The surprisingly complex tRNA subsystem: part 1—generation and maturation

Royal Truman

tRNAs serve as linker molecules to translate mRNA sequences into protein sequences, but these require extensive processing by preexisting protein-based molecular machines. 5′ leader sequences and 3′ extensions must be removed using complex ribonucleases having virtually no similarity across prokaryotes, mitochondria, and eukaryotes. In plants, RNase P is a protein-only enzyme. Introns must also be removed but the location of the splice junctions differ for eukaryotes and archaea.

Aminoacylation occurs on the trinucleotide motif '*CCA*' as a free RNA end on all tRNAs, a motif necessary for processing in ribosomes. tRNA genes encode the CCA sequence in *E. coli* and related bacteria, Gram-positive *Bacillus subtilis,* and some related bacteria. However, in nearly all eukaryotes and archaea the CCA sequence is not encoded and must be added post-transcriptionally.

All cells examined also encode an enzyme to repair the CCA sequence if damaged. The enzyme can discriminate against damaged tRNAs, causing defective ones to be degraded. A second quality control process found in all three kingdoms of life involves adding a second CCA to defective tRNAs as a degradation tag.

These observations support the view that cells could not have arisen without an overall plan.

The cellular genetic system is irreducibly complex, being composed of several collaborating sub-systems based on DNA, RNA, and proteins. Remarkably, the individual subsystems themselves, like mRNA (messenger RNA), DNA, ribosomes, and so on cannot exist independently without one or more of the others (figure 1). There is no simple evolutionary starting point. One could reason that since all organisms depend on DNA, perhaps this polymer was present first.

DNA, however, consists of nucleotides manufactured from pre-existing enzymes (protein-based molecular machines), and a new DNA strand requires a former one to already exist to serve as a sequence template. DNA replication requires many additional molecular machines such as polymerases, helicases, and topoisomerases. All these molecular machines are only produced by translating pre-existing mRNA (which comes from DNA) on pre-existing ribosomes. But the RNA and protein sequences used by ribosomes are themselves coded for on DNA.

Ever more circular dependencies are identified the deeper one looks. Transcription of portions of DNA to produce mRNA requires that RNA polymerases already exist, and translation only works if suitable tRNAs are already available (derived from DNA) and charged by aaRS (aminoacyl tRNA synthetases) which are additional molecular machines. This charging is not accurate enough without proofreading enzymes (yet more molecular machines). And none of the systems mentioned work without energy-supplying ATP or GTP molecules (created by complex machines requiring dozens of proteins). Everything requires something else.

It seems natural to conclude that an initial top-down design was necessary and then everything set up correctly. But this common-sense thinking contradicts the dominant naturalist paradigm. No form of intelligent planning nor contribution can be considered. But the deeper the subsystems themselves are examined the more dependencies are discovered. This is a dilemma. Evolution cannot work without enough minimal functionality to ensure survivability.

Documenting all the additional interdependencies in the genetic system seemed overwhelming, so we'll focus on the simplest subsystem for now: merely producing a collection of suitable tRNAs to serve as adaptor molecules needed to translate codons to the intended amino acid at ribosomes. Genetic code translation cannot occur without tRNAs, so they must have been present initially. To simplify, we will not consider the need for aminoacyl-tRNA synthetases, nor the proteins tRNAs need to interact with at ribosomes. These are separate subsystems which will require additional series of papers. We will even neglect details such as how Stop codons are recognized without requiring their own tRNAs.[1,2]

Parts1 and 2[3] of this series cover processing of tRNAs, and chemical modifications for generating functional tRNAs, respectively. Part 3[4] covers additional quality control measures, and part 4[5] how customized RNA fragments derived from tRNAs are used for signalling and regulatory purposes. Finally, in part 5[6] we examine mathematically why there can be no simple starting point for the tRNA portion of the genetic code, and why evolution is incapable of developing stepwise the numerous complex processes that utilize tRNAs.

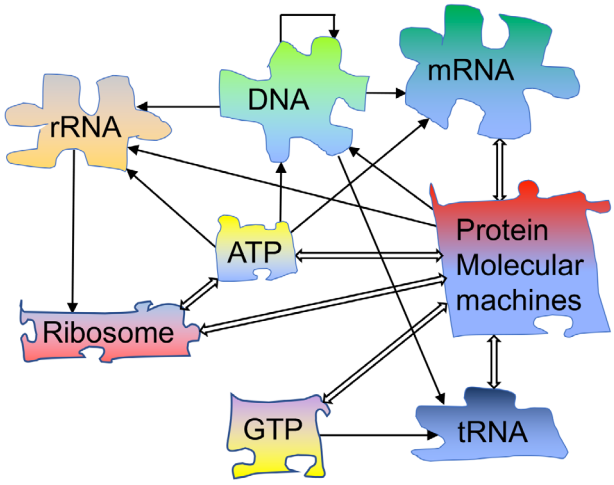

Figure 1. The cellular genetic system is composed of several mutually dependent subsystems. Single and double arrows represent an immediate, direct dependency for a particular subsystem to function. Dozens of protein-based molecular machines are indispensable such as DNA and RNA polymerases; topoisomerases; helicases (for nucleotide, amino acid and cell membrane synthesis); transcription factors; error correcting enzymes; aminoacyl tRNA synthetases, etc.

Overview and general information on tRNAs

The single-cell eukaryote yeast contains approximately 3 million tRNA molecules and the bacterium *E. coli* about 200,000.[7] tRNA is present in tens of millions of copies in a human cell and is the most abundant kind of RNA.[8] tRNA genes are found on mitochondrial and nuclear chromosomes, and nuclear-originating tRNA transcripts can be imported into mitochondria.[8] High tRNA gene sequence diversity is found in all mammalian genomes. Even the simplest eukaryote *Saccharomyces cerevisiae* has about 272 tRNA genes comprising 53 unique DNA sequences.[8,9]

Abundance, chemical modification, and charging levels (covalent attachment of amino acid to a tRNA in a high-energy state) are regulated to ensure proper decoding of mRNA in different cell types and cell states in environment-dependent ways. Translation is most active during the cell cycle when the cell doubles its protein mass before fission. During this time, the amount of tRNA increases, the relative proportion of *isoacceptors*[10] (tRNAs charged with the same amino acid but having different anticodons) change, most tRNAs get charged with activated amino acid, and chemical modification levels for most tRNAs are high.

When cells are stressed or cycle progression is halted, the major protein synthetic activity switches to stress response and other regulatory proteins.[8]

Spatial and chemical features of tRNAs must serve as informational signals to guide interactions with various components related to the translation apparatus. As an example, dozens to hundreds of different tRNA molecules must differ sufficiently in a cell so that the correct one can be reliably aminoacylated by synthetases with the right amino acid.[9] This includes the tRNAs having different primary structures which must recognize synonymous codons (i.e. representing the same amino acid).[11]

In spite of considerable sequence variation, all cellular tRNAs must fold into almost identical three-dimensional structures to fit the relatively narrow tRNA binding sites of the ribosome during translation elongation. The folded cloverleaf secondary structure (figure 2) of every kind of tRNA requires Watson–Crick or wobble pairs in the four helical stems to produce a consistent structure. There are always either four or five inverted repeats that are responsible for formation of the stem-loop.[12] The structure must also optimize the specificity of codon-anticodon interactions between mRNA and tRNA.[11] Therefore, many chemical modifications are necessary to change the biophysical properties of the individual tRNAs before they can be used,[13] as we will discuss in part 2.

Genes encoding tRNAs for some anticodons appear in dozens of copies in individual organisms, while others are universally absent. This variability reflects the fact that mature tRNAs and degradation fragments derived from them regulate many cellular processes including protein expression level, translation accuracy, adequacy of protein folding, and even mRNA stability,[14] details we will discuss here and in part 2[3] and part 3[4] of this series.

Obtaining tRNAs from DNA is not a trivial matter and informational patterns are involved to demarcate the genes. Computer algorithms have been designed to identify tRNA genes using features such as the presence of highly conserved residues located at predictable distances from each other. This revealed at least 610 tRNA genes in the reference human genome.[8] Of these, 423 appear able to produce a stable tRNA structure, having 264 different DNA sequences.[8] The eukaryote tRNA genes are also preferentially localized in the nucleolus along with RNase P (discussed below), to facilitate mass production,[9] whereby RNase P seems to associate directly with the Pol III transcription apparatus through interactions with the transcription factor TFIIIB, which binds upstream of tRNA genes.[9]

I. tRNA gene variety

Scan-SE is a widely used bioinformatic tool to predict tRNA genes, based on predicted clover leaf secondary structure produced, sequence similarities across organisms, and various diagnostic intragenic tRNA promoter sequences.[8,15] Tools like these can produce false positives, but nevertheless as of 2014 approximately 600,000 tRNA genes were registered on the tRNA gene database.[16] The number of tRNA genes vary considerably across organisms, with the *Plasmodium falciparum*[17] parasite containing the lowest number of tRNA genes known for a eukaryotic cell (a total

of 46 nuclear genes encoding 45 tRNA isoacceptors). Two different genes encode the initiator and elongator $tRNA^{Met}$.[18]

Isodecoders (tRNAs with different body sequences but the same anticodon).[19] Isodecoder expression varies widely in human tissues and is similar across mammalian genomes. The many isodecoders must carefully avoid being mischarged by the 19 other tRNA synthetases and therefore include hindering features for quality control purposes.[8] That so many similar isodecoders are present across many kinds of organism implies their dedicated use for comparable biological purposes.

Each isodecoder tRNA set is generated from a pool of tRNA genes. For example, 35 genes code for $tRNA^{Ala}$(AGC) in the human genome.[20] Isodecoders have differential, tissue-specific functions. The five $tRNA^{Arg}$(UCU) genes recognize the AGA codon of arginine. One of them (Tr-20) is very different in both sequences and the presence of introns. Tr-20 is primarily expressed in the central nervous system where it accounts for ~60% of all $tRNA^{Arg}$(UCU) expression. One mutation which significantly reduces the expression of Tr-20 isodecoder increases ribosome pausing specifically at AGA codons and results in neurodegeneration, likely due to increased protein misfolding or degradation.[8]

Isoacceptors.[10] Since different codons can represent the same amino acid (i.e. synonymous codons) it is not necessary to have tRNAs with all $4 \times 4 \times 4 = 64$ possible trinucleotide anticodon patterns if the system is properly designed. And indeed, in humans, tRNA isoacceptors are used to translate 12 amino acids: two for Glu, Lys, Gln; three for Ile, Val, Thr, Ala, Gly, Pro; four for Ser; and, five for Leu, Arg. For the six amino acids that are represented by only two codons each (Phe, Tyr, His, Asn, Asp, Cys) a single tRNA is present. In these cases, the wobble anticodon nucleotide (with the exception of $tRNA^{Cys}$) can be modified to more efficiently identify and decode both codons.[8] It is important not to overlook the engineering challenge to ensure that the correct logic is implemented. A particular tRNA anticodon must correctly identify either one and only one mRNA codon, or two, three or four specific ones. Misinterpretation would lead to incorrect protein sequences.

The logic and implementation of synonymous codons is taken up in more detail in part 2. Relying on fewer tRNAs saves the cell energy and material and decreases risk of damaging mutations (the fewer the tRNAs genes, the less which could be damaged). The concept of synonyms also minimizes the damage caused by mutations on protein-coding genes at the synonymous position, since in many cases the correct amino acid would still be identified during translation.

While the advantages just mentioned may be apparent, this requires careful upfront design and will not arise by chance nor evolutionary trial-and-error. Once the genetic code logic was selected, the tRNA 'hardware' must be designed able to implement the redundancy logic with high precision. As an example, codon patterns UU*U* and UU*C* are meant to both represent amino acid phenylalanine (nucleotides U and C are to be treated as interchangeable in the third position). But they must not be treated as equivalent in the second and first positions. U*C*U must be translated as serine and *C*UU as leucine.

The 'hardware' design must ensure that the tRNA anticodon and codon H-bond interactions work exactly as necessary in three dimensions during translation at a specific location in a ribosome. We now see why chemical modifications in tRNAs are vitally important and also foreknowledge of what the 'natural' structures would look like for various possible chemical solutions.

Incidentally, not to overlook the obvious, once the tRNA hardware specifications were designed, this must be taken into account. In our example, if amino acid leucine is to be specified, then all incorrect UU*U* or UU*C* codons must be modified before 'going live'. (Acceptable codons for leucine would be UU*A*, UU*G*, CUU, CUC, CUA, or CUG). This principle must be applied for all tRNA cases of single, double, triple, and quaternary redundancy for all protein-coding genes throughout the entire genome, to prevent translational chaos.

II. tRNA maturation processes

tRNAs are composed of 73–90 nucleotides and have a characteristic cloverleaf secondary structure made up of the D-loop, T-loop, variable loop or extra arm, and the anticodon stem-loop (ASL) (see figure 2). Coaxial stacking of the T- and D-loops helps to produce the necessary three-dimensional shape.[7,21]

In bacteria, tRNAs are usually extracted from long transcripts by nucleases. In eukaryotes, however, transcription factors recognize specific sequence patterns shared by all tRNA coding regions to mark the genes as individual transcription units to which transcription factors and Pol III bind.[9] Although little or no critical effort has been invested in challenging the notion of common ancestry for tRNA gene sequences from an evolutionary point of view, one unexpected set of observations has been difficult to overlook. This involves the unrelated ways Gln-tRNA are formed. Genome sequence and biochemical analyses provided a major surprise, revealing that different biosynthetic routes are used in each of the three kingdoms.[22]

These kinds of examples reveal that eukaryotes are based on unique designs and contradict the assumption they evolved from prokaryotes. In fact, the uniqueness of so many eukaryote genes and cellular processing with no plausible relationship to prokaryotes is common knowledge. This has led many to question the notion of common ancestry, as the recent case of Chinese Professor Change Tan who then converted from being a committed atheist and evolutionist to a passionate creation scientist.[23]

Eukaryotic cells employ separate DNA-dependent RNA polymerases (Pol I–III) to transcribe precursors such as

ribosomal RNA (pre-*rRNA*) and mRNA (pre-mRNA). Pol III is dedicated to transcribing small RNAs such as pre-tRNAs.[24,25] The primary product of transcription of a tRNA gene must then be modified to become functional.

The precursors generally contain extra nucleotides at both their 5' and 3' ends and sometimes in the middle, which will be discussed next.

Removal of 5′ leader sequences from pre-tRNAs

The 5' leader sequence is removed from tRNA precursors by ribonuclease P (RNaseP), an enzyme found in all organisms (figure 3).[21] To date, the only known exception to the requirement for RNase P activity is *Nanoarchaeum equitans*, which presumably can produce leaderless tRNA transcripts.[12] The prokaryotic RNaseP are about 10% protein. Archaeal and eukaryotic enzymes are 50–70% protein, and the proteins are larger and share no sequence similarity to each other or to the prokaryotic proteins.[21] Notice again how the data does not imply common ancestry.

Unexpectedly from an evolutionary point of view, RNase P is a protein-only enzyme (PRORP) in plants. Also remarkably, some RNase P versions in the organelles of various organisms require an RNA component and others don't.[24] Without a collection of functional tRNAs translation cannot occur but the prerequisite processing molecular machines are complex. For example, in budding yeast, the RNase P complex consists of the RNA subunit RPR1, and nine essential proteins.[24,26,27]

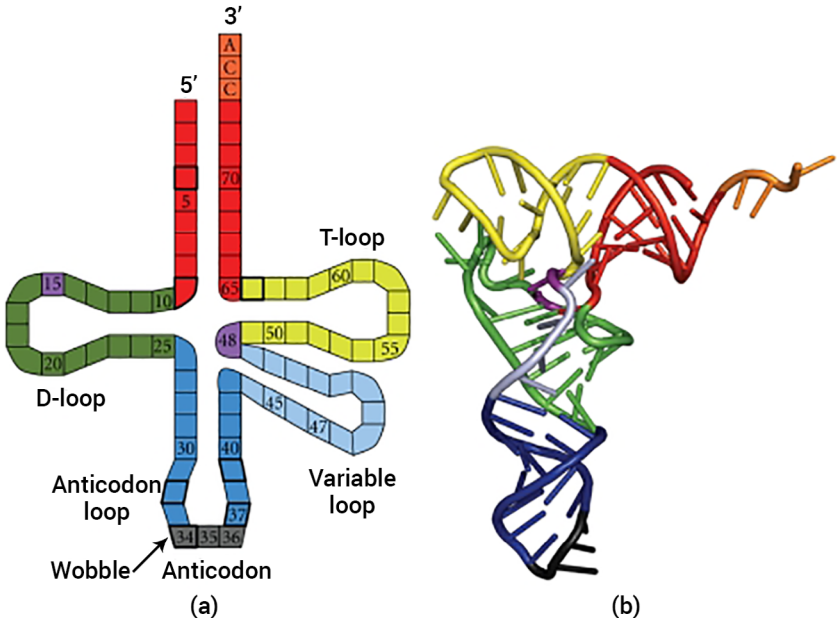

Image: ResearchGate/CC BY 3.0

Figure 2. tRNAs. (a) Schematic representation, showing the D-loop (green), anticodon loop (blue), anticodon (grey), variable loop (light blue), the T-loop (yellow), the acceptor stem (red), and the CCA aminoacyl binding site (orange). The Levitt base pair is shown in purple. (b) Tertiary structure of a yeast $tRNA^{Phe}$, rendered with PyMOL (1ehz, in the Protein Data Bank).

Almost all yeast nucleus-encoded tRNAs are transcribed as single pre-tRNAs with ~12 extra leader nucleotides on the 5' end.[25] There are three interesting facts about RNase P in yeast:[25]

1. Mitochondrial and nucleolar forms of the enzyme are encoded by unrelated genes. Nucleolar RNase P consists of nine proteins and a single essential RNA (RPR1) encoded in the nucleus. In contrast, the mitochondrial enzyme contains only a single nuclear encoded protein, Rpm2, and a single RNA (RPM1). Moreover, RPR1 and RPM1 RNA parts differ extensively in length and sequence.
2. There are extensive differences in RNase P structure when examined using an evolution-theoretic phylogenic tree. Unlike the bacterial, archaeal, and eukaryotic forms of RNase P, which are ribozymes with varying numbers of protein subunits, higher plant mitochondrial and nuclear versions are now known to be protein enzymes. This is not consistent with a possible evolutionary explanation that mitochondria are the result of endocytosis of aerobic bacteria by large anaerobic bacteria (Boston University biologist Lynn Margulis's Endosymbiotic Theory). A putative single origin of organisms having the genetic code and ubiquitous presence of a ~12 extra leader sequence would predict similar RNase P enzyme complexes throughout nature and not multiple dramatically different designs across various taxa.
3. Most of the protein subunits of the RNase P enzymes are shared with RNase MRP, which is involved in pre-rRNA processing. In other words, the same proteins are involved in producing tRNAs and ribosomes, implying the familiar theme of a concurrent origin of both systems.

For most tRNAs, RNase P generates the mature 5' end. However, generation of the $tRNA^{His}$ 5' end requires an additional step, addition of a 5' G (G_{-1}), catalyzed by Thg1. G_{-1} addition to $tRNA^{His}$ is essential for its aminoacylation.[25]

Removal of 3′ trailer sequences

In bacteria and yeast, removal of 3' extensions from pre-tRNA is a biochemically complicated process, involving both exo- and endonucleases.[28,29] In some organisms, exoribonucleases perform 3'-end tRNA processing, while in others, endonucleases do so. Complete 3' processing requires ribonuclease II, polynucleotide phosphorylase, and ribonucleases T and/or PH.[26,27]

Almost all yeast nucleus-encoded tRNAs are transcribed as single

pre-tRNAs with ~12 extra 3' trailer nucleotides, which need to be removed.[21,25] Yeast Rex1 is a 3' to 5' exonuclease that participates in the processing of pre-tRNA trailers as well as in the processing of other RNAs such as ribosomal 5S rRNA, 5.8S rRNA.[25] Ribosomes without tRNAs serve no purpose, and *vice versa*. We point out that enzymes are precisely designed to operate in both genetic subsystems while avoiding collateral damage to all the other RNA present in the cell. Surely this was so from the very beginning of cells. Or were all these enzymes somehow present before the genetic code and thus when they would be needed?

Addition of CCA 3′ end

All tRNAs require the CCA sequence at their 3'-end (figure 2). This pattern protrudes from the acceptor stem as a single-stranded motif and is recognized by the aminoacylation enzymes. Once inside the ribosome, the CCA sequence is also necessary to interact with a changing ensemble of ribosomal proteins and RNA elements as the tRNA moves through the ribosome.[30] The CCA end is also necessary for accurate recognition by the RNA component of bacterial RNase P in bacteria.[12]

Remarkably, sometimes this CCA end is encoded in the tRNA genes and other times it must be enzymatically added after the tRNA precursor has been formed, unexpected if a common ancestor had existed. In *E. coli* and related bacteria, the tRNA genes encode the CCA sequence, which is also the case of Gram-positive *Bacillus subtilis* and some related bacteria.

In other Gram-positive bacteria, nearly all eukaryote tRNA genes (both cytoplasm and mitochondria), and archaea, the CCA sequence is not encoded in tRNA genes and must be added post-transcriptionally using a complex series of reactions.[21,25, 30] This sequence is acquired and maintained by stepwise nucleotide addition using an unusual RNA polymerase that does not use a nucleic acid template for nucleotide selection.[30] Kinetic analysis implies additional protein factors assist in the release after CCA has been added.[30]

Further discrediting the notion of common ancestry, bacteria such as *Aquifex aeolicus*, *Deinococcus radiodurans*, and *Synechocystis sp*. use two class II CCA enzymes, one to add C (Cytosine) precisely at positions 74 and 75, and an A (Adenine)-adding enzyme for addition 76.[30] The two enzymes are fundamentally different in the arrangement of secondary structures of individual domains and share virtually no amino acid sequence resemblance.[30]

Quality control of the CCA sequence. Even organisms that do encode CCA in tRNA genes contain a CCA enzyme, which is responsible to repair the CCA sequence if damaged.[30] Kinetic work performed on *E. coli* revealed that for each of the three nucleotide additions, the enzyme has an innate ability to discriminate against damaged tRNAs by scrutinizing the integrity of the tRNA substrate. Flawed tRNA can arise from incorrect nucleotide sequences, incorrect processing and folding, and incomplete post-transcriptional modifications (discussed in part 2). Also, some bacterial and eukaryotic tRNAs carry introns in the anticodon loop that have not been removed.[30] This discrimination causes defective tRNAs to be rapidly degraded by RNA surveillance mechanisms found in all cells, preventing them from entering the ribosome machinery.[30] This will be further elaborated on in part 3 of this series.

Two structurally distinct classes of CCA enzymes are known and they differ in the mechanism of nucleotide selection.[30] The Iarchaeal enzymes are members of class I nucleotidyl transferase family, whereas bacterial and eukaryotic enzymes are members of class II.[30]

Structurally unstable tRNAs are also actively eliminated by CCA-adding enzymes which add a second CCA, leading to a 3'-terminal CCACCA sequence, which serves as a specific degradation tag.[31–33] This occurs in all three kingdoms of life. A study in 2015 using *Archaeoglobus fulgidus* elucidated the mechanism involved.[31] The CCA-adding enzyme monitors in particular the acceptor stem when deciding between CCA and CCACCA addition. The CCACCA addition generates a single-stranded tail that can be recognized by various exonucleases. A recent study showed that tRNAs ending in CCACCA, but not CCA, were degraded *in vitro* by RNase R from *E. coli* and by yeast Rrp44.[32]

Intron removal

tRNA genes often contain introns which must be removed.[21] The percentage of tRNA genes that contain introns varies across organisms, ranging from 0% in some bacteria like *E. coli*, to ~5% in Drosophila, *C. elegans*, mouse, and human genomes, to >50% in some archaeal genomes.[24,25] 59 of the 274 yeast nuclear tRNA genes contain an intron, ranging from 14 to 60 nucleotides.

Problematic for evolutionism, there is no similarity in the sequences at the splice junctions.[9] Yeast and vertebrate tRNA introns are always located one base 3' to the anticodon, but introns appear in other locations in tRNA genes in archaea.[25] The catalytic splicing endonuclease subunits are similar across archaea and humans, although the archaeal enzyme recognizes a helix-bulge-helix structure in pre-tRNAs but the budding yeast and human enzymes don't.[9] There also appears to be fundamental differences in the subcellular location of pre-tRNA splicing among different organisms.[9]

Removal of introns from pre-tRNAs in eukaryotes and archaea is performed by a protein-based endonuclease consisting of three polypeptide subunits,[26,27] whereas in bacteria and organelles intron removal proceeds via a group I self-splicing reaction.[9]

Removal of tRNA introns is essential in yeast and other fungi for tRNAs to decode codons. In some cases chemical modifications require the intron to first be removed but several examples are known where essential chemical modifications only occur on intron-containing pre-tRNAs.[25] The implication for these organisms is that both the introns

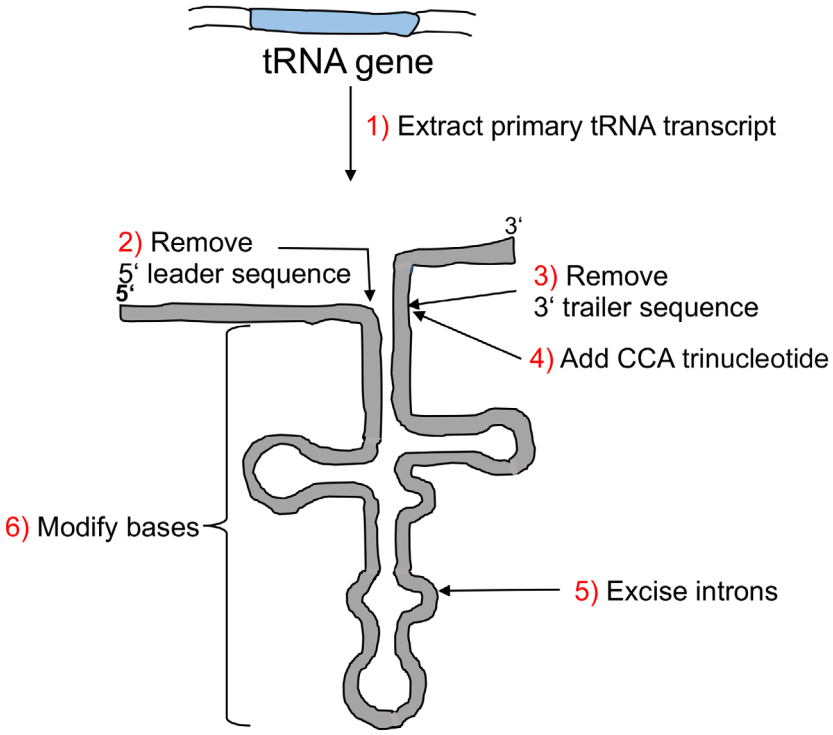

Figure 3. Transcription in eukaryotes produces a pre-tRNA that is cleaved on the 5′ end by RNase P and on the 3′ end by an endonuclease so that CCA can be attached. Many nucleotide modifications also occur. Many tRNA genes have introns which much be removed.

and the intron-removing enzymes were present from the very beginning.

The pre-tRNA splicing reaction occurs in three steps in yeast, involving several protein enzymes:[25]

1. Remove introns from pre-tRNAs. In yeast and vertebrates, the tRNA splicing endonuclease is composed of proteins Sen2, Sen34, Sen15, and Sen54. Sen15 and Sen54 sequences are very different between yeast and vertebrate cells and are absent from the archaeal genomes.
2. Ligate the 5’ and 3’ exons. This a very complex process catalyzed by tRNA ligase Trl1. The yeast mechanism is found also in plants. However, vertebrates and archaea ligate the tRNA halves directly by a 3’–5’ ligase activity.
3. Remove the residual phosphate at the splice junction. This is catalyzed by a phosphotransferase.

Thus, even though step one of tRNA splicing is similar for archaea, yeast, plants, and vertebrates, completion of the splicing reaction in yeast and plants requires two steps (5’–3’ ligation and 2’phosphotransferase), whereas in archaea and vertebrates, completion requires a single one-step ligation.[25]

Preparation of initiator aa-tRNA

Ribosomes must know where to begin translating mRNAs, and the reading frame as units of three nucleotides at a time must begin with the correct single nucleotide. Specially prepared initiator aminoacyl-tRNAs with a methionine attached are used for this purpose. In archaea and eukaryotes, the initiators are made directly by methionyl-tRNA synthetase, and then bind to initiation factors that deliver them to the ribosome. In bacteria, mitochondria, and chloroplasts the situation is completely different, since the Met-tRNA^{fMet} produced by the synthetase must first be formylated by a specific formyltransferase before it can be used to initiate protein synthesis.[22] The N-formyl group is removed later from mature proteins by peptide deformylase, leaving the canonical methionine as the mature N terminus.[22]

Conclusions

tRNAs are not merely sections of RNA to which activated amino acids are attached. Many protein-based molecular machines are necessary to process the pre-tRNAs and to biochemically transform tRNAs into usable form. The 5’ leader sequence must be removed using complex ribonucleases based on unrelated RNA and protein components across prokaryotes, mitochondria, and eukaryotes. RNase P is a protein-only enzyme (PRORP) in plants, a bad surprise for those believing in the RNA World theory. Removal of 3’ extensions from pre-tRNA is a biochemically complicated process, involving both exo- and endonucleases. If evolution is true, what came first, the tRNA genes or the genes needed to process the pre-tRNAs?

The alternative sources of the CCA tag are not predicted by a common ancestor scenario. Suppose the original organisms’ tRNA genes included the CCA sequence. What came first, the tRNA genes or the genes necessary to sever the 3’-end of the tRNA precursors at a precise position to expose the CCA tag as a single-strand of RNA? Since the tRNA genes of the archaea, some Gram-positive bacteria, the putative free-living ancestors of mitochondria, and the eukarya lack the CCA motif, were they all mutated away later? This process would have imposed a high cost in energy, material, and reproductive time and thus a selective disadvantage, since during the process of CCA loss the molecular equipment for both CCA-generating options would be needed. Once one scheme of such vital importance for survival worked it should not have changed.

Alternatively, suppose none of the genes contained the CCA initially. What came first, the tRNA genes lacking CCA, the genes to sever the tRNA precursors precisely at

nucleotide position 74, or the gene(s) to add CCA at the severed position?

The fact that bacteria such as *Aquifex aeolicus*, *Deinococcus radiodurans*, and *Synechocystis sp*. don't use one enzyme to repair damaged CCA but have two unrelated enzymes, one to add a 'C' precisely at positions 74 and 75, and the other an 'A' at position 76[29] points to separate designs.

What is the most natural way to explain how adding a second CCA sequence at the 3'-end of tRNAs came about to serve as an information signal, a Boolean yes/no instruction whether to degrade that unstable tRNA? As intelligent agents we frequently use variables to represent objects and events to perform this kind of logic. Evolution would have no plan. All the components to process such logic would need to arise by chance in organisms which supposedly were working just fine before.

In Part 2 we continue by documenting how tRNAs cannot work without dozens of biochemical transformations of the nucleotides, which require protein-based enzymes to have been present from the very beginning.

References

1. Ivanov, V., Beniaminov, A., Mikheyev, A., and Minyat, E., A mechanism for stop codon recognition by the ribosome: a bioinformatic approach, *RNA* **7**:1683–1692, 2001.
2. Bertram, G., Innes, S., Minella, O., Richardson, J.P., and Stansfield, I., Endless possibilities: translation termination and stop codon recognition, *Microbiology* **147**:255–269, 2001.
3. Truman, R., The surprisingly complex tRNA subsystem: part 2—bio-chemical modifications, *J. Creation* **34**(3):87–94, 2020.
4. Truman, R., The surprisingly complex tRNA subsystem: part 3—quality control mechanisms, *J. Creation* **35**(1), submitted.
5. Truman, R., The surprisingly complex tRNA subsystem: part 4—tRNA fragments regulate processes, *J. Creation* **35**(1), submitted.
6. Truman, R., The surprisingly complex tRNA subsystem: part 5—evolutionary implausibility, *J. Creation* **35**(1), submitted.
7. Barciszewska, M.Z., Perrigue, P.M., and Barciszewski J., tRNA – the golden standard in molecular Biology, *Mol. BioSyst.* **12**(12):12–17, 2016.
8. Pan, T., Modifications and functional genomics of human transfer RNA, *Cell Res* **28**:395–404, 2018.
9. Hopper, A.K., Pai, D.A., and Engelke, D.R., Cellular dynamics of tRNAs and their genes, *FEBS Letters* **584**:310–317, 2010.
10. *Isoacceptors* are tRNAs having different anticodons but are charged with the same amino acid. (Anticodons are the three nucleotides on tRNAs which bind to an mRNA codon). Isoacceptors are used to translate synonymous codons which represent the same amino acid in the genetic code.
11. Sharp, S.J., Schaack, J., Cooley, L., Burke, J.D., and Soil, D., Structure and transcription of eukaryotic tRNA gene, *Critical Reviews in Biochemistry* **19**(2):107–144, 1985.
12. Shepherd, J., and Ibba, M., Bacterial transfer RNAs, *FEMS Microbiology Reviews* **39**(3):280–300, 2015.
13. Krutyhołowa, R., Zakrzewski, K., and Glatt, S., Charging the code—tRNA modification complexes, *Current Opinion in Structural Biology* **55**:138–146, 2019.
14. Rak, O., Dahan, O., and Pilpel, Y., Repertoires of tRNAs: the couplers of genomics and proteomics, *Annual Review of Cell and Developmental Biology* **34**(1):239–264, 2018.
15. Lowe, T.M., and Chan, P.P., tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes, *Nucleic Acids Res.* **44**:W54–W57, 2016.
16. Abe, T., Inokuchi, H., Yamada, Y., Muto, A., Iwasaki, Y., and Ikemura, T., tRNADB-CE: tRNA gene database well-timed in the era of big sequence data, *Frontiers in Genetics* **5**(114):1–7, 2014.
17. *Plasmodium falciparum* is a unicellular protozoan parasite that causes falciparum malaria in humans. The parasite is transmitted through the bite of an *Anopheles* mosquito.
18. Ng, C.S., Sinha, A., Aniweh, Y., Nah, Q., Babu, I.R., Gu, C., Chionh, Y.H., Dedon, P.C., and Preiser, P.R., tRNA epitranscriptomics and biased codon are linked to proteome expression in *Plasmodium falciparum*, *Mol. Syst. Biol.* **14**:e8009, 2018.
19. *Isodecoders* are tRNA genes having the same anticodon but different body sequences.
20. Torres, A.G., Reina, O., Attolini, C.S-O., and de Pouplana, L.R., Differential expression of human tRNA genes drives the abundance of tRNA-derived fragments, PNAS **116**(17):8451–8456, 2019.
21. Raina, M. and Ibba, M., tRNAs as regulators of biological processes, *Front Genet*. **5**:171, 2014.
22. Ibba, M., and Söll, D., Aminoacyl-tRNAs: setting the limits of the genetic code, *Genes & Dev.* **18**:731–738, 2004.
23. Personal communication with Dr Change Tan.
24. Hopper, A.K., and Nostramo, R.T., tRNA processing and subcellular trafficking proteins multitask in pathways for other RNAs, *Front. Genet.*, 20 February 2019.
25. Hopper, A.K., Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast *Saccharomyces cerevisiae*, *Genetics* **194**:43–67, 2013.
26. Reiner, R., Ben-Asouli, Y., Krilovetzky, I., and Jarrous, N., A role for the catalytic ribonucleoprotein RNase P in RNA polymerase III transcription, *Genes Dev*. **20**(12):1621–35, 2006.
27. Grosjean, H., *DNA and RNA Modification Enzymes: Structure, mechanism, function and evolution,* CRC Press, Taylor & Francis Group, 2018. See chapter 23.
28. *Endonucleases* are enzymes that cleave the phosphodiester bond within a polynucleotide chain in the middle (endo) of the recognition sequences. Exonucleases cleave the ends of recognition sequences.
29. *Exonucleases* are enzymes that cleave nucleotides one at a time from the 3' or the 5' end ('exo') of a polynucleotide chain.
30. Hou, Y.M., CCA addition to tRNA: implications for tRNA quality control, *IUBMB Life* **62**(4):251–260, 2010.
31. Betat, H., and Mörl, M., The CCA-adding enzyme: a central scrutinizer in tRNA quality control, *Bioessays* **37**(9):975–982, 2015.
32. Wilusz, J.E., Whipple, J.M., Phizicky, E.M., and Sharp, P.A., tRNAs marked with CCACCA are targeted for degradation, *Science* **334**(6057):817–821, 2011.
33. Wellner, K., Betat, H., and Mörl, M., A tRNA's fate is decided at its 3' end: collaborative actions of CCA-adding enzyme and RNases involved in tRNA processing and degradation, *Biochimica et Biophysica Acta (BBA)—Gene Regulatory Mechanisms* **1861**(4):433–441, 2018.

Royal Truman *has bachelor's degrees in chemistry and in computer science from State University of New York; an MBA from the University of Michigan (Ann Arbor); a Ph.D. in organic chemistry from Michigan State University; and a two-year post-graduate 'Fortbildung' in bioinformatics from the Universities of Mannheim and Heidelberg. He works in Germany for a European- based multinational.*