</b>
<b>Example letter of submission</b>	<b>9</b>
<b>Example Editor response to the submission</b>	<b>10</b>
<b>Example response to the editor and reviewers.</b>	<b>10</b>
<b>Introduction to Responses to Reviewers</b>	<b>11</b>
Response to Reviewer 1	11
Response to Reviewer 2	13
<b>Example editor response to revision.</b>	<b>15</b>
<b>Thesis attribution statement.</b>	<b>15</b>
<b>Elements of style.</b>	<b>16</b>
<b>Use Oxford Commas</b>	<b>16</b>
<b>Use clear declarative topic sentences.</b>	<b>16</b>
<b>Do not use "this" as the subject of a sentence.</b>	<b>16</b>
<b>Avoid "This is..." "There exists..."</b>	<b>16</b>
<b>Never use the word "ones".</b>	<b>16</b>
<b>Never use the word "respectively".</b>	<b>16</b>
<b>Use the active voice.</b>	<b>17</b>
<b>Complete all comparisons.</b>	<b>17</b>

Use “which” and “that” correctly.	17
Do not confuse kinetics and thermodynamics	18
Use parallel construction.	18
Do not tell the reader what is interesting or significant or intriguing important.	19
Use consistent nomenclature and syntax throughout a document.	19
Use signal words.	19
Do not use contractions.	19
Use strong topic sentences.	19
Tenses (adapted from Nature, Effective Writing)	19
<i>Analyze and Improve your writing.</i>	20
<b>Figures:</b>	22
Adobe Illustrator	22
Work Flow	22
Graphical File Types and Uses	22
<b>VECTOR</b>	22
AI (Illustrator)	22
PDF	22
EPS	23
SVG	23
<b>PIXILATED</b>	23
PSD	23
JPG	23
PNG	23
GIF	23
TIFF	23
Colors	23
<i>Dropbox. Organization of Manuscript Folders.</i>	25
<i>Credits for this document.</i>	26

## General Principles

Doing science is only part of our job. Communicating science it is another part. Firstly, writing our science teaches us our science. Explaining complex concepts causes is to think deeply and leads to ever deeper understanding. We reduce complex concepts so that people who are interested but not expert can understand and appreciate them. If we write things that are rigorously correct but cannot be understood by non-experts than we will not succeed.

Each section of a manuscript should be organized in a clear and logical thread. A topic follows from the preceding topic. Topics are introduced by headings. Headings evolve as the manuscript is written. Initial headings are the points listed in an outline. In the final manuscript, headings can be as frequent as one every other paragraph. Headings make it easy for readers to follow the progression of the manuscript, and to scan to find what they want.

Every section of a manuscript (abstract, introduction, results, methods...) has a well-defined function. Each section should be functionally independent of the rest of the manuscript. Readers often scan a manuscript to find selected information. The majority of readers will read only portions of your manuscript. A manuscript must be constructed to accommodate scanning. Most will be interested only in the abstract, introduction or discussion. Some will look only at the figures and figure legends. A few will read the results and methods sections. When was the last time you read a manuscript from front to back?

## The Outline

Make an initial outline of your paper before your experiments are completed. The outline should contain a list of figures. As the outline is being worked out, make draft figures. For your initial outline and figures you will be missing some of the data. Anticipate the data. Sketch graphs of the anticipated data. These graphs of hypothetical data are an important part of the hypothesis-testing process that ensures that you understand your experiments and are moving as directly as possible towards a publishable manuscript.

Once the outline of the paper is complete and the initial figures are constructed, write the first draft of the results section. Start this process by describing your figures in detail - as if to a person who cannot see them.

## The Abstract

[adapted from Nature; Instructions to Authors] The abstract is written last, but must be constructed very carefully, as it is the most accessible (electronically) and widely read part of the manuscript. The abstract and the key words direct search engines to your paper (or not). The

abstract must stand alone, and can be redundant with the most important parts of the manuscript. The abstract will be read by people who do not have access to the manuscript.

The abstract should contain the following elements [adapted from Nature].

- (i) Broad introduction for any a scientist in any discipline.
- (ii) Detailed background for scientists in your field.
- (iii) Statement of the problem.
- (iv) Summary of the primary result.
- (v) Explanation of the significance of the result.
- (vi) Statement of general context of the result.
- (vii) Statement of broader context.

The origin of life on earth poses some of the most profound and exciting questions in science and philosophy, and tests our understanding of chemical and biological principles. Looking into deep time, Woese and Fox sketched out a timeline of life as a tree with three primary branches. Their telescope was the ribosome, which exists in every cell and at its core is universally conserved. Structures of ribosomes in three-dimensions, which are now available, allow one to dissect primordial molecules, reactions and events. We ask if the timeline of life can be visualized beyond the root of the tree, to the biochemical origins. Here, using a three-dimensional comparative method, based on structural fingerprints in known expansions in eukaryotes, we establish a comprehensive and coherent model of the evolution of the prokaryotic ribosome. This atomic level model reconciles the histories of the LSU, the SSU, tRNA and mRNA to a common timeline. In the model, rRNAs grow by accretion; recursively expanding and adding successive layers, iteratively growing, subsuming and freezing the rRNA. The model outlines the timing of acquisition of functions such as catalysis, decoding, energy transduction and translocation. The SSU was smaller than the LSU during initial subunit assembly, which was mediated by tRNA. At initial subunit assembly, proto-mRNA bound to the SSU and acted as a cofactor. Proto-mRNA positioned the activated ends of tRNAs within the peptidyl transferase center, which catalyzed the production of the earliest polypeptides. The model allows us to visualize essential processes during the transition from chemistry to biology. This work helps explain how RNA joined forces with the first polypeptides to create the ribosome. These molecular machines embody a symbiotic relationship, that of protein and nucleic acids, which took root during the origin of the ribosome and evolved into life as we know it.

## The Introduction

The purpose of the introduction is to frame your results in the context of the literature. The introduction should contain 1) a general background, 2) a specific background that justifies the need for the research, 3) the specific purpose of the research, 4) objectives (hypothesis), and 5) a summary of the approach taken. Within the first four paragraphs, the introduction must contain a short clear statement of what was done, as in:

"Here we use computation and experiment to determine if  $\text{Fe}^{2+}$  can substitute for  $\text{Mg}^{2+}$  in RNA folding and catalysis."

Certain rhetorical devices allow you to unambiguously distinguish your new work from previous background work. "'Here we use...'", means we did it. "Previously it was shown..." or "Hud and coworkers demonstrated..." means someone else did it and it is already published.

## The Results

This section is a logical and objective presentation of the experiments conducted and of the results obtained without interpretation of their significance. The results section is directed at scientists with expertise in methods used in the paper. The results section should convince an informed and skeptical reader that the conclusions are fully supported by the experiments. In the results section you should describe your figures so that they would be interpretable to a blind person – who could not see the figures.

Following is an example paragraph taken from the results section Anton's L2 paper. Note (a) the heading, (b) The topic sentence that states the paragraph's conclusion in the first sentence, written in an active sense, (c) the level of detail in the description, (d) the integration of figures, tables and text, (e) relatively simple and declarative sentence structure, and (f) the unambiguous distinction between Anton's results and other work.

*Polarization.* RNA and magnesium ions enhance the stability of the D2-AMN complex indirectly, without direct interaction with the rProtein. Stabilization is seen by analysis of an analogous complex that omits the rRNA and the magnesium ions, in the form of (H<sub>2</sub>O)<sub>4</sub>-AMN (Figure 3B). This omission causes the water molecules to depolarize, attenuating the interaction energy. As shown by the results of the NEDA analysis (Table 5) the polarization component in the D2-AMN complex (-39.1 kcal/mol) is substantially larger than that in the (H<sub>2</sub>O)<sub>4</sub>-AMN complex (-19.3 kcal/mol). Thus, the polarization of water molecules by the magnesium ions plays a significant role in the stability of the rRNA-magnesium-water-protein complexes. Our result here is consistent with previous observations (18) that the pKa of water molecules decreases upon incorporation into the magnesium first shell.

## The Discussion

The first goal of the discussion section is to restate the results for a general reader who does not read the results section. In addition, the discussion must explain the importance of the results, place them in context of what is already known, and suggest future directions.

Although some combine the results and discussion sections in a published manuscript, we do not. It can be useful to write a combined results and discussion section in initial drafts. This way, in the beginning, one can avoid making sometimes difficult decisions on what constitutes results and what constitutes discussion. In later drafts the combined sections can be disentangled to give a distinct results section and a discussion section.

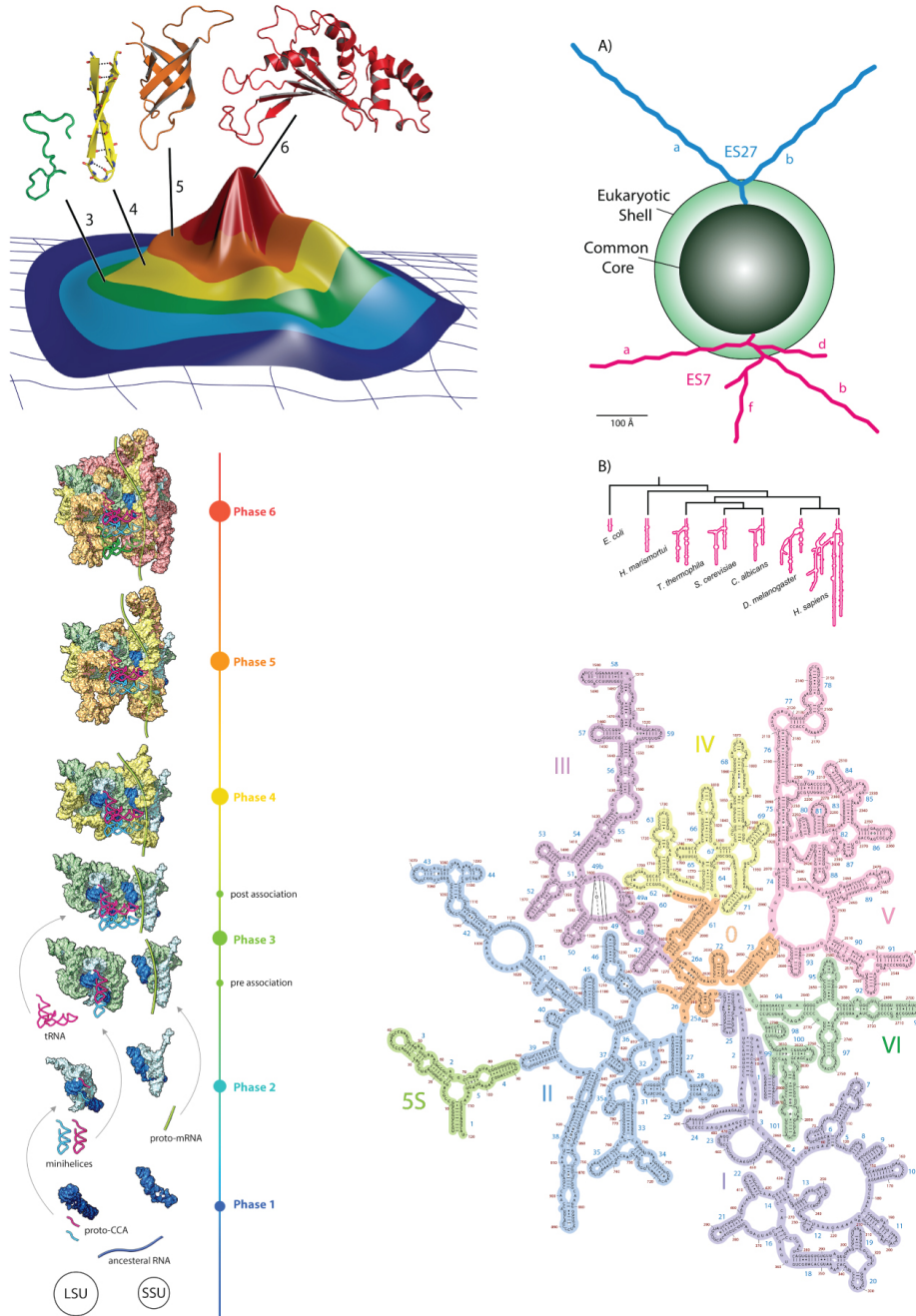
## The Figures

We want our figures to be exceptionally clear, informative, esthetic and creative. We have the technology and expertise to illustrate any conceivable concept. Let's do it right.

Figures provide the ultimate organizing structure of a scientific manuscript. Every figure must have a clear and well-defined purpose. Making a good figure is a laborious and iterative process. The final version of every figure from the Williams group must be in Adobe Illustrator. The version we submit to the journal will be written out from the Illustrator file as a jpg or png. Do not use Powerpoint for making figures for papers of theses.

### The Central Figure.

Every manuscript should contain a Central Figure that broadly summarizes the state of knowledge and can be understood by generalists. The Central Figure is intended to be useful to people inside and outside our lab, in talks and reviews. We want others to want to repurpose our central figures in reviews and talks. The Central Figure is often a clear schematic that summarizes the information in manuscript and the current state of knowledge.



Central Figures: Nicholas' folding-fitness figure (top left), human expansion segments figure (top right), Chad/Anton's schematic of ribosome evolution (bottom left) and our secondary structures (bottom right). The success of a central figure is indicated by how frequently it is used by people inside and outside our lab - in reviews, talks, icons for a apps, etc.

## Figure Legends

Within a Figure Legend, call each figure panels once, only once and in the correct order.

### No, don't do this

**Figure 1: RNA Goldilocks landscapes are predicted by simulations.** (A) Simulations reveal the influence of  $[Mg^{2+}]$  on the chemical lifetime of an RNA that is cleaved more slowly in the folded state than in the unfolded state (black/red line). The Goldilocks peak is highlighted in red. The lifetime of a never folded RNA is shown by a dashed line ( $k_u = 1 t_{rel}^{-1}[Mg^{2+}]_{rel}^{-1}$ ). The lifetime of an always folded RNA is shown by a dotted line ( $k_f = 0.2 t_{rel}^{-1}[Mg^{2+}]_{rel}^{-1}$ ). An RNA that shifts between unfolded and folded states shifts between unfolded and folded lifetimes, to establish a Goldilocks peak. Goldilocks behavior requires conversion from unfolded to folded and a slower cleavage constant of folded vs. unfolded RNA ( $k_f < k_u$ ). (B) The two-state reaction mechanism. U is unfolded RNA and F is folded RNA. (C-F) Effects while other parameters are held constant of (C)  $K_D$  (D)  $n$  (E)  $k_f$ , and (F)  $k_u$ . Each parameter was varied by multiplication or division by  $1+(0.1 \times 2^i)$  ( $i = 1, 2, 3, \dots, 8$ ). For this representation,  $[Mg^{2+}]$  was converted to  $[Mg^{2+}]_{rel}$  where  $1 [Mg^{2+}]_{rel} = K_D = 0.022 \text{ mM } Mg^{2+}$ . Lifetime ( $t$ ) was converted to  $t_{rel}$  where  $t_{rel} = 1$  when  $[Mg^{2+}]_{rel} = 1$  and the cleavage constant(s) are always  $1 t_{rel}^{-1}[Mg^{2+}]_{rel}^{-1}$ .

### Yes, do this.

**Figure 1: RNA Goldilocks landscapes are predicted by simulations.** (A) Simulations reveal the influence of divalent metal concentration  $[M^{2+}]$  on chemical lifetime. A Goldilocks peak, highlighted in red, is observed if the RNA is cleaved more slowly in the folded state than in the unfolded state (black/red line). The lifetime of a never-folded RNA is shown by a dashed line ( $k_u = 1 t_{rel}^{-1}[M^{2+}]_{rel}^{-1}$ ). The lifetime of an always-folded RNA is shown by a dotted line ( $k_f = 0.2 t_{rel}^{-1}[M^{2+}]_{rel}^{-1}$ ). An RNA that shifts between unfolded and folded states shifts between unfolded and folded lifetime traces, establishing a Goldilocks peak. (B) A two-state (folded/unfolded) reaction. U is unfolded RNA and F is folded RNA. (C) The effect of varying  $K_D$ , while holding all other parameters constant, in a simulation of folding landscape. (D) The effect of varying  $n$ , the Hill coefficient, on the folding landscape. (E) The effect of varying the cleavage rate constant for folded RNA  $k_f$ , on the folding landscape, (F) The effect of varying the cleavage rate constant for unfolded  $k_u$  on the folding landscape.  $[M^{2+}]_{rel} = [M^{2+}]/K_D$ .  $t_{rel} = 1$  when  $[M^{2+}]_{rel} = 1$  and the cleavage constant(s) are always  $1 t_{rel}^{-1}[M^{2+}]_{rel}^{-1}$ .

## The References.

We use endnote (see final section). There have been instances in which we have offended people, even our friends, by not citing them. We must work very hard to avoid this. Citing someone is a way of validating their work.

## Internal review before submission.



When a manuscript is thought to be complete and ready for submission, review a printed copy of the manuscript. The first author should conduct the final review. Read a hard copy of manuscript, in its entirety, in one sitting.

Some things to look for (this list is not comprehensive):

1. Figure calls:
  - a. Ensure each figure is called, in the right order and that the call points to the correct figure.
2. Acronyms:
  - a. Acronyms are defined both in the abstract (if used there, try to avoid it) and in the main body of the text.
  - b. Acronym definitions appear at the first occurrence. Use the acronym in place of its definition throughout the remainder of the document.
  - c. Acronyms should not be defined in headings or figure legends.
3. Units:
  - a. Appropriate and consistent use of symbols (e.g., **uM** vs. **μM**). Make sure that micromoles have not been converted to millimoles.
4. Figures and Figure Legends:
  - a. Check every figure and legend for accuracy and completeness.
  - b. Check the scale and font size of figures. Will the text be legible at published size? The minimum allowable font is 8 pt.
5. Citations:
  - a. Ensure that numbers are sequential.
  - b. Ensure that the numbers correspond to the correct reference. Do not place endnote entries in textboxes.
  - c. Ensure that journal abbreviations are correct.

## Peer Review.

Editor and reviewer comments are frequently unkind and maybe even ignorant. If we anticipate and respond appropriately to concerns of editors and reviewers, we maximize the chances that our papers will be accepted and will not be rejected or go back for extra rounds of review.

### Example letter of submission

Dr. *delete*  
Editor-in-Chief,  
Journal of *delete*

Dear *Delete*,

We submit the manuscript, "*Delete*" for consideration for publication in *Journal of Delete*.

The ribosome is the most ancient biological assembly, retaining a molecular record of early evolutionary history. We have developed methods of structural and sequence analysis of the ribosome that allow us, for the first time, to quantitatively understand and visualize the true extent of conservation of the ribosome.

To define the common core of the ribosome in three dimensions we have established a new statistic that we call Pairing Adjusted Sequence Entropy (PASE). This is a statistic that simultaneously characterizes conservation of sequence and of base pairing. PASE provides a net measure of similarity that controls for differential restraints on base paired nucleotides compared to unpaired nucleotides in RNAs with conserved three-dimensional structure. We combine PASE with structural analysis to define the common core of cytoplasmic rRNA.

Our results demonstrate that fully 90% of prokaryotic rRNA is contained within the common core, which provides a structural and functional foundation of rRNAs of cytoplasmic ribosomes of all living species. Our characterization of the common core allows us to investigate the relationship between ribosomal size, geological time, and organismal complexity.

Loren Williams is the corresponding author. All authors have reviewed the manuscript and approve submission. No authors have conflicts of interest. The work described in this manuscript has not been submitted to other journals or published previously.

We thank you for your consideration.

### Example Editor response to the submission

Dear Prof. Williams:

The in-depth review of your manuscript by the editors and the peer reviewers is now complete. Based on their assessment, it is clear that your manuscript requires a substantial revision before it can be considered further for publication in *delete*. Comments from the editors and external reviewers are included below.

We invite you to revise your manuscript within 60 days and submit it for further consideration. A delayed submission will be treated as a new submission. If you need an extension, please contact *delete* by e-mail (*delete@gmail.com*) before the deadline.

Most importantly, the revised manuscript will be subject to editorial and external reviews, and its eventual acceptance depends on the reviewers' and editors' enthusiasm. Note that manuscripts invited to be revised are accepted at a very high rate. But, manuscripts deemed to require more than one major revision are often rejected. So, it is critical that you revise it to satisfy all editorial and reviewer concerns. A re-review by original and new reviewers may raise additional concerns so anticipate them in advance and revise thoroughly.

### Example response to the editor and reviewers.

Dear Sir or Madam:

We submit our revision of the manuscript, "*deleted*" for consideration for publication in *deleted*. We have carefully considered each of the reviewer comments and the editor comments. We have modified the manuscript in accordance with these comments and created a detailed

enumeration of our responses in the following pages. We found the comments to be constructive and useful, and made modifications to the manuscript in affirmative response to them.

We thank you for your consideration.

## Introduction to Responses to Reviewers

Save all editor and reviewer comments to file called “reviewer\_comments.docx”. Never edit or change this file. Copy the file that to another file named rxr\_v1\_initials.docx. Edit the rxr file. Split up the reviewer comments so that each specific comment/question/criticism is a discreet paragraph with a heading; “Reviewer # Comment #”. Partition the comments finely.; no paragraph should contain comments on multiple subjects. Do not delete or modify any of the reviewer comments. After each reviewer comment make and label an “Author Response” paragraph (see below). The responses should be short, concise, to the point, and affirmative. The comments should refer to figures, sections, and page numbers where the manuscript and supplementary materials have been revised. Make it easy for reviewers and editors to compare the revised manuscript with the response to reviewers.

### Response to Reviewer 1

Reviewer 1 Comment 1: In their manuscript, *deleted* et al defined the universal structural core of ribosomal RNA pertinent to the modern living organisms. For this purpose, they combined information on the 3D structure of the ribosomes with sequences of rRNA and their secondary structures from phylogenetically diverse organisms. The statistical framework used by the authors is based on a solid theoretical background of information theory. It is clear that the authors have intimate knowledge of the ribosome structure and relevant literature. This part of the work is very meticulous and sound, I appreciate the hard work and congratulate the authors on this achievement.

**Author Response: We thank the reviewer for the kind comments.**

Reviewer 1 Comment 2: Further, the authors used this information to make inferences regarding the evolution of the ribosome, which are described in the last sections of the manuscript and summarized in Figure 5. This is the part of the manuscript with which I have most of my problems. While I find many of the authors’ findings very interesting, I could not agree with the way they are presented and discussed.

**Author Response: We agree that there were deficiencies in the presentation and in our method of analysis. We have modified the manuscript to accommodate the criticisms of the reviewer, as described in the sections below.**

Reviewer 1 Comment 3: The authors present a model of evolution in which rRNA is getting expanded over time. In Figure 5, axis y corresponds to the length of rRNAs and axis x to the time. The immediate issue is that what is shown in the figure is the evolution of ribosomal rRNA within a single lineage that led to humans. If that was the authors' goal, it should be clearly stated and described – the work relates only to the evolution of a single lineage. If the goal was to make inferences about rRNA evolution in general than it should address the following questions.

- (a) Does rRNA grow in prokaryotes?
- (b) Does rRNA grow in plants, fungi, protozoa?
- (c) Do the rates differ?

At present, we observe a wide distribution of rRNAs of various lengths, while Figure 5 gives a single value per time point and gives an impression that there is a monotonous growth in all lineages. This is simply not true.

We agree with the reviewer that this section of our manuscript was incomplete and poorly explained. Our original manuscript did not provide a reasonable explanation of our model, our assumptions, our data and the uncertainty in our model. The reviewer's comments here, and in the sections below, are very insightful have impelled us to improve our narrative, and our approach to analyzing this data. We have revised the results section on the size-evolution of the ribosome. We substantially changed and improved Figure 5, which now specifies lineages in a correct way.

Reviewer 1 Comment 4: The model of rRNA growth is plausible and is supported by the empirical evidence. Yet, this is only one of many possible scenarios and the authors' data do not reject alternative scenarios. Thus, I think the authors need to present their speculations in the manner that would allow some room for alternatives. For example, we could not exclude a possibility that the primordial ribosomes were larger and they went through the reduction of rRNA size at some points in the course of the evolution.

Author Response: We have incorporated these reviewer suggestions, as indicated in our response to Reviewer 1 Comment 3, above. In the section on Size Evolution of the Ribosome, we discuss the limits of our model and ascribe sources of error. We note the possibility that ribosomes got larger and were subsequently reduced.

Reviewer 1 Comment 5: What was before the ribosome? The authors seem to equate the ribosome to the Life itself. Perhaps it indeed could be equated to the modern Life (but so it is DNA since it is the genome material used by all cellular organisms, yet we have reasons to believe that DNA is a relatively novel invention). Unless we propose spontaneous emergence of ribosomes, they had to evolve somehow from something else. Did small and large subunit emerged simultaneously? A scenario was suggested where the large subunit evolved from a non-templated peptidyl transferase ribozyme, while the small subunit later provided "decoding help" (Baranov et al PMID: 19479032). Irrespective of whether such scenario reflects the reality, it has been shown that 70S ribosomes could be made from a single rRNA without causing significant phenotypical alterations (Orelle et al PMID: 26222032). To me, this suggests that the reason of two subunits existence is indeed most likely historical/evolutionary and not biochemical. If so, what are the chances that the two subunits emerged at the same time as Fig. 5 suggests? I think that the left part of Figure 5 is very hard to deduce and it would be better to design it so that the high degree of uncertainty would be clear. If the authors want to suggest a specific scenario they should be clear about the possibility of many others.

Author Response: The questions posed here by the reviewer are interesting and timely and are indeed related to the work described in the present manuscript. In part, is our interest in the origins of translation drove us to attempt to quantitatively define the common core. We have published a series of papers on the origin of translation. These papers attempt to address exactly some of the questions posed above by the reviewer. For example:

1. Petrov AS, Gulen B, Norris AM, Kovacs NA, Bernier CR, Lanier KA, Fox GE, Harvey SC, Wartell RM, Hud NV, & Williams LD (2015) History of the Ribosome and the Origin of Translation. *Proc Natl Acad Sci USA* 112:15396–15401.
2. Petrov AS, Bernier CR, Hsiao C, Norris AM, Kovacs NA, Waterbury CC, Stepanov VG, Harvey SC, Fox GE, Wartell RM, Hud NV, & Williams LD (2014) Evolution of the Ribosome at Atomic Resolution. *Proc. Natl. Acad. Sci. USA* 111(28):10251-10256.

Reviewer 1 Comment 6: Also, the model presented relies on the existence of LUCA. LUCA is a very powerful and useful abstract construction but it may never have existed in the reality. Horizontal gene transfer plays a major role in the evolution of modern prokaryotes and it is very likely that genetic information was exchanged even more freely in the past. Thus, we could not exclude a possibility of multiple life origins on earth that followed by series of divergent and convergent events with no LUCA. We really do not know. I am not suggesting that the authors shouldn't use LUCA, but it would be helpful to give it some hypothetical qualifier at some point in the beginning of the manuscript and in the discussion.

Author Response: We have incorporated this reviewer suggestion. In the section of the Discussion with the subtitle Ribosomal Evolution, we describe limitations and possibilities of LUCA, in accordance with the suggestion above.

Reviewer 1 Comment 7: Speaking of convergent events, the common view on the origin of eukaryotes is that they emerged as a result of the endosymbiotic relationship between archaea and eubacteria. The genome of the former became a nuclear genome and the genome of the latter became the mitochondrial genome. Plants underwent a secondary endosymbiont capture and the situation in some protists is even more complicated. The authors refer to the ribosome encoded in the nuclear genome as THE eukaryotic ribosome. But why? Is mitochondrial ribosome not eukaryotic? Is it prokaryotic then? I guess we couldn't say that as they differ substantially. If we look at the evolution of both eukaryotic ribosomes in the lineage that led to humans than we see opposite trends: rRNA of nuclear ribosomes expanded, while rRNA of mitochondrial ribosomes was shortened and replaced with proteins.

Author Response: This reviewer comment indicates a certain sloppiness in our narrative. The present work deals exclusively with cytoplasmic ribosomes, not mitochondrial or chloroplast ribosomes. Organellar ribosomes are complicated in that the ultimate ancestry of organellar rRNA is bacterial, and while many organellar rProteins are derived from the nucleus. We have clarified in several places that our focus here is cytoplasmic ribosomes.

Reviewer 1 Comment 8: These are my major critical points, I think if the authors address them in their revision and present their specific evolutionary model/scenario in a more balanced and less resolute way, leaving room for other possibilities while being more clear in their terms regarding evolution, eukaryotic ribosomes, etc., this could make it a remarkable paper.

Author Response: We thank the reviewer for the insightful and constructive comments, that have helped us improve our manuscript in a fundamental way.

Reviewer 1 Comment 9: (Minor comments) The Dataset S1 is mentioned in the manuscript, but I did not find it among the manuscript files. It would be great if the authors provide the sequence alignment for the core structure.

Author Response: That was our mistake. We will ensure the MSA is included in the supplementary materials.

Reviewer 1 Comment 10: (Minor comments) In the introduction, rRNA is spelled as Rrna.

Author Response: Fixed.

Reviewer 1 Comment 11: (Minor comments) "Human ribosomes are about 240 nucleotides longer than ribosomes of the last" – should refer to rRNA, not ribosomes. And again, there are two very different ribosomes in humans (see one of my above points).

Author Response: Fixed

## Response to Reviewer 2

Reviewer 2, Comment 1: In this work, the authors propose a novel metric for interpreting nucleotide sequence data within the framework of rRNA secondary structures. Building on their previous investigations into the evolution of the ribosome, the motivation of this work is to find a sequence-based metric for conservation of ribosomal structure that accurately reflects the biological context of sites within the multiple sequence alignment of rRNA genes. The authors develop a PACE scoring metric, that directly measures sequence diversity on a higher, qualitatively distinct level than individual sites, taking into account their base-pairing in the secondary structure of the rRNA molecule. The authors present a carefully curated database of rRNA sequences including this structural site-context information, as well as demonstrate that the PASE metric effectively improves the

structural mapping of the conserved ribosomal "core". They additionally develop a sequence-blind structural mapping statistic with which to further extend this analysis, and better evaluate the MSA. As I understand it, this information is more robust to mutational saturation than individual sequence identities, as these may be selectively neutral, with only the base-pairing itself being under selection. The manuscript is well-written and organized. The arguments are clear, and the methodologies are apparently sound. The authors do a good job of establishing the significance of the ribosome as a core functionality of the earliest life, and of the principles behind reconstructing the earliest history of this system with a comparative genetics approach. The database they generate will be especially useful for future evolutionary approaches, such as phylogenetic reconstruction and ancestral sequence reconstruction.

**Author Response:** We thank the reviewer for the encouragement. We have incorporated this sentiment into the manuscript and in the revised manuscript have done a better job of explaining the significance of our work.

I have only the following relatively minor comments and suggestions:

Reviewer 2, Comment 1: The authors make the general observation that the distances generated from the PACE scoring seem to follow the general evolutionary history of rRNA, that is, that bacterial and archaeal sequences are the two groups of the closest evolutionary distance, followed by Archaea and Eukarya. A very natural extension of this work would be to actually generate and present a phylogeny based on this new distance metric. For example, generating a phylogenetic tree for the selected rRNA sequences in the database using evolutionary distances calculated from nucleotide S-matrices, vs, evolutionary distances calculated from a distance matrix generated from the PACE scoring. While not necessary to the validity or the reasoning of the paper as presented, this additional analysis would greatly add to its impact and significance, in my opinion, and seems a very natural immediate extension of the work.

**Author Response:** This is an excellent idea which we are pursuing. We have made several preliminary phylogenetic tree reconstructions using the PACE scores. The results suggest that the variation in the more conserved single stranded regions provides information on deeper branching events, while variation in the faster evolving double helical regions tend to contribute to the differentiation and separation of the species-specific branches. We believe we can systematically exploit this information for improved phylogenetic tree reconstructions. However, it is complicated and will take a careful and prolonged effort to properly understand and balance the two types of signals. Therefore, this analysis will be described in a subsequent publication.

Reviewer 2, Comment 1 (Page 4, line 44): The authors discuss unpaired nucleotides and paired nucleotides in terms of "units of structure". Are these also "units of selection"? If so, some argument should be made here more explicitly about the evolutionary significance of these units in terms of selection vs. neutral processes, how therefore these are a more reasonable evolutionary unit from which to devise a metric of conservation.

**Author Response:** Yes, we do believe base pairs are units of selection in helices. We have made note of this in the discussion, and cited another paper where this possibility is described.

Reviewer 2, Comment 1: (page 5, line 18) Rrna should be changed to rRNA.

**Author Response:** Fixed

Reviewer 2, Comment 1: (page 14 line 8) It would be interesting to show some higher-resolution information from Table 3, for example, a per-site scatterplot of PASE vs. Shannon entropy for paired positions, in order to show how robust the PASE and Shannon indices are, across different levels of conservation.

**Author Response:** This is another good idea. However, rather than representing the data as a scatter plot we have taken the difference and mapped it onto the secondary structure. This has been incorporated into the manuscript as Figure S8.

Reviewer 2, Comment 1: (Figure 5) This is a very interesting figure, and clearly shows the trend of increasing rRNA length across the lineages mapped. However, the figure may be slightly misleading to some readers. Effectively, the authors have drawn a cladogram, with a selection of nested clades labeled along the time axis, and their representative rRNA lengths mapped. So, it is not that the ribosome is increasing over time (is it increasing over time within Bacteria or Archaea? I don't think so?) Its that more highly derived complex metazoans show a trend of increasing the size of their ribosomes. This is a subtle but important difference that should be clarified and addressed in the text.

Author Response: The reviewer is correct. Reviewer 1 identified the same problem. We have modified and clarified Figure 5. It now specifies sizes within specific lineages. This comment was very helpful to us.

### Example editor response to revision.

Following is the best kind of response to a revision:

Editor Comments to the Author:  
Dear Dr. Williams and colleagues,

Many thanks for your thorough revision. I am satisfied that the reviewers comments have been adequately addressed. I am recommending that the manuscript is accepted for publication.

Best wishes,  
*deleted*

### Thesis attribution statement.

Every PhD student in the Williams lab should put a statement of attribution of work at the beginning of *every* chapter of their thesis. If work is reproduced from a manuscript, all the authors of the manuscript should ok the attribution statement. The following is an example attribution statement.

Portions of this chapter are adapted from previously published work: *give the full citation to the previous work*. The author of this document contributed to the work in this chapter by conducting SHAPE footprinting experiments in Figure 1, SHAPE data analysis in Figure 2, and Creation/design of Figure 4. The author of this document wrote portions of the initial draft, helped edit and organize the manuscript, and wrote the initial draft of the response to reviewers.

## Elements of style.

### Use Oxford Commas

RNA, DNA, and protein came down in the pellet.

### Use clear declarative topic sentences.

Here are some example topic sentences.

Ions such as  $\text{Na}^+$ ,  $\text{K}^+$  and  $\text{Mg}^{2+}$  are required for folding of RNAs into compact structures and for conferring functionality to RNAs.

Cytoplasmic ribosomes contain a 'common core' [4] consisting of rRNA and rProteins with conserved structures in essentially all extant species (Figure 1).

The coordination and geometry of di-magnesium centers of RNA show certain similarities to those of di-iron and di-manganese centers within catalytic sites of proteins (Figure 1).

Here is a terrible topic sentence.

Figure 1 shows the hydrogen bonding of ribosomal protein L4 with rRNA in the exit tunnel.

### Do not use "this" as the subject of a sentence.

"This" must always be followed by a noun.

No: This is a fast reaction.

Yes: This reaction is fast.

### Avoid "This is..." "There exists..."

No: This is...

No: There is...

No: There are...

No: That is...

No: There exists...

### Never use the word "ones".

No: In the kinetic realm only a few species (the fast-reacting ones) react with the compressor.

Yes: In the kinetic realm the few fast-reacting species react with the compressor.

### Never use the word "respectively".

No: Carbon and nitrogen are green and blue, respectively.

Yes: Carbon is green and nitrogen is blue.



### Use the active voice.

No: It was observed that the RNA degraded.

Yes: The RNA degraded.

Yes: We observed the RNA degrade.

### Complete all comparisons.

"Greater" requires "than". "More" requires "than". "Less" requires "than".

No: Sodium chloride is more soluble.

Yes: Sodium chloride is more soluble than barium chloride.

Do not use "less" when you mean "fewer"

### Use "which" and "that" correctly.

*That is a defining and restrictive pronoun. Which is a nondefining and nonrestrictive pronoun.*

The lawn mower that is out of gas is in the garage.

“that is out of gas” tells the reader which lawn mower is under consideration. It *restricts* the sense of the sentence to the lawn mower in the garage. This sentence implies that there are additional lawn mowers that are not in the garage and that are not out of gas. A reasonable next sentence might state, “Please get the other lawn mower, from the basement, and mow the lawn.”

The lawn mower, which is out of gas, is in the garage.

“which is out of gas” adds parenthetical information about the lawn mower in the garage, and should be set off by commas. This information could be removed without changing the basic intent of the sentence. A reasonable next sentence might state, “Please mow the lawn, but fill the tank first.”

The RNA that is cleaved entered the gel.

“that is cleaved” tells the reader which RNA is under consideration. It restricts the RNA being discussed to the RNA that has entered the gel. This sentence implies that there is additional RNA that is uncleaved and did not enter the gel. A reasonable next sentence might state, “In contrast, intact RNA stayed in the wells.”

The RNA, which is cleaved, entered the gel.

“which is cleaved” adds parenthetical information about all the RNA under consideration. All the RNA here is cleaved and all of it entered the gel. A reasonable next sentence might state, “This experiment confirms that RNA is not stable to high magnesium concentrations.”

### **Additional examples of which v that**

The enzyme cleaves RNA, which has a 2' hydroxyl group.

This sentence means that all RNA is cleaved by the enzyme and, parenthetically, that RNA has a 2' hydroxyl group.

The 'which...' phrase adds information but is not essential to the meaning of the sentence.

Note the comma before "which".

The enzyme cleaves RNA that is chemically modified.

The enzyme only cleaves RNA that is chemically modified. The RNA that is cleaved is restricted, to chemically modified RNA. There must be RNA that is not chemically modified and might not be cleaved.

The 'that...' phrase is essential to the meaning of sentence.

Note the lack of comma before "that".

My bike, which was locked, has been stolen.

Maybe I have another bike that was not stolen, maybe not.

My bike that was locked has been stolen.

Only my bike that was locked was stolen. I must have another bike that was not stolen.

Examples of correct usage of 'which'.

These GC-rich rRNA elements are now seen to contain G-quadruplexes (Figure 2), which are stacked tetrads of guanine bases.

As positive controls for our computations and experiments we used the ADAM10 G-quadruplex, which gives a G-score of 42.

### **Do not confuse kinetics and thermodynamics**

The word stable should be restricted to the thermodynamic sense and should not be used in reference to kinetic phenomena.

Never use the phrase "kinetic stability". It is an oxymoron.

No: Hydrogen and oxygen do not react because they are kinetically stable.

Yes: Hydrogen and oxygen do not react because they are in a kinetic trap.

Yes: At one atmosphere total pressure, hydrogen and oxygen are persistent at room temperature, but are unstable.

Yes: Water does not react because it is stable.

No: DNA is more stable than RNA.

Yes: DNA is in a deeper kinetic trap than RNA.

Yes: DNA is more persistent than RNA.

### **Use parallel construction.**

No: This protein is soluble, stable and is a monomer.

Yes: This protein is soluble, stable and monomeric.

### Do not tell the reader what is interesting or significant or intriguing important.

No: It is significant that RNA can catalyze chemical reactions and maintain genetic information.

Yes: RNA can catalyze chemical reactions and maintain genetic information.

No: Intriguingly, RNA can catalyze phosphoryl transfer,  
Show, don't tell. Make your writing interesting. Do not command the reader to be interested.

### Use consistent nomenclature and syntax throughout a document.

Technical writing, unlike creative writing, must be consistent, simple and direct. Every paper has a well-defined internal glossary.

### Use signal words.

contradiction is signaled by: in contrast, but, however, nevertheless, although.

cause/effect is signaled by: as a result, consequently.

reason/conclusion is signaled by: therefore, because, hence.

addition is signaled by: additionally, moreover, furthermore.

### Do not use contractions.

Use "It is" not "It's".

### Use strong topic sentences.

Never start a paragraph with, "Figure 3 shows blah blah blah."

### Tenses (adapted from Nature, Effective Writing)

Use past, present, and future verb tenses as you do in ordinary writing and speaking. In a technical paper, most of sentences are in the past tense, some are in the present tense, and very few are in the future tense.

The past tense reports the past: what you did, what someone reported, and what happened in an experiment.

The present tense expresses general truths, such as your conclusions and atemporal information (about what the paper does or covers).

The future tense communicates perspectives: what you will do in the coming months or years.

#### Past tense

##### Work done

Woese and Fox **sketched** out a universal tree of life.

We **collected** NMR spectra from . . .

Yonath **determined** the structure of the ribosome.

### **Work reported**

Petrov **reported** a similar growth rate . . .

Hud **observed** fast reactions . . .

### **Observations**

The mice in Group A **developed**, on average, twice as much . . .

The number of defects **increased** sharply . . .

The conversion rate **was** close to 95% . . .

## **Present tense**

### **General truths**

The Woese and Fox tree of life **contains** three primary branches.

The universality of translation **extends** beyond sequence homology to three-dimensional structure.

The ribosome **reads** mRNA and **catalyzes** peptidyl transfer.

### **Atemporal facts**

Section 4 **explains** the difference between . . .

Franklin's paper **provides** a framework for . . .

This paper **presents** the results of . . .

## **Future tense**

### **Perspectives**

In later experiments, we **will study** the role of . . .

The influence of water activity **will be** the object of future research . . .

## **Analyze and Improve your writing.**

(a) Write the first draft of a paragraph in logical order. Then take the last sentence, which is the conclusion, and move it to the beginning of the paragraph, to convert it to the topic sentence.

(b) Copy and paste the first sentence of each paragraph of the introduction (or any other section) to form a single artificial paragraph. This new paragraph should read cleanly and in a logical progression.

(c) Convert the paragraphs of a given section into different text colors. Try re-ordering the sentences and paragraphs. Experiment.

(d) Break up a paragraph by inserting carriage returns after each sentence. Read each sentence independently of the others. Try re-ordering the sentences.

(e) Read your text out loud.

(f) Have your computer read your text out loud to you.

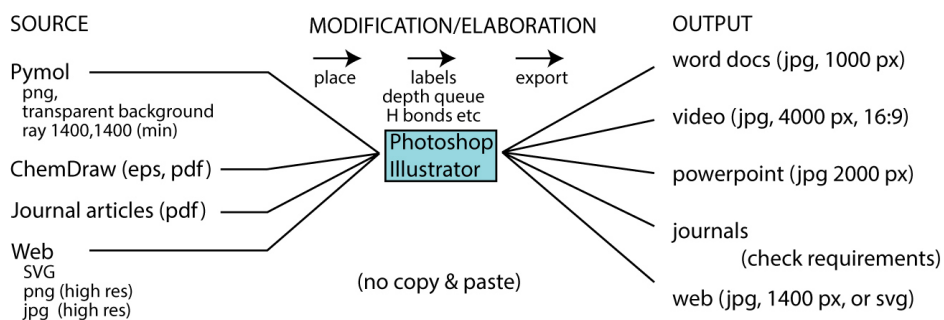
(g) Use an online dictionary and thesaurus as you write.

## Figures:

### Adobe Illustrator

We have a site license for Adobe Creative Suite. The learning curve is a bit steep, but use it. All final figures are made in Illustrator, which are converted to high res jpg or png prior to submission. Do not make figures with Powerpoint.

### Work Flow



## Graphical File Types and Uses

### VECTOR

In a vector file, every dot, line and object is represented by an equation. Vector files are editable by Illustrator or Inkscape. Illustrator is expensive with a steep learning curve, but is widely used, powerful and exact. We have a site license for the adobe suite, which includes Illustrator. Inkscape is a free and open source vector editor that works pretty well.

### AI (Illustrator)

The Illustrator file type is raw and fully editable, containing original lines and elements in a modifiable format. Imported raster elements are not editable in Illustrator. Illustrator files contain layers. Illustrator files are proprietary and are not portable; you cannot open an ai file without Illustrator. Essentially any other file type can be easily exported from Illustrator.

### PDF

Portable document files can be read, opened and edited by Illustrator and can be converted to Illustrator files (sometimes they are password and you have to break into them, check with Loren on how to do this). This portable document format is readily transferable and can be opened by web browsers.

## EPS

Encapsulated postscript files are the predecessor of PDF and can be opened and edited by Illustrator. Chemdraw files can be saved as eps and imported as vector objects into Illustrator.

## SVG

Scalable Vector Graphics files are an open source vector format. Wikipedia offers svg files. The svg file is raw and fully editable, with layers, containing all original lines and elements in a modifiable format. If you want to make vector format files available to others, use svg. Inkscape reads and writes svg format.

## PIXILATED

(A raster image is a grid of dots called pixels where each pixel is assigned a color)

## PSD

Photoshop files are the only pixelated format with layers, and are used for photographic-quality images.

## JPG

Joint Photographic Experts Group files are flat (no layers) and have efficient and easily controllable compression. JPGs do not support transparent backgrounds (always have colored or white background). Use JPG for inserting images into word documents (large images cause problems in word).

## PNG

Portable Network Graphics are flat lossless portable files (uncompressed) that support transparent backgrounds. Use PNGs for submitting images to journals.

## GIF

Graphics Interchange Format files support transparent backgrounds, are lossless and support basic animations. GIFs can be exported with customized settings that reduce the amount of color and image information, reducing the file size.

## TIFF

Tagged Image File Format are very high quality lossless files with deep bit-channels. Use TIFF for data (gels and images). They are large and slow but necessary for data. TIFFs can be scanned (the intensity of bands can be accurately integrated).

## Colors

CMYK stands for **c**yan, **m**agenta, **y**ellow and **k**ey (black). Files in this format are optimized for physical printing.

RGB stands for **r**ed, **g**reen and **b**lue. Files in this format are be optimized for the web, or anything on a screen.



## Dropbox. Organization of Manuscript Folders.

(share with ldw@gatech.edu)

Organize your dropbox **directory** in the following way:

### **\_2020\_yrname\_topic**

#### **manuscript**

manuscript\_v12\_ldw.docx (*increment the version number and substitute your initials when you edit*)

**z\_archive** (*keep the directory clean and neat. Put previous versions here. Do not delete anything.*)

#### **figures**

##### **illustrator**

figure\_1\_CD\_spectra.ai (*use the figure number and short descriptor*)

figure\_2\_EMSA.ai

**z\_archive**

##### **jpg**

figure\_1\_CD\_spectra.jpg (*horizontal res = 2000 horizontal dots, no compression*)

figure\_2\_EMSA.jpg

**z\_archive** (*keep the directory clean and neat*)

**tiff** (*if necessary*)

**pymol\_scripts\_etc** (*source files, scripts, pdb files*)

#### **miscellaneous**

##### **submission\_documents**

letter of submission\_v3\_asp.docx

##### **supplementary\_materials**

supp\_mat\_v3\_asp.docx

##### **illustrator\_sm**

figure\_1s\_MSA.ai

figure\_2s\_sec\_struct.ai

**z\_archive**

##### **jpg\_sm**

figure\_1s\_MSA.jpg

figure\_2s\_sec\_struct.jpg

**z\_archive**

##### **endnote\_library**

yourname\_ **topic**.enl

yourname\_ **subject**.Data

##### **reviews\_papers**

## Credits for this document.

This information was pulled together over a long period of time and is based on guidance from my graduate and postdoctoral advisors Barbara Ramsay Shaw and Alexander Rich, and from my coauthors and collaborators including Nicholas Hud, Jennifer Glass and Jessica Bowman, from the real *Elements of Style* (Strunk & White), *Writing a Paper* by George Whitesides, and possibly some other sources that I do not remember.